UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Information content of performance measures

Jiang, N.

Link to publication

Creative Commons License (see https://creativecommons.org/use-remix/cc-licenses):
Other

Citation for published version (APA):
Jiang, N. (2020). *Information content of performance measures*.

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)

Download date: 20 Oct 2020

Nan Jiang

Information Content of Performance Measures

Information Content of Performance Measures

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. ir. K.I.J. Maex

ten overstaan van een door het College voor Promoties ingestelde commissie,

in het openbaar te verdedigen in de Aula der Universiteit

op woensdag 14 oktober 2020, te 16.00 uur

door Nan Jiang

geboren te Shandong

Promotiecommissie:

Promotor:	prof. dr. J.F.M.G. Bouwens	Universiteit van Amsterdam
Copromotor:	prof. dr. P. Kroos	Universiteit van Amsterdam
Overige leden:	dr. P. Casas-Arce	Arizona State University
	dr. M. Epure	Universitat Pompeu Fabra
	prof. dr. S.M. Morais Lourenço	Universidade de Lisboa
	prof. dr. V.S. Maas	Universiteit van Amsterdam
	prof. dr. D. Veenman	Universiteit van Amsterdam
	prof. dr. F.H.M. Verbeeten	Universiteit van Amsterdam

Faculteit Economie en Bedrijfskunde

Acknowledgement

I have wondered a couple of times throughout my PhD years what am I going to write in my thesis acknowledgement, but when the moment finally approaches, I deeply realize that words can only do so much as to express one's sentiments. I really had to try to keep this section at a reasonable length, because if I would name everybody that has ever taught me, guided me, and inspired me over the years, I run the risk of making this longer (and perhaps even more interesting) than the thesis itself.

My largest debt of gratitude is to my supervisor Jan Bouwens. Jan has granted me incredible degrees of freedom in exploring what I want to do, and how I want to do it; but whenever I'm stuck somewhere, he always manages to inspire me and guide me through the dark with his abundant experience and genius insight. He has introduced me to the fascinating world of doing field research on management accounting topics, where I find my passion truly resides. He is always supportive rather than restrictive of the decisions I make, and gives me the advice in my best interest. At times I feel that he knows me better than I know myself. Besides supervision, Jan also sets an example of the kind of scholar that I'm always longing to become: he has genuine curiosity for the world around him, he never lets himself stay in the comfort zone, and has never ceased exploring whatever challenges him intellectually as well as benefits the society. Jan is a respectable supervisor, a knowledgeable mentor, a dedicated colleague, and a candid friend.

I am also highly indebted to my co-supervisor Peter Kroos, this thesis could not have been in the current shape without his very insightful and constructive suggestions. I have greatly benefited from his approach to research, he has critical understanding of management accounting research, and, at the same time I learned a lot from him in terms of packaging and positioning my research, doing presentations, and writing in the way that appeals to the readership of academic journals. The practical advice that he gave me has gradually equipped me with the mentality of establishing myself not only as a PhD student but also as a junior academic. It is great fun working with Peter, what I particularly admire about him is that even when going through tough times, his passionate and contagious laughter never fades one bit. Peter, I sure will miss your laughter and the (sometimes lame) jokes we shared!

I'm also deeply appreciative to the other members of my doctoral committee, Pablo Casas-Arce, Mircea Epure, Sofia Lourenço, Victor Maas, David Veenman, and Frank Verbeeten, my thesis has greatly benefited from their thoughtful evaluation and feedback at various stages of the manuscript. Besides, their approach to research and my conversations with them over the years have helped to sharpen my thoughts and inspired me to think deeper from different perspectives.

I would also love to thank everyone, especially my fellow PhD colleagues at the Accounting Department at University of Amsterdam. I cannot say enough how much I enjoyed the caring and supportive atmosphere in the department. I enjoyed all the ideas we shared, the lunches and drinks we had, the trips we went on together, and many others. I would especially like to give a huge credit to the supporting team at UvA, thanks for making my life at UvA smooth and trouble-less. I could imagine spending my PhD years in no better way, the PhD life at UvA has been everything that I could ever ask for.

I greatly appreciate the generous financial support from Limperg Institute, their sponsorship made possible my research visit at Arizona State University, which turned out to be a unique and treasurable experience. My heartful thanks and appreciation goes to my two mentors at ASU: Michal Matějka and Pablo Casas-Arce. They have always been available and responsive whenever I feel that I need their help, and they have kept offering invaluable advice to my research and career. I have enormously benefited from my talks with Michal, I get to deepen my understanding about management accounting research from a refreshing new perspective, and he also encouraged me to be aware and learn to communicate my research to a broader audience. Pablo is incredibly intelligent and humble at the same time. His experience and talent in field studies give him a sharp angle into research, and his comments are always very constructive and spot-on. It is such an enjoyment working with him. The interactions I had with all colleagues and fellow PhDs at ASU have broadened my understanding of accounting research and academia with their perspective from the other side of the Atlantic. In addition, their warm company and hands-on support have loaded me with the mental fortitude and motivation to keep delivering and enjoying the life where I'm kind of alone but certainly not alone in Mountain Standard Time Zone.

Over the years of my research master and PhD I have been fortunate enough to make a number of good friends that took on the same journey as me in pursuing their PhD, many of whom have already launched a successful initial step into their academic career. I enjoyed so much the talks I had with every one of you about research

and well beyond, and your examples of perseverance and productivity deeply encouraged me through hard times. We share each other's joy and struggle. Knowing you are there for me means a lot to me, despite the physical distance for most of the time. My journey through the PhD would not have been as fun without you.

No words are ever enough to describe my gratitude to the family I'm blessed to have. Being the only child to my parents, it isn't always easy for them as I am on the other end of the continent for most of the time, but they have never once in my life override their wish on top of how I want to live my life. They raise me like they fly a kite, they let me freely explore what the sky has to offer, while always hold the string firmly in their hand, so that I always know I have them at my back. Mom, Dad, Grandma, Grandpa, aunts and uncles, thank you for your unconditional love and support for making me what I am. You are my heroes, my biggest fortune in life.

In Chinese we say 三十而立, that one establish herself as an independent, standalone person at the age of thirty. For me, the thirty age mark coincides with my date of defense, and has therefore a dual meaning: upon finishing the PhD, I will also gradually establish myself as a standalone researcher. All the wisdom, support, and kindness that I have been blessed with throughout the years, I will do my best to embrace and pass along.

Nan Jiang

September 2020

List of Co-authors

Chapter 2, entitled *Adding controls: Do junior managers respond the same as senior managers?* is joint work with my promotor Jan Bouwens. Jan have been involved in the data collection process as we use proprietary data from practice, and he has also been involved in developing the theoretical framework and positioning the paper. I have been responsible in carrying out the empirical analysis and interpretation of the results.

Chapter 4, entitled *Seniority and Use of Subjectivity in Performance Evaluation* is joint work with my promotor Jan Bouwens. Similar as Chapter 2, Jan have helped with data collection process and theory development. I have been responsible for the empirical analysis and positioning the paper as in the current shape.

Contents

Chapter 1 Introduction 3

Chapter 2 Adding controls: Do junior managers respond the same as senior managers? 7

2.1 Introduction 7

2.2 Literature and Hypothesis 10

2.3 Sample and Research Setting 13

2.3.1 Research setting Description 13

2.3.2 Survey and interview evidence from project managers 16

2.3.3 Regression Model Description 17

2.4 Empirical Results 19

2.4.1 Descriptive Statistics 19

2.4.2 Regression Results 21

2.5 Discussion and Additional Analysis 24

2.5.1 Discussion of empirical results 24

2.5.2 Alternative explanations 26

2.5.3 Robustness checks of empirical results 28

2.6 Conclusion 31

Chapter 3 Performance Appraisal on Employee's Future Potential 47

3.1 Introduction 47

3.2 Hypothesis Development 51

3.3 Research Setting and Data Description 54

3.4 Main Tests and Results 57

3.4.1 Change in Patterns of Supervisor Appraisal 57

3.4.2 Association of Supervisor Appraisal with Objective Performance Measures 59

3.4.3 Supervisor Appraisal in Predicting Future Performance 60

3.4.4 Firm's Use of Supervisor Appraisal in Making Promotion Decisions 62

3.4.5 Discussion of Empirical Results 64

3.5 Additional Analysis and Robustness Checks 66

3.5.1 Effect of Supervisor Appraisal on Employee Career and Compensation 66

3.5.2 Alternative Purpose of Supervisor Appraisal 67

3.5.3 Alternative Measure of Supervisor Appraisal 67

3.5.4 Effect Over Time and Cross-sectional Evidence 68

3.6 Conclusion 69

Chapter 4 Seniority and Use of Subjectivity in Performance Evaluation 89

4.1 Introduction 89
4.2 Theory and Literature Review 92
4.3 Data and Research Settings 95
4.4 Empirical Models and Results 97
4.4.1 Regression Methods 98
4.4.2 Analysis and Result 98
4.5 Additional Analysis and Discussion 100
4.5.1 Subjective compensation in predicting future performance 100
4.5.2 Explicit performance measure to in picking up agent's effort 102
4.5.3 Nonlinear relationship between subjectivity and tenure 103
4.6 Conclusion and Extensions 104
Summary in English 117
Summary in Dutch 118

Chapter 1 Introduction

My dissertation aims at advancing our understanding of organizational use of performance measures and how it enhances management control systems. Performance measures, along with incentive systems and delegation of decision rights, are the three fundamental and highly interdependent elements of the management control system in an organization (Abernethy, Bouwens, and van Lent 2004, Brickley, Smith, and Zimmerman 2015). The choice and use of performance measures plays a central role in achieving the optimal organizational performance. According to agency theory, the principal aims at properly motivating the agent to exert effort in the best interest of the principal. Since effort is unobserved by the principal, performance measures are selected such that they are informative to the principal about the agent's choice of effort (Holmstrom 1979). However, the role played by performance measures extend well beyond reflecting the agent's current effort choice. In the following three chapters of my dissertation, I conduct my research using data and managerial practice from the field, and through the lens of economic theory to shed light on a broader understanding of the information contained in performance measures.

Company's use of performance measures can contain information about the desired direction of effort that the principal expect the agent to exert. Chapter 2, coauthored with Jan Bouwens, is entitled *Adding controls: Do junior managers respond the same as senior managers?* In this chapter, we examine how project managers in a construction firm respond to the introduction of an additional performance measure on working capital. We put forward the idea that introducing the performance measure on a specific aspect of the employee's performance, on top of the general performance measure (profit) already in place, can serve as directing employee's effort to the direction desired by the firm, while it may not work in a uniform way for all employees. We find the level of working capital significantly decreased after the introduction of this policy, indicating that the introduction of performance measure on working capital is effective. Meanwhile, we further find that the outcome is mainly driven by the performance from junior project managers, senior managers however, rather than decreasing their working capital, show instead a slight increase in their working capital levels. The finding lends support to our prediction, that for less experienced agents, the introduction of a specific performance measure emphasizes the principal's intended direction of exerting effort. However, for more experienced agents, who have already figured out different domains of the job and

are better aware of balancing their effort, the introduction of this performance measure infringes the senior agents authority of doing things "in their own way"; the implicit understanding between the principal and agent is disrupted, and as a result the more senior agents are resistant to follow the policy.

Apart from directing and recognizing employee's effort, performance measures also serve the role of identifying and gauging employee's overall value of human capital. Performance measures, if used properly, can contain information about the extent to which principal distinguishes her assessment of the agents' current performance from their future potential. In a dynamic and comprehensive working environment in particular, it is crucial for companies to recognize their employees' strength and weakness, and identify employees with future potential. Such practice requires reliable indicators with a forward looking nature, where standalone, objective performance measure can fall short for this purpose, while subjective performance measure may suffer from bias and imprecision (e.g., Moers 2005, Bol 2008, 2011). In Chapter 3 entitled *Performance Appraisal on Employee's Future Potential*, I identify a situation where performance appraisal on employee's overall competence becomes a more precise metric in delegating supervisor's assessment of employee's future potential. The idea I put forward is that performance appraisal is more informative about employee's future potential in the presence of other form of reward for his past performance. I find in a professional service firm, after the firm introduced a performance based bonus to its employees, performance appraisal provided by the supervisor becomes more predictive of the employee's future performance, and also matters to a greater extent in explaining employee's chance of promotion. The findings suggest that without a performance based bonus, the supervisor has to put together both her recognition of employee's past performance as well as her assessment of the employee's on-going contribution in the performance appraisal, and the latter thus conveys a blurred message of the employee's overall competence. The performance based bonus grants the supervisor a better instrument to represent her assessment of employee's future potential more precisely, without compromising the effective recognition of employee's past performance.

Although companies choose performance measures to reflect employee's effort choice that adds value to the company, such measures contain information about an employee's overall contribution only to a limited extent. Performance measures used for evaluation are usually specified at the inception of the (written) contract and before any performance outcome is realized. However, the pre-set performance measures can be

imperfect in capturing employee's overall performance, especially in a comprehensive working environment. Chapter 4 is coauthored with Jan Bouwens and entitled *Seniority and Use of Subjectivity in Performance Evaluation*. We study in this chapter where a firm applies a universal bonus formula to all its employees, and we look into the question that when the pre-set formula cannot be tailored to each individual employee, what alternative way the firm can resort to in order to better recognize each employee's contribution. We find that the firm makes ex post adjustment to the bonus predicted by the pre-set formula, and the extent of adjustment increases with the employees' seniority as represented by their tenure in the firm. This is the case as seasoned employees gradually evolve into a relational contract with the principal, and they are more likely to engage in value-adding activities beyond the pre-set performance measures, the performance measure is thus imperfect in capturing their overall contribution. In provision of this, the principal effectively recognizes the employee's effort that resides out of what is specified in the formal contract by compensating his effort subjectively. We extend previous literature on use of subjective performance evaluation, that the use of subjectivity can increase the congruence of an incentive contract.

Overall, the three chapters offer innovative insight in organizational use of performance measures, both in theory and in practice. Chapter 2 provides theoretical and empirical evidence that by choosing performance measures, the principal communicates to the agents what direction of effort is expected from them. Economic and organizational theories suggest that employees are different in terms of their experience and relationship with their principal (e.g., Aghion and Tirole 1997, Benabou and Tirole 2003, Gibbons and Henderson 2012), Chapters 2 and 4 speak to the importance of taking into account the heterogeneity of the agents when applying performance measures to them. Moreover, identifying employee's overall competence and gauging their human capital is becoming a focal topic for many organizations, Chapter 3 offers initial insight into an approach that increases the precision of such measures. The dissertation generally conveys the message that the when using performance measures, principal needs to understand who the agents are, and what they are capable of doing. Overall, these chapters shed light on the emerging topics in management accounting, and open up for more research opportunities.

References:

Abernethy, M. A., Bouwens, J., & Van Lent, L. (2004). Determinants of control system design in divisionalized firms. The Accounting Review, 79(3), 545-570.

Aghion, P., & Tirole, J. (1997). Formal and real authority in organizations. Journal of political economy, 105(1), 1-29.

Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. The review of economic studies, 70(3), 489-520.

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

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

Brickley, J. A., Smith, C. W. and Zimmerman, J. L., (2015), Managerial Economics and Organisational Architecture, Irwin McGraw Hill, Boston.

Gibbons, R., & Henderson, R. (2012). Relational contracts and organizational capabilities. Organization science, 23(5), 1350-1364.

Holmstrom, B. R.(1979). Moral hazard and observability. Bell Journal of Economics, 10(1), 74-91.

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

Chapter 2 Adding controls: Do junior managers respond the same as senior managers?

2.1 Introduction

We study, in a construction firm, whether and how project managers who are initially evaluated on their project profitability respond to the introduction of an additional performance measure: working capital levels. Our prediction is that project managers, conditional on their seniority, will respond differently to the new policy.

Principals who want to ensure their agents make the right decisions can take the position that they are using specific performance measures to show, in detail, whether agents took the desired course of action or can use measures as a summary of the decisions their agents made. If a firm uses detailed measures of performance, it implies to its agents the desired course of action. This requires the principal to know in advance what the best course of action is. Depending on the knowledge of the agent, using detailed measures can bring costs or benefits to the firm. Agents capable of taking appropriate courses of action may feel their choices are unduly limited. For them, the use of a summary performance measure would allow the discretion to do as they see fit, given the conditions faced by the firm. This would benefit the principal, assuming the agent is motivated to act in the interest of the firm. On the other hand, agents who are uncertain about the best courses of action may be helped by more detailed performance measures, which can guide their effort choices. Following the course of action described by the set of performance measures will be consistent as long as the conditions facing the firm have not changed. But if conditions change, the uncertain agent may not be able to adapt and may make decisions that, all else equal, differ from the those that would be made by the agent assessed according to the summary measure.

Our study takes issue with these two conditions. We study a construction firm that initially evaluated its project managers on a summary measure: the profitability of their project portfolio. However, shareholders forced the firm to reduce its working capital. Top management responded with the introduction of a working capital measure, on top of the profitability measure. Project managers were instructed to reduce their working capital levels and were advised that this measure would be included in their performance evaluation.

We exploit this change, which is arguably a natural experiment, to assess how agents respond to the introduction of specific performance measures.

Based on theory, we expect that project managers will respond differently to the introduction of the working capital measure, conditional on their seniority. We predict that senior managers are more likely to consider the introduction of the additional measure as a restriction of their decision rights.

To enhance project profitability, project managers negotiate with clients to affect costs levels (e.g., change material specifications compared to the original contract to cut costs) and can push clients to pay their balances or can tolerate delayed payments. The better course of action depends on the specific conditions, and it is likely that agents acquire skills over time to evaluate potential courses of action. To the extent that profitability is the sole measure of performance, agents can decide whether to push clients to pay their balances or to curry favor with clients so as to reduce other costs, while de facto allowing clients to extend their credit term. The very introduction of an additional measure (i.e., working capital) curbs the agent's discretion in terms of selecting what the best course of action is when it comes to pursuing (project) profitability. We therefore expect that agents who know the best course of action will be less tolerant to the use of these controls than agents who have less experience and fewer skills.

We study how project managers respond to this infringement of their decision rights. We predict that senior managers will show their discontent with the restriction of their decision rights and that this sentiment will surface in the extent to which working capital improvements are achieved. For less experienced project managers, on the other hand, we expect that they are more willing to implement the new policy which will surface in the working-capital level improvement they achieve.

We find evidence that the firm's overall working capital levels decrease significantly after the introduction of the working capital performance measure, which suggests that, at the firm level, the new policy works. However, we observe a dispersion in the way the project managers react to the introduction. Before the new policy, senior managers (measured by their tenure in the firm) had outperformed junior managers in terms of managing their working capital levels and cash conversion cycle. Afterwards however, junior managers improve their levels of working capital, while this is hardly the case for the senior managers. These findings suggest that senior managers resist the instruction on how they should work.

We apply several robustness tests to examine whether the attribution points to resistance as the explanation for our findings. These tests suggest that the findings are *not* driven by the way the projects are allocated to managers, by how much scope different agents have to improve their performance, or by career concerns. The results lend further evidence that senior project managers consider the introduction of the specific performance measure an infringement of their authority. In addition, we also find evidence that the predicted dispersion (senior managers show a higher level of working capital after the policy) is not observed for their performance in profit, i.e., the association of tenure with realized revenue does not show a significant negative change after the policy. This evidence further suggests, consistent with our prediction, senior managers resist specifically the shareholder's introduction of the policy that effectively curbs their decision scope, rather than a "revenge" to the principal.

Our study contributes to literature in several ways. First, the literature provides evidence on how agents respond to controls. On the one hand, analytical (Frey, 1993), experimental (Falk and Kosfeld, 2006) and archival work (Campbell, Epstein, and Martinez Jerez, 2011) suggests that agents perform at lower levels if their principal imposes controls to monitor their actions. On the other hand, Nagin et al. (2002) find, in their field experiment, that call center employees seek opportunities to shirk when their actions are less likely to be monitored. It is therefore not clear when controls reduce or enhance performance. The findings of our paper suggests that the answer to that questions depends on the length of the relation of the agent with the firm.

Our second contribution pertains to the experience of agents. Campbell et al. (2011) show that experienced workers can better explore how to best direct their effort. They find that performance suffers if a principal curbs the decision rights of senior agents. This is not the case for junior agents facing equally low levels of the right to explore. Our work relates to this idea. However, unlike in the setting of Campbell et al. (2011), our agents face an unanticipated change in their right to explore how to direct their effort. This allows us to study how such a change affects the effort choice of senior versus junior employees. Related to this argument, Ittner and Larcker (2001) have suggested that research should factor in the multiple objectives performance measure have. We identify conditions where the same performance measures affect the motivation of the agent differently, conditional on their seniority.

In addition, our paper also contributes to the learning literature (Argote and Greve 2007). The conventional understanding is that experienced (senior) agents can adapt more quickly to a new situation than inexperienced (junior) agents. However, our study shows that experience is not a sufficient condition for adaptability; when experienced (senior) agents resist a change, adaptation is less likely to occur. Thus, experience can predict rigidity, not adaptability.

The remainder of the paper develops as follows. In section 2.2, we describe the theory and develop the hypothesis of our study and review the relevant literature. In section 2.3, we introduce our data and research setting as well as the empirical model to test the main hypothesis. We present empirical results and their interpretation in section 2.4. In section 2.5, we provide additional analyses to address alternative explanations. We conclude our study in section 2.6.

2.2 Literature and Hypothesis

Two agents who perform the same task may differ in their tolerance for control. Personalities aside, agents will differ in how well they have developed their skills and put them into practice. To the extent that agents have sufficient knowledge to independently make decisions, they would prefer that other people not interfere with their decision-making. Aghion and Tirole (1997) argue that principals who limit their agents' decision rights, despite the fact that these agents know what the best course of actions is, are likely to meet opposition in the form of effort reduction. Ceteris paribus, performance will deteriorate. Typically, application of specific performance measures entails decision right reduction, as they describe desired actions and norms that must be met. Merchant (1985, p. 29) refers to such specific measures as action controls. In their empirical study Abernethy, Bouwens and Van Lent (2004) show that decentralization is positively associated with the use of summary measures, i.e., measures that let the agent decide what the best course of action is.

Agents who have yet to develop the necessary skills are unlikely to resist the implementation of additional performance measures that more or less specify desired actions. In fact, the additional performance measures may help them select the right course of action and may help improve their skills. For instance, Dye (2004) argues that individuals can learn from performance outcomes recorded in a set of performance measures, provided that these measures vary cross-sectionally or over time. The introduction of a specific performance

measure in addition to a summary measure may elicit a similar effect. This idea is also consistent with the literature that takes issue with career concerns (Fama, 1980). According to this literature the principal infers the true state of the world from the messages sent by the agent. Agents in these models have strong incentives to exert early efforts and to build a reputation of a highly productive work (Harris and Holmström, 1982 and Gibbons and Murphy, 1992).

Skills, experience and tenure

Agents typically develop their skills as they accrue experience over their tenure. The literature provides ample evidence that agents with little experience have less knowledge than experienced agents. Hambrick and Fukutomi (1991) argue that a new external CEO lacks knowledge of the task, including facts, trends, contacts, and procedures that pertain to the successful conduct of the CEO's role in the firm but this disadvantage can be overcome with the increase in tenure. In addition, as agents serve their firm, their commitment to the CEO's vision is also more likely to become inculcated as tenure progresses. In a comprehensive meta-study, Sturman (2003) presents an abundance of evidence in support of the idea that organizational experience instills work-related skills and knowledge and that these accumulate with the agent's tenure. Moreover, the longer an agent works for a firm, the more likely it is that the agent understands what matters for the firm's long-term overall progress. Bonner (1990) and Choo and Trotman (1991) provide evidence to suggest that experienced auditors tend to have more task-specific knowledge and a better knowledge structure, which enhances information collection and supports better judgment. Casas-Arce, Martinez-Jerez, and Narayanan (2017) find that inexperienced agents are more likely to rely on new information than seasoned workers. They argue that these agents are more likely to rely on such information because of their lack of experience.

Tenure and control

In a setting where the firm aims to tap the knowledge of its agent and the agent has the necessary skills, it makes little sense to use specific performance measures that impair the potential for agent to fully exploit his knowledge. In fact, it is likely for seasoned agents to resent such specific controls, as their application suggests that the principal knows better than the agent what the best course of action is. For instance, Campbell et al. (2011) find, in their field study, that agents whose decisions rights are trimmed via specific

controls are less likely to make value-increasing decisions on how to best serve individual clients. This finding is consistent with the work of Aghion, Dewatripont and Stein (2008), who argue that agents may start to look out for other job opportunities when they are no longer free to make decisions suited to the situation they face. Dewatripont, Jewitt and Tirole (1999a, b) model a similar situation and argue that prescribed tasks may implicitly express insufficient appreciation for the decision-making talent of the agent.

In a similar vein, it is argued that specific performance measures may even negatively relate to the outcome measure (Dewatripont et al, 2000). For instance, input measures that restrict the use of raw materials may harm (the output measure of) profit. In such cases, agents are likely to reduce their effort to make decisions that would improve profit. Principals can restore the link between the existing knowledge level and measured performance if they remove the specific performance measure. Similarly, Dessein (2005) argues that the principal signals her belief to the agent inasmuch as she hands over control. Aghion and Tirole (1997) argue that agents who have the authority to make decisions on their own have incentives to collect and process information. Constraining these agents with specific performance measures could lead them to refrain from information collection, as such detailed controls limit the viable set of decision options.

Specific measures may also entail influence costs (Milgrom 1988). These costs arise because agents may try to convince their principal that the use of the specific measure does not apply or leads to inefficiencies. In the situation that the agent is right, influence costs distract the agents as well as the principal from activities that deserve attention.

However, specific performance measures are not necessarily bad. For instance, choice of performance measures helps the firm to communicate its objective to the employees (e.g., Abernethy, Dekker and Schulz 2015). Less skilled agents are more likely to accept the principal's restriction of decision-making, as they may want to use the mandated measures to help them direct their attention. To the extent that input measures yield desired levels of performance, agents may even be encouraged (Ittner, Larcker, and Randall 2003; Ittner, Larcker, and Meijer 2003; Kaplan and Norton 1992). This is consistent with the idea advanced by Benabou and Tirole (2003), who argue that agents are encouraged to do the right thing if they are helped in arriving at the desired decision. This idea is akin to the work of Holmstrom and Milgrom (1990, 1991), who show that better agents produce better results if responsibilities for a task are split. By introducing a specific

performance measure that tells the agent where to direct attention, the principal de facto imposes a split in responsibilities as the agent can rely on the selection the principal made in advance.

The above line of reasoning leads to the following prediction: As tenure progresses, it becomes less likely for agents to accept the introduction of additional specific performance measures. As agents accumulate skills, they arguably can better make the necessary trade-offs between different actions to support their decision. Junior agents may still have to develop their skills. We propose that senior agents with more experience have developed more skills than agents with short tenure. Specific performance measures that impair decision rights are therefore less likely to be accepted by senior agents than by junior agents. In addition, specific performance measures may help junior agents to direct their attention to the right course of action. That is, a specific performance measure functions like a predetermined choice to the agent, as such a control, in effect, operates as a norm or even an instruction to the agent (Merchant, 1985). This instruction helps agents because they do not need to tradeoff between the choice underlying the specific performance measure and other opportunities; the choice is already made (by the principal). The specific performance measure thus enables the junior agents to focus on the remaining choices they can make.

In line with the reasoning above, we express our expectation in the following hypothesis:

H1: Junior agents allocate their effort in response to the introduction of a specific performance measure in the implied direction, while senior agents are less likely to do so.

2.3 Sample and Research Setting

2.3.1 Research setting Description

The data and operational information for this study were obtained from a large construction firm in Europe, referred to as DC. DC offers a range of activities, including building, mechanical, and electronical construction projects offered to both governments and the private sector. The firm runs simultaneously thousands of projects that vary in scale and scope, where each project is assigned to a project manager. The project managers are responsible for the operational and financial results of their projects, and they are typically engineers holding professional degrees. Some of them have taken courses in accounting and finance. As engineer professionals, the responsibilities and decision scope of project managers are overall similar

within DC, regardless of their tenure within the firm. The internal accounting system keeps close track of each project's financial state, including estimated revenue, accumulated costs, and revenue realized by the time. To control its resource consumption, the firm monitors its cash conversion rate and its working capital. However, shareholders and analysts were critical of the working capital level of the firm. This led firm managers to conclude that project managers paid insufficient attention to (on time) cash collection.

To re-direct the attention of its project managers, the firm decided to impose on them a specific performance measure to accelerate the cash conversion cycle. The introduction of the specific performance measure was accompanied by a series of videos. In October 2014, the CEO released a video that called upon all project managers to step up their effort to decrease the working capital level and hence accelerate cash conversion. This video was immediately followed by another one relaying a training course that all project managers had to take. This video explained how project managers could improve their working capital levels and how to best approach the managers of the firms with whom they ran projects. In the meantime, the firm took several measures to make clear to the project managers that they should take working capital levels seriously. During their progress report meetings with financial controllers and supervisors, project managers were expected to be able to explain what they had done to improve their cash conversion rate. The firm appointed special "cash champion" who report to the COO, to make sure that cash conversion did improve.

It took the firm only two weeks to release the videos, organize the mandatory training, and implement the tighter supervision. In the weeks that followed, project managers did experience these tighter controls, with their supervisors and financial controllers ensuring that they strived to reduce their working capital levels. We refer to these actions as the introduction of the "specific performance measure."

As working capital had been highlighted so prominently, working capital levels had also been made a criterion in performance appraisals (final quarter 2014). The firm decided that it did not want to tie variable pay to working capital performance (nor did they tie project profitability to the variable pay levels). In the previous situation, project managers who reported favorable profit numbers over an extended period would—conditional on client satisfaction—be promoted to a higher fixed salary. This metric was now extended with the measure of working capital level. Hence performance appraisals regarding profit and working capital matter in the longer run.

As financial relations with suppliers are centralized, the major part of the working capital that project managers can manage is accounts receivable. They can do so by sending out invoices more frequently, for instance, with each milestone that is achieved in completing the project. This would lead them to send out invoices more frequently than that they had in the past. As each invoice has a fixed repayment term, project managers would then have to follow up if the client didn't pay promptly. Hence, besides invoicing more frequently, they also have to reach out to their clients more often.

In the spirit of our hypothesis, we expect that project managers respond differently to the introduction of the specific performance measure, conditional on their tenure. Our aims are to examine (1) to what extent the project managers respond to the specific measure and (2) the enforcement (project controllers and supervisors asking project manager to explain what they did to keep working capital levels low). In addition to project profitability, project managers' performance is indicated by their level of working capital and the speed with which they achieve cash conversion.

Data obtained from DC's internal accounting system allows us to measure several indicators of conversion efficiency: (percentage of) working capital, cash conversion cycle, lead time between when the project is started and the first invoice is released, and the (total) number of invoices issued to the client. The detailed definition of each performance indicator and how each is calculated is reproduced in Table 1.

Our original sample contains all individual projects DC undertook over the period of the first quarter in 2013 to the last quarter in 2015. In the internal accounting system, the first data entry for each project shows the estimated revenue. We use this number to proxy for project size. Some revenue entries show a zero or a negative number for administrative reasons (projects that had been closed in the past but now require re-work for which the client cannot be charged) or pure error. After consultation with the firm, we eliminated zero and negative revenue projects from the sample. In addition, we also eliminated a few observations where realized revenue was missing to ensure the project information was valid. We collected working capital data, invoice data, and data that allowed us to calculate the number of days that the invoice was outstanding. In addition to financial information for each project, we also collected data of the individual project managers who run the projects: their tenure, age, and information on how many projects a manager runs and whether the project is government project (measure of complexity). Due to a system upgrade by the end of 2013,

some projects are assigned an invalid project manager identifier. We keep our sample to the observations where we can credibly track to an individual project manager. This resulted in a final sample of 28551 valid project-quarter observations, with 92 unique project managers. Because of the system upgrade, valid observations with project manager information in 2013 are relatively small, compared with those in 2014 and 2015, with the number of observations in each year being 1236, 19218, and 8097, respectively.

2.3.2 Survey and interview evidence from project managers

In addition to the project related data obtained from the firm's internal control system, we also gathered information from a survey of a subsample of project managers to support our empirical findings. With the help of the financial controller of DC, we managed to have survey evidence from approximately 20 project managers at the end of year 2015 about their job tasks and interaction with their supervisor. We believe it is a representative sample of all project managers in our empirical analysis (n=92), the tenure of project managers who take the survey ranges from 1 to 25 years, with a mean of 8.4 years, comparable to the full sample.

Although the limited number of respondents restricts us from drawing statistical conclusion from the survey, we do observe a strong correlation between the tenure of project manager and their likelihood of being proactive and taking the initiative (e.g., agree to a greater extent to questions like "I take the initiative to discuss the progress of the projects" and "When it seems that we do not follow the schedule, I make sure that we do get back to the original schedule"). This lends support to our main assumption in the paper, that more experienced managers employs authority in their work. In addition, we also observe that longer tenured project managers tend to be less satisfied with their supervisors (e.g., agree to a lesser extent to questions such as "treat project managers as his / her equal" or "has no 'hidden agenda'") and unhappy with the way supervisors involve standard procedures in their work (e.g., supervisors of longer tenured project managers are perceived to be more likely to "requires project managers to follow standard procedures"). This further supports the idea we put forward that more experienced project managers appreciate the authority in getting their work done. Considering that the survey is taken one year after the change in working capital policy, senior manager's negative attitude could be due to their reported increase in the frequency of meeting their supervisors.

Overall, evidence from the survey confirms that the introduction of the policy change in working capital does have an effect on project manager's daily work, mainly with supervisor's closer monitoring of project manager's performance in terms working capital level. Meanwhile, evidence from the survey provides further insight that senior project managers are more reluctant to accept the additional performance measure on working capital as opposed to their junior peers, besides the empirical evidence we will present later in the paper.

2.3.3 Regression Model Description

Recall that the introduction of the specific performance measure of working capital reduction was issued early October 2014. Project managers were also informed that, starting from that quarter onward, their performance in working capital management would be part of the performance appraisal. As the shareholders prevailed upon the firm to reduce working capital, the executive team had little choice but to follow up with measures that would yield real improvements. Hence, as far as the project managers were concerned, the introduction of the specific performance measure arguably constituted an exogenous shock. In our setting, all project managers were exposed to the introduction of the same specific performance measure. As a consequence, we cannot run an outright difference-in-difference test. We instead consider the study a natural experiment. The data does allow us to study the responses of different agents to the introduction of the working capital measure. That is, we predict and examine whether, conditional on their tenure, agents respond differently to the introduction of the specific performance measure of working capital. We test this expectation with the following regression model:

$$WC\%_{i,t}=\alpha+\beta1*Tenure_{i,t}+\beta2*After_{t}+\beta3*Tenure*After_{i,t}+\gamma*Controls+\epsilon_{i,t} \quad . \qquad (1)$$

Our dependent variable *WC%* in equation 1 represents the working capital level, scaled by the estimated (budgeted) revenue of the project, and the suffixes of *i* and *t* represent each individual project, and finally *t* represents one of the 12 quarters over the years 2013 to 2015. We conduct our empirical analysis with observations at project-quarter level, as the project managers are evaluated each quarter, based on the performance of every project each project manager is responsible for. In our robustness tests, we use other measures of working capital efficiency as alternative dependent variables.

Tenure is measured as the total time that has elapsed between the year the project manager joined DC and the end of 2015. We examine the effect of project manager's tenure on the level of working capital of the projects they manage, which resembles taking the partial derivative of *WC%* by *Tenure*. The coefficient $\beta 1$ captures any level differences of the trend of working capital management in the period before the specific performance measure was introduced, when comparing the short-tenure and long-tenure groups. With their experience accumulated over time, longer-tenure (senior) project managers understand that a lower working capital level benefits the firm, and they are likely to have already worked on managing working capital levels even before the firm points it out. Therefore, we expect that longer-tenure (senior) project managers will be better at managing working capital than their junior counterparts. We expect the sign of $\beta 1$ to be negative (i.e., WC% is lower when project manager's tenure is longer).

After is a time indicator that takes on the value of 1 for observations later than the third quarter of 2014 and a value of 0 otherwise. We took the fourth quarter of 2014 as the event date, as the policy change and the training program were all organized at the beginning of that quarter and the introduction of the new working capital control took immediate effect. The coefficient on *After*, $\beta 2$, captures the overall change in outstanding working capital before and after the introduction of the specific performance measure. We expect the sign of the coefficient on *After* to be negative (i.e., after the introduction of the policy, WC takes on lower levels than before).

The major coefficient of interest is $\beta 3$, which captures the incremental (or subtractive) sensitivity of the project's outstanding working capital to its manager's tenure. According to our hypothesis, senior managers may resist the introduction of the metric of working capital level, as it constitutes an infringement of their decision rights. We therefore expect the magnitude of the negative sensitivity of working capital on project manager's tenure to decrease, in other words, we expect $\beta 1$ is negative and $\beta 3$ is positive .

In addition to *Tenure*, we include other variables that may affect levels of working capital. We include (the overall) *Size* of the project to control for factors other than those reflected in size that affect working capital levels (e.g. a strategic project, which may entail higher working capital levels than projects that are of a less strategic nature). We control for the number of projects, *#proj*, calculated as the total number of projects one project manager is responsible for at the current quarter, to account for the potential that project managers

may face descending levels of control over their working capital as the number of projects rises. *Age* is the project manager's age, which reflects the project manager's personal characteristics that may affect working capital levels. In addition, *Age* is arguably related to career potential (high for juniors, low for seniors) and controlling for this factor ensures that it is less likely to explain our main results. The variable *Gov* indicates whether the project is a government project and is included to capture the complexity of the project. From our interviews, we learned that government projects are complex, as they feature much red tape in terms of filing for permits or dealing with delays when citizens challenge permits in court.

2.4 Empirical Results

2.4.1 Descriptive Statistics

We present the descriptive statistics of dependent, independent, and control variables in Table 1. Panel A listed a few indicators of the conversion efficiency of the projects. The (absolute) amount of working capital has an average of 2690.93, with a relatively large dispersion (standard deviation equals 71380), which comes down to an average of 8.43% of the total revenue of the project. The projects have an average cash conversion cycle of a bit less than a month (26.97 days), and this cycle varies noticeably (standard deviation 99.35). The lead time between the date that the project commences and when the first invoice is issued amounts on average to 1.32 quarters, while for most observations, the lead time is one quarter (p75=1). The number of invoices sent to the client in each quarter is on average 0.78. The total number of invoices sent for each project per year has a mean of 1.65 and standard deviation of 9.08.

(insert Table 1 about here)

Table 1 Panel A and the above interpretations give a general description of the conversion efficiency condition of DC. As the firm presents working capital reductions as a strategic priority, we expect that project managers will on average try to improve their working capital levels. Hence, on average, working capital is expected to decrease after the policy change. We divide the sample into two parts — the observations before the fourth quarter of 2014, when the policy change was implemented, and those of later periods. We present the average values of the different cash conversion efficiency indicators of the two subsamples in Table 2. Note that the number of observations before and after the shock is more or less

comparable (around 14000 observations before and after). Consistent with the aim of the firm, the working capital measures indicate a reduction in working capital, following the policy change, with the working capital reduced from 4106.8 to 1861.5 and its percentage from 13.38 to 4.08. Moreover, the cash conversion cycle is reduced from 40.24 days to 15.11 days, lead time reduced from 1.35 quarters to 1.28 quarters, and both of the variables of the number of invoices sent to the client increased. All variables change significantly after the policy change, evaluated at 99% confidence level, in the direction consistent with the intended outcome that cash conversion efficiency increases. The change in the average cash conversion efficiency rates suggest that the firm was successful in implementing its working capital policy.

(insert Table 2 about here)

In Table 1 Panel B, we reproduce the descriptive statistics of the project's demographic variables. There are 92 unique project managers in our sample, whose tenure in DC range from two years to 42 years, with an average tenure of 14.49 years (median 10 years). The youngest project manager is 26.9 years old and the oldest 71.4, the average age of the project managers being 47.97 years old. The number of projects each manager is responsible for per quarter is summarized at manager-quarter level. On average, each project manager is responsible for 59.61 projects per quarter, and the distribution is largely right-skewed: Half of the managers take charge of 11 or less projects per quarter, while the maximum can be 1779 projects. The reason for the large number is that projects are described on a units basis, each representing a separate contract. This is consistent with what we observe in project size: the average project size equals estimated revenue of 49577, while most of the projects are much smaller in scale (p25=430) than the larger ones (p75=7247.9, smaller than 49577). The mean of *Gov* indicates that there is only a small fraction (2%) of projects operated for the government. For analysis later on, *Tenure, Age* and *#proj* denote the manager characteristics who runs each specific project, and all variables are at project-quarter level.

The correlation of the above variables is shown in Table 3. The conversion efficiency variables are correlated with each other (the correlations between working capital and working capital percentage and cash conversion cycle are 0.09 and 0.18 respectively, both significant at the 99% confidence level), lending support that these variables are reliable indicators of project conversion efficiency. Note that *Tenure* is significantly negatively correlated with working capital (-0.02, p=0.00) and with cash conversion cycle (-

0.03, p=0.00), which provides primary evidence that more senior project managers are in general performing better in terms of conversion management. *Tenure* correlates positively with likelihood to take on governance projects (0.22, p=0.00) and negatively with the number of projects managed at the current quarter (-0.15, p=0.00), which suggests senior project managers are more likely to be responsible for more complex and integrated projects.

(insert Table 3 about here)

2.4.2 Regression Results

In this section, we summarize the results of our tests. Our main results are reproduced in Table 4. We estimate several specifications of the model, where in all cases *WC%* (working capital scaled by budgeted project size) is the dependent variable. We cluster the standard errors by individual project to account for autocorrelation over time. In addition, across all specifications we include quarter fixed effect to take into account the seasonal variations due to the construction firm's nature of business. Column (1) presents the result of the baseline regression model with control variables described in the previous section. We examined the variance inflation factor (VIF) to test for potential multicollinearity, and the statistics are well below conventional level to raise multicollinearity concerns. The coefficient of $\beta 1$ on *Tenure* is significantly negative (-0.231, p-value=0.00). This is consistent with the predicted sign on *Tenure* and lends empirical support to the idea that, before the introduction of the specific performance measure, managers with longer tenure had been better at controlling working capital levels than their more junior counterparts. The coefficient $\beta 2$ on the time indicator *After* is significantly negative (-15.10, p-value=0.00), which is consistent with our prediction. This evidence, along with the difference of the average working capital level before and after the introduction of the specific performance measure shown in Table 2, suggests that the introduction of the specific performance measure helps reduce working capital on an average basis.

(insert Table 4 about here)

To test our main hypothesis, we turn to the coefficient on the interaction term *Tenure*After*. The coefficient $\beta 3$ of *Tenure*After* is positive and significant (0.563, p-value=0.00). Note that the sign of the coefficient on interaction term switches, compared to the main effect reflected in the coefficient $\beta 1$ on *Tenure* (reflective of the situation before the policy change was implemented), and supports our hypothesis that managers with

longer tenure respond differently to the policy change than those with shorter tenure. The coefficients indicate that, while an additional year of experience will, on average, reduce a project's working capital percentage by about 0.23% before imposing the specific performance measure, the effect does not stay the same afterwards; that is, an additional year of experience will instead increase the project working capital percentage by 0.332% (-0.231+0.563), after the new policy takes effect. This implies that the introduction of the specific performance measure does lead managers to respond differently, conditional on tenure. Note that, *after* the introduction of the working capital metric, working capital decreased on an average basis ($\beta 2<0$). We use this result to estimate how many years of tenure offsets the average improvement. According to the estimated coefficients of the baseline model, the working capital improvements agents achieve over their tenure are neutralized for agents with a tenure longer than 26.82 years (derived from (-0.231)**Tenure*<0.332**Tenure*-15.10). In our sample, 17 out of the 92 project managers have a tenure in DC longer than 26.82 years. The statistical evidence is consistent with our hypothesis that the introduction of the specific performance measure led the senior agents to make less progress than junior colleagues, and for a subgroup of senior managers, their performance with regard to working capital percentage even deteriorates after the specific performance measure of working capital was introduced.

The coefficients on some of the control variables are worth noting. We observe that, on average, projects run for government accumulate 6.187% more working capital than private projects and that projects run by busier project managers entail slightly smaller working capital percentages (coefficient on *#proj* equals -0.005%, p-value=0.00). Note also that, with each incremental year of age, the working capital percentage goes down on average with 0.391% (at 99% confidence level).

We include the *Age* variable to account for the potential career concerns, and the significant coefficient validates the inclusion of age. We believe that career concerns of project managers with similar ages are comparable, and the fact that we find a significant result for the main variables of interest after controlling for *Age* implies that the we can attribute the findings to the project manager's tenure, instead of the possibility that senior managers are less concerned with working capital because it does not affect their career prospects. In our robustness tests, we conduct more sophisticated tests.

We show the results from alternative specifications in the next two columns of Table 4 to examine the robustness of our result. It is reasonable to believe that project manager's personal characteristics will also contribute to their ability to manage working capital and are not fully captured by the control variables such as *Age* and *#proj*, therefore, we run the baseline regression model with project fixed effect, and the result is presented in Table 4, column (2). Since we use the end of our sample period as the reference point to define *Tenure* and *Age*, these two variables stay unchanged for a specific project manager, and are therefore dropped out in the regression when project manager fixed effect is included. Including project manager fixed effect increases explanatory power of the regression (R-squared 11%) and the sign of most coefficients stays unchanged. The coefficient on our main variable of interest *Tenure*After* stays positive and significant (0.264, p-value=0.00), and is comparable with the results in column (1) in terms of economic significance (i.e., it takes now 6.171/0.264=23.37 years of tenure to offset the average increase in working capital percentage).This result further sheds light on our hypothesis that agents respond differently to the introduction of the specific performance measure of working capital conditional on their experience.

In Table 4, column (3), we tabulate the regression result including client fixed effect. The main action project managers may take to reduce working capital is issuing invoices more frequently to their clients and following up and urging their clients to pay. The concern that keeps managers from doing so is that clients may get annoyed and become less cooperative about other factors that determine project profitability. We run the regression including client fixed effect (two observations dropped because of missing client identifier), and we observe the result still holds for the main variables. This confirms that our result is indeed driven by managers' response to the introduction of the specific performance measure, rather than that managers would be concerned with any specific clients.

We run additional robustness analysis that we do not tabulate the results. To account for potential concern of extreme values, we winsorized the project size (i.e., replaced the smallest 1% size by the value of the 1% observation as well as the largest 99% by the value of the 99% observation) and run the regression again; the results show that winsorizing the sample leaves the statistical significance of coefficients unaltered. Similarly, it is reasonable to expect that project managers pays more attention and effort in management the working capital of the large projects they are responsible for, regression analysis for observations that are among the top 10% of project size exclusively provides consistent result. We cluster standard errors by time (year-

quarter), and the results are again consistent with our original results. In addition, to address the potential concern that projects operated for the government are systematically different in terms of cash conversion, we exclude observations which *Gov*=1 and the result still holds.

To summarize, the empirical evidence from our regression models lends support for our hypothesis that project managers with longer tenure differ from managers with shorter tenure in how they respond to the introduction of the specific performance measure. Where managers with shorter tenures improve their performance in cash collection after the introduction of the specific performance measure of working capital, those with longer tenures seem less willing to do so.

2.5 Discussion and Additional Analysis

2.5.1 Discussion of empirical results

Our results suggest that, conditional on the agent's tenure, that person responds differently to policy changes that affect his or her decision rights. Our theory predicts that senior agents, given their experience accumulated over time, are better positioned to make informed decisions and thus expect the principal to take their decision at face value.

Our finding is also consistent with what relational contract would predict. Gibbons and Henderson (2012) specify two distinct dimensions that determine these relations: credibility and clarity. Over time, clarity and credibility accumulate to the point where principals and agents enter into a steady state in which both parties know what to expect, and both the principal and agent benefit from their mutual understanding: the agent with multiple encounters with the principal chooses effort levels the principal desires (in this case, managing working capital levels), even when they are not explicitly specified, and the agent knows he or she will be compensated appropriately. The introduction of the specific performance measure represents principal's interruption of the agent's decision right, and resembles an infringement of the mutual understanding between the principal and the agent. Therefore, relational contract theory would predict that the agents would respond negatively to such an introduction of additional performance measure. The reason is that the action signals that the principal no longer believes the agent will make the right trade-offs. On the other hand, more junior agents may still have to learn to make the trade-offs and may therefore be less concerned about their

relationship with the principal in this situation. In fact, they may feel that the change helps them make better decisions. Our results lends support to these arguments. Senior agents start to perform worse after their decision rights are lessened while the juniors start to perform better.

To provide further evidence on the findings we described above, we present how working capital levels differ between agents who differ in their tenure before and after the introduction of the control. We divided the observations into senior and junior managers at the median number of *Tenure* (10 years).

In Table 5 we present the average working capital percentage of projects run by junior and senior managers, aggregated both before and after the control, respectively. We observe that projects led by senior managers perform significantly better in terms of WC% performance before the specific performance measure was introduced (2.19%) than those led by junior peers (18.18%, difference significant at 99% confidence level). This is in line with our prediction that senior managers have established the mutual understanding with the principal, that they incorporate working capital into the set of decisions they make without the principal explicitly asks for it, and in the meantime anticipate the principal to properly appreciate such decisions. After the working capital measure was introduced, however, junior managers perform significantly better than before (3.16%, t-stat=18.36), while senior managers on the contrary, report on average an even higher working capital percentage than they used to have (5.24%, t-stat=2.87). This implies that senior managers react to the control imposed on them differently than their junior peers. Senior managers who used to outperform their junior colleagues before the introduction of the specific performance measure of working capital now shows an adverse outcome. That is, the mean working capital percentage of projects led by senior managers (5.24%) is even higher than that of junior managers (3.16%, t-stat=-2.35). This evidence is consistent with the idea that the introduction of such specific performance measure on working capital resembles an infringement of the authority that senior managers used to be implicitly granted, and that senior managers seem to resist being "told what to do". In addition, by further separating the level of working capital percentage over each quarter, we see that working capital percentage of projects managed by senior managers continues to overtake the level of those managed by junior managers. Due to the limited time period after the event, we have to take cautions in interpreting such observation, yet the data suggests the adverse effect of performance measure on working capital to senior managers persist at least after a few quarters, this lends us further support that it is unlikely that senior managers see it more as an non-event.

(insert Table 5 about here)

2.5.2 Alternative explanations

Our empirical findings may be subject to alternative explanations, and we attempt to address them in this subsection.

Difference in project (re-)allocation. Our results may be due to project assignment in that junior managers get different projects assigned than senior managers, in terms of size or complexity, before and after the control was implemented. The different patterns we observe for senior and junior managers may also be associated with the possibility that, after the policy change, senior managers are (re-)allocated to projects for which it is inherently much harder to reduce working capital, while junior managers get easier projects.

To examine this concern, we first present the descriptive statistics for project size and complexity. From the results presented in Table 6 Panel A, we infer that projects that are run by senior managers are in general comparable in size (mean size for senior managers slightly larger than junior managers but difference is not significant but more complex (proxied by whether it is a government project). This validates the inclusion of project size and government project as control variables in our regression analysis. Furthermore, what matters is, for each subgroup, whether the projects assigned to individual managers are significantly different before and after the policy change. In Table 6 Panel B, we show the mean size and governance indicator of projects run by senior managers before and after the change as well as the descriptives for projects managed by junior managers. We observe that, the average size and type of projects assigned to junior managers does not show significant change after the policy takes place. For senior managers, projects assigned to them are similar in size (difference in mean insignificant) while are less likely to be government projects (0.053 before and 0.025 after, t-stat=7.39). This is due to the fact that after 2014 quarter four (the policy change), there are fewer projects from the government operated by the firm (272 government projects, or 1.94 in percentage before 2014Q4, and 200 government projects or 1.37 in percentage after 2014Q4). In other words, there is no evidence to believe that the way to allocate the projects are systematically different before and after the policy change. The evidence shows that it is unlikely that project reallocation accounts for our results.

(insert Table 6 about here)

Alternative performance measures. We also provide evidence that other performance outcomes, such as project revenue, do not show patterns of change similar to those we observe in working capital after the specific performance measure of working capital had been introduced. Realized revenue is the actual revenue the company realizes adjusted for any capital charges after the project is conducted, as opposed to budgeted revenue. We therefore re-run the regression presented in Table 4 but take realized revenue as the dependent variable. We show the result of specifications with and without project manager and client fixed effects, and we find insignificant coefficients on *Tenure*, *After*, and the interaction term *Tenure*After* (Table 7). The result re-assures that the senior project managers only show resistance with regard to the specific performance measure (working capital) recently introduced, but not that they also give up on what they have always been expected to deliver. Together with our descriptive statistics in Table 6, this additional analysis suggests that project managers did not get assigned different projects in terms of size. The finding is also interesting in the way that it suggests that the senior managers are particularly resistant to the policy change imposed by the shareholders, i.e., the senior agents respond against the party that introduced the specific performance measure on working capital, rather than posing a "revenge" to the principal by shirking.

It is worth noting that consistent with the argument we present to develop the hypothesis, that senior managers are better aware of the negative consequence of stepping up the cash collections process (e.g., the client may get irritated), one may expect that the effect of the interaction term *Tenure*After* on realized revenue in this analysis is positive. The fact that we do not find significant positive coefficient indicates that maintaining a better client relationship have a long-term effect on project revenue, and since we only have up to one year observations after the policy change, the observed effect is still yet to materialize.

(insert Table 7 about here)

Manager's career perspective. Career concerns might also explain our results. That is, responding more to the firm's expectations may benefit junior managers more than their senior colleagues, who are closer to retirement and thus less worried about career advancement. We address this concern by adding a career variable that equals to one for managers closer to retirement (i.e., *Age*>60), and the result on main variables remains unchanged (result untabulated). This evidence shows that managers' career concerns do not explain our findings.

Too good to improve? Another concern may be that senior managers are so good that they have already reached at working capital levels where they simply have less scope to further improve working capital levels. We provide evidence that this argument is unlikely to explain our results. First of all, the economic significance of the coefficients in our main analysis indicates that for a subgroup of senior managers, their working capital performance does not just stay the same but even deteriorates after the policy change. In Table 5, we also see that the mean WC% of senior managers after the new metric was introduced is significantly higher than that of their junior peers (5.24% versus 3.16%, difference t-stat=-2.35). This indicates that there is ample room to improve for senior managers.

To further shed light on the possibility to improve performance, we compare the relevant information in DC's annual reports and those of another construction firm in the same European country and listed on the same stock exchange. (Note that, in the annual reports, the working capital level and percentage is aggregated at firm level and is therefore not directly linkable to those at project level in our data.) We find that, for both firms, the percentage of working capital to noncurrent assets has decreased by 27%. (In both firms, working capital is negative, which means the percentage increased in absolute value.) But be aware that DC is three times as large in total revenue and market capitalization, so working capital performance should have been more efficient than it was. Using an industry peer as a benchmark hardly lends support for the belief that, for some managers, their working capital has reached the optimal level. In addition, in 2015, DC reported an achievement of -8.1% working capital percentage (with regard to total revenue) and set a target of -10% for the following year; it beat this target in 2016. This suggests that there exists room to further improve working capital levels for both junior and senior managers.

2.5.3 Robustness checks of empirical results

Alternative specifications. The empirical evidence presented so far is robust to a number of alternative specifications. We applied four alternative indicators of conversion efficiency as the dependent variable in addition to the WC% measure in our main regression: the natural logarithm of (1+WC%), the absolute amount of working capital (*WC*), cash conversion cycle (*CCC*), and the time that elapses between project initiation and releasing the first invoice (*Pending*). Control variables are included, standard errors are

clustered by projects and, for all four regressions, we use the initial (unwinsorized) sample. The regression results using the three alternative dependent variables are presented in Table 8.

(insert Table 8 about here)

As in the previous analyses, a lower amount of all three dependent variables reflects a better cash conversion rate. Hence the prediction of the sign of the coefficients on the three explanatory variables (*Tenure*, *After*, and *Tenure*After*) remains unchanged (negative, negative, and positive, respectively). The regression results are consistent across all specifications using the alternative measures of working capital percentage (*WC%*) presented in Table 4.

Where does the result become significant? A concern may be that the effect we describe is relevant for only a small group of our sample (e.g., project managers who are about to retire). To see whether that is true, we created dummy variables to indicate different levels of seniority. We find (untabulated) that the interaction between tenure and the event (after) start to become significant from seniority levels of 10 years and higher. This result does not suggest that it is just the manager who is close to retirement who resents the specific performance measure — indeed half of the population shows some level of resistance. The results also suggest that resistance against the introduction of the specific performance measure of working capital is likely the driving factor. That is, it is more the infringement of the decision rights of the more senior agents that drives our results: After 10 years or more, the agents believe the principal should expect them to do the right thing.

As the results on our interaction start to become significant at the median, we present these results in Table 9, using *WC%, WC,* and cash conversion cycle (*CCC*) as our dependent variables. The results are robust for all specifications, lending support for the idea that it is not just the agents close to retirement who resist the additional control.

(insert Table 9 about here)

***Assumptions of analysis*.** Our analysis is akin to the difference-in-difference method, which assumes that the common trend assumption is met. This assumption requires that all observations follow a similar trend, absent the exogenous shock. In other words, the only reason that individual observations differ is because of

the shock. We empirically examine the common trend assumption with a placebo test. The idea is to randomly pick an alternative point as a placebo "shock" before the actual policy change takes place, and re-run the regression. We re-run the regression assuming that the shock occurred in the first quarter of 2014, rather than the fourth quarter of 2014 (which is the actual point in time of the shock). The results are reproduced in Table 4, column (4). The coefficients on *Tenure* (0.220, p-value=0.15), After (3.194, p-value=0.24) and the interaction term *Tenure*After* (-0.233, p-value=0.15) turn out to be insignificant. The number of observations in column (4) is smaller because we restrict our sample period from 2013Q1 up until the actual policy change in 2014Q4. The placebo test further confirms that the parallel trend assumption holds and therefore lends support to our empirical evidence.

Cross sectional evidence with regard to client type. We explained earlier in the paper that the major action a project manager can take to lower the working capital level is to "chase after" their clients and inquire them to pay back accounts receivables on time. In further understanding this mechanism, we provide cross-sectional evidence conditional on types of clients each project is designated. The cross-sectional evidence of the main regression analysis is presented in Table 10. From columns (1) and (2) we can see that the observed effect is concentrated among non-government projects. This is due to the fact that projects operated for governments are billed with a pre-set schedule and procedure, there is limited slack the project managers could take among the government projects in order to respond to the policy change in managing working capital. In further discover evidence from client types, we find that among all projects, 12,740 has a billed party as other departments of the DC corporate (e.g., DC Water & Energy, DC Northwest Region, etc.), and 15,811 are projects billed for outside clients. We further conduct the analysis separately for projects billed to outside clients versus projects billed within DC, and present the result in columns (3) and (4). As we can see, the observed pattern (heterogeneous reaction among senior and junior managers) are concentrated among projects that belong to outside clients. This finding is consistent with our argument puts forward earlier, that senior managers understand the critical consequence of "chasing after" client's tale in long term performance, and thus they are less likely to take actions towards outside clients as opposed to within DC clients. In sum, the cross-sectional evidence conditional on types of clients entails further support to the mechanism that drives the heterogeneous reaction that we propose.

(insert Table 10 about here)

2.6 Conclusion

We employ a unique setting and document how a well-intended change in a control mechanism may be opposed by some agents. We argue and find that junior agents are more likely to respond positively to the introduction of a metric for working capital improvement than their senior colleagues. Our evidence potentially advances understanding of how a uniform implementation of a specific performance measure may produce different effects, depending on an agent's skills and relationship with the principal. That is, our analyses suggest that firms may want to be cautious with reducing the decision rights of senior employees, as they may resist such a change.

As with all studies, ours has limitations. We can only collect some data to control for the (personal) characteristics of the agents included in the study. We do not know, for instance, the individual managers' level of risk aversion. Our relatively short study period also restricts deeper understanding of the before and after effects of the control mechanism. That is, after our sample period, senior agents may have decided to give up their resistance. We leave it to future studies to examine how such factors may relate to the uses of control systems.

References

Abernethy, M.A., J. Bouwens and L. Van Lent (2004). Determinants of control system design in divisionalized firms, *The Accounting Review,* 79(3): 545-570.

Abernethy, M. A., Dekker, H. C., & Schulz, A. K. D. (2015). Are employee selection and incentive contracts complements or substitutes?. Journal of Accounting Research, 53(4), 633-668.

Aghion, P., and J. Tirole (1997). Formal and real authority in organizations. *Journal of Political Economy*, 105: 1–29.

Aghion, P., M. Dewatripont, and J. Stein (2008). Academic Freedom, Private-Sector Focus and the Process of Innovation. *RAND Journal of Economics*, 39: 617–635.

Argote, L., and H. Greve (2007). A Behavioral Theory of the Firm—40 Years and Counting: Introduction and Impact. *Organization Science*, 18(3): 337–349

Baker, G., R. Gibbons, and K.J. Murphy (2002). Relational Contracts and the Theory of the Firm. *Quarterly Journal of Economics*: 39–84.

Bénabou, R., and J.Tirole (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies*, 70: 489–520.

Bolton, P., and M. Dewatripont (2011). Authority in Organizations: A Survey. In R.
Gibbons, and D.J. Robert, eds., *Handbook of Organizational Economics Princeton*:
Princeton University Press.

Bonner, S. E. (1990). Experience effects in auditing: The role of task-specific knowledge. *The Accounting Review*: 72–92.

Campbell, D., M. J. Epstein, and F. A. Martinez-Jerez (2011). The learning effects of monitoring. *The Accounting Review*, 86.6: 1909–1934.

Casas-Arce, P., F.A. Martínez-Jerez, and V.G. Narayanan (2017). The Impact of Forward-Looking Metrics on Employee Decision Making: The Case of Customer Lifetime Value. *The Accounting Review*, 92(3): 31–56.

Choo, F., and K. T. Trotman (1991). The relationship between knowledge structure and judgments for experienced and inexperienced auditors. *The Accounting Review*: 464–485.

Dewatripont, M., I. Jewit, and J Tirole (1999a). The economics of career concerns, Part I: comparing information structures. *Review of Economic Studies*, 66: 183–198.

Dewatripont, M., I. Jewit, and J Tirole (1999b). The economics of career concerns, Part II: application to mission and accountability of government agencies. *Review of Economic Studies*, 66: 199–217.

Dewatripont, M., I. Jewitt and J. Tirole (2000). Multitask Agency Problems: Task Separation and Clustering. *European Economic Review*, 44: 869–877.

Dye, R.A. (2004). Strategy Selection and the Choice Among Performance Measures. *Management Science*, 50(12): 1624–1637.

Falk, A., and M. Kosfeld (2006). The Hidden Costs of Control. *The American Economic Review* 96(5): 1611–1630.

Fama, E. F,(1980). Agency Problems and the Theory of the Firm, *The journal of Political Economy*, 1980: 288–307.

Frey, B. (1993). Does monitoring increase work effort? The rivalry with trust and loyalty. *Economic Inquiry*, 31(4): 663–670.

Gibbons, R., and R. Henderson (2012). Relational contracts and organizational capabilities. *Organization Science* 23.5: 1350–1364.

Gibbons, R. and K. J Murphy. 1992. Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," The Journal of Political Economy 100 (3): 468–505.

Hambrick, D. C., and G. DS Fukutomi (1991). The seasons of a CEO's tenure. *Academy of Management Review*, 16.4: 719–742.

Harris, M. and B. Holmström. 1982. A Theory of Wage Dynamics. *The Review of Economic Studies* 49 (3): 315–333.

Holmstrom, B., and P. Milgrom (1990). Regulating Trade Among Agents. *Journal of Institutional and Theoretical Economics*: 85–105.

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

Ittner, C.D., and Larcker, D.F. (2001), Assessing empirical research in managerial accounting: a value based perspective. *Journal of Accounting and Economics*, 32, pp. 349-410.

Ittner, C., D. Larcker, and M. Meyer (2003). Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review*, 78 (3): 725–758.

Ittner, C., D. Larcker, and T. Randall (2003). Performance implications of strategic performance measures in financial services firms. *Accounting, Organizations and Society*, 28 (7/8): 715–741.

Kaplan, R. S., and D. P. Norton (1992). The balanced scorecard: Measures that drive performance. *Harvard Business Review*, 70 (1): 71–79.

Lee, F., A. Edmondson, S. Thomke, and M. Worline (2004). The mixed effects of inconsistency on experimentation in organizations. *Organization Science*, 15 (3): 310–326.

Merchant, K.A. (1985). *Control in Business Organizations*. Ballinger Publishing Company.

Milgrom, P.R. (1988). Employment contracts, influence activities, and efficient organization design. *Journal of Political Economy*, 96, 42–60.

Milgrom, P.R. and J. Roberts (1992). *Economics, Organization and Management*. Prentice-Hall, Upper-Saddle River, N.J.

Nagin, D. S., J. B. Rebitzer, S. Sanders, and L. J. Taylor (2002). Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. *American Economic Review*, 92 (4): 850–873.

Sturman, M. C. (2003).Searching for the inverted U-shaped relationship between time and performance: Meta-analyses of the experience/performance, tenure/performance, and age/performance relationships. *Journal of Management*, 29(5): 609–640.

Wiseman, R. M., and L. R. Gomez-Mejia (1998). A behavioral agency model of managerial risk taking. *Academy of Management Review*, 23 (1): 133–153.

Appendix A Variable Definition

WC	The absolute amount of working capital of each project managed by a project manager, calculated as the sum of accounts receivable and work-in-progress subtracted by accounts payable in the firm's internal accounting system.
WC%	WC% denotes working capital (WC) scaled by project size. We keep the observations with WC% ranging from -100% to 100%.
CCC	Cash conversion cycle, calculated as (days sales outstanding + days inventory outstanding - days payables outstanding) in the firm's internal accounting system. We keep the observations with CCC ranging from -365 to 365.
Pending	Number of quarters between when the project is initiated and when the first invoice is sent to the client.
#invoice	Number of invoices sent in the current quarter with regard to each project.
#invoices_tt	Total number of invoices sent with regard to each project (since last fiscal year).
Size	Size of each project, measured as the estimated revenue of the project.
log(1+Size)	Natural logrithmn of project size (size+1)
Gov	Gov indicates the project type. It takes the value of 1 if the project is done for the government and 0 otherwise.
Tenure	Tenure denotes total time from when the project manager started working for DC till end of year 2015.
Age	Age represents age of the project manager who runs the project.
#proj	Total number of projects one project manager is responsible for at the current quarter
After	Takes the value of 1 if the observation is on or later than the fourth quarter of 2014 and 0 otherwise.

Table 1 Descriptive Statistics

Panel A shows the descriptive statistics of indicators of conversion efficiency, and Panel B shows descriptive statistics of demographic characteristics specific to each project. Specific definitions of variables are presented in Appendix A. Observations are at project-quarter level unless indicated otherwise. *Tenure* and *Age* are summarized at the project manager level. *#proj* is summarized at the manager-quarter level.

	N	Mean	min	p25	Median	p75	max	sd
Panel A: Indicators of conversion efficiency								
WC	28551	2,690.93	-4,824,901	-223	52.39	976	2,918,206	71,38
WC%	22696	8.43	-100	-17.8	0	38.23	100	48.65
CCC	23313	26.97	-364.8	0	0	82.03	365	99.35
Pending	28551	1.32	-2	1	1	2	4	0.81
#invoices	28551	0.78	0	0	0	0	145	4.09
#invoices_tt	28551	1.65	0	0	0	1	464	9.08
Panel B: Demographic variables – Project-quarter level								
Size	28551	49,577	0.1	430	1,224.4	7247.9	18,865,759	454,436
Gov	28551	0.02	0	0	0	0	1	0.13
Manager and Manager-quarter level								
Tenure	92	14.49	2	4.5	10	24	42	11.95
Age	92	47.97	26.9	40.3	50.65	54.9	71.4	9.93
#proj	479	59.61	1	2	11	32	1,779	176.4

Table 2 Mean Conversion Efficiency Indicators Before and After the Control

This table presents the mean value of each indicators of conversion efficiency of project-quarter observations before and after introducing the specific performance measure (fourth quarter of 2014). The Difference column shows the result of a two-sided t-test of each variable in the Before and After column. *, **, *** represent results significant at the 90%, 95%, and 99% levels, respectively. Specific definitions of variables can be found in Appendix A.

	Before 2014Q4		***After 2014Q4***		***Difference***
	N	Mean	N	Mean	t-value
WC	13980	4106.8	14571	1861.5	2.65***
WC%	10616	13.38	12080	4.08	14.43***
CCC	11004	40.24	12309	15.11	19.43***
Pending	13980	1.35	14571	1.28	7.88***
#invoices	13980	0.71	14571	0.84	-6.72***
#invoices_tot	13980	1.28	14571	2.00	-2.63***

Table 3 Correlation Coefficient of Variables

This table presents the correlation coefficients of variables used in this study. *, **, *** represent results significant at the 90%, 95%, and 99% levels, respectively. The p-value is presented in the parentheses. Specific definitions of variables can be found in Appendix A.

	WC	***WC%***	***CCC***	***Pending***	***#invoices***	***#invoices_tt***	***Tenure***	***Age***	***Size***	***#proj***	***Gov***
WC	1										
WC%	0.09***	1									
	(0.00)										
CCC	0.18***	0.49***	1								
	(0.00)	(0.00)									
Pending	0.01***	0.19***	0.02***	1							
	(0.00)	(0.00)	(0.00)								
#invoices	0.06***	0.02***	0.05***	-0.02***	1						
	(0.00)	(0.00)	(0.00)	(0.00)							
#invoices_tt	0.04***	-0.003	0.03***	-0.01***	0.82***	1					
	(0.00)	(0.68)	(0.00)	(0.00)	(0.00)						
Tenure	-0.02***	0.001	-0.03***	0.03***	0.03***	0.02***	1				
	(0.00)	(0.92)	(0.00)	(0.00)	(0.00)	(0.00)					
Age	-0.01**	-0.04***	-0.09***	-0.04***	-0.009	-0.02***	0.43***	1			
	(0.02)	(0.00)	(0.00)	(0.00)	(0.13)	(0.00)	(0.00)				
Size	0.02***	-0.01*	-0.00	-0.02***	0.16***	0.17***	0.00	-0.03***	1		
	(0.00)	(0.05)	(0.60)	(0.00)	(0.00)	(0.00)	(0.40)	(0.00)			
#proj	-0.03***	-0.05***	-0.09***	-0.04***	-0.13***	-0.12***	-0.15***	-0.06***	-0.05***	1	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		
Gov	0.004	0.01***	0.000	0.03***	0.01	0.006	0.22***	0.15***	0.14***	-0.10***	1
	(0.48)	(0.06)	(0.93)	(0.00)	(0.09)	(0.27)	(0.00)	(0.00)	(0.00)	(0.00)	

Table 4 Effect of Tenure on WC%

This table presents the empirical results of the following regression:

$$WC\%_{i,t}=\alpha+\beta1*Tenure_{i,t}+\beta2*After_{t}+\beta3*Tenure*After_{i,t}+\gamma*Controls+\epsilon_{i,t}. \quad (1)$$

Details about each specification are included in each column. "Yes" in Quarter FE, Project manager FE, and Client FE indicates the quarter fixed effect, project manager fixed effect, and client fixed effect is included in the specification, respectively. We cluster standard errors by projects across all specifications. *, **, *** represent results significant at the 90%, 95%, and 99% levels, respectively. P-values are presented in the parentheses. The F-stats of testing b1+b3=0 is reported when applicable. Specific definitions of variables can be found in Appendix A. Column (4) shows the result of a placebo test of the parallel trend assumption. *After* takes the value of 1 if the observation is on or later than the 2014Q1 and before 2014Q4, and 0 if the observation is before 2014Q4.

		Dependent Variable: WC%			
	Pred. Sign	(1)	(2)	(3)	(4)
Tenure	**-**	-0.231***	-	-	0.220
		(0.00)	-	-	(0.15)
After	**-**	-15.10***	-6.171***	-10.59***	3.192
		(0.00)	(0.00)	(0.00)	(0.24)
Tenure*After	+	0.563***	0.264***	0.234***	-0.233
		(0.00)	(0.00)	(0.00)	(0.15)
Age		-0.391***	-	-	-0.992***
		(0.00)	-	-	(0.00)
Log(1+Size)		-1.769***	-2.715***	-2.459***	-2.106***
		(0.00)	(0.00)	(0.00)	(0.00)
#proj		-0.005***	0.003**	0.000	-0.006***
		(0.00)	(0.04)	(0.80)	(0.00)
Gov		6.187**	8.008*	-	12.04**
		(0.02)	(0.09)	-	(0.01)
Quarter FE		Yes	Yes	Yes	Yes
Project Manager FE		No	Yes	Yes	No
Client FE		No	No	Yes	No
Cluster SE		prj	prj	prj	prj
F-stat (b1+b3=0)		50.82***	-	-	0.03
obs		22696	22696	22694	15925
R^2		0.05	0.11	0.18	0.07

Table 5 WC% Before and After Policy Change for Groups of Managers

This table presents the mean working capital percentage (*WC%*) before and after the introduction of the specific performance measure of working capital, shown separately for the projects managed by junior versus senior project managers. The Diff column shows the result of a two-sided t-test of each group in the Before and After column. The Diff row shows the result of a two-sided t-test of the comparison with Junior versus Senior in each period. *, **, *** represent results significant at the 90%, 95%, and 99% levels, respectively. **Junior** refers to projects managed by junior managers (whose tenure is no more than 10 years), and **Senior** refers to the ones managed by senior managers (whose tenure is longer than 10 years). **Before** accounts for observations from the first quarter of 2013 up until the fourth quarter of 2014, and **After** accounts for the rest of observations.

	Before		**After**		
	#obs	Mean	#obs	Mean	Diff (t-stat)
Junior	7427	18.18	6742	3.16	18.36***
Senior	3189	2.19	5338	5.24	-2.87***
Diff (t-stat)		15.63***		-2.35**	

Table 6 Mean Size and Type of Project Allocated Among Subgroups

Panel A presents the mean statistic of *Size* and *Gov* for projects assigned to Junior and Senior project managers, respectively. Panel B presents for the projects assigned to Junior and Senior managers separately, and the mean statistic of *Size* for observations before and after the specific performance measure. The Diff column shows the result of a two-sided t-test of the two variables in each cell. *, **, *** represent results significant at the 90%, 95%, and 99% levels, respectively. **Junior** refers to projects managed by junior managers (whose tenure is no more than 10 years) and **Senior** refers to the ones managed by senior managers (whose tenure is longer than 10 years). **Before** accounts for observations from the first quarter of 2013 up until the fourth quarter of 2014, and **After** accounts for the rest of observations.

Panel A: Mean Size and Gov of projects assigned to each subgroup

	#obs	Mean *Size*	Diff (t-stats)	#obs	Mean *Gov*	Diff(t-stats)
Junior	18484	47907		18484	0.006	
Senior	10067	52644	-0.84	10067	0.036	-19.32***

Panel B: Mean Size and Gov of projects assigned to each subgroup before and after the implementation of the specific performance measure

	Junior				
	#obs	Mean *Size*	Diff (t-stats)	Mean *Gov*	Diff (t-stats)
Before	10012	45734		0.006	
After	8472	50475	-0.68	0.005	0.48
	Senior				
	#obs	Mean *Size*	Diff (t-stats)	Mean *Gov*	Diff (t-stats)
Before	3968	50970		0.053	
After	6099	53733	-0.31	0.025	7.39***

Table 7 Regression Result of Alternative Performance Measure used as DV

This table presents the regression result using realized revenue, which is an alternative performance measure other than the conversion efficiency, as the dependent variable. Details about each specification are included in each column. "Yes" in Quarter FE, Project manager FE, and Client FE indicates the quarter fixed effect, project manager fixed effect, and client fixed effect is included in the specification, respectively. We cluster standard errors by projects across both specifications. *, **, *** represent results significant at the 90%, 95%, and 99% levels, respectively. In the parentheses, p-values are presented. *Realized Rev* is the realized revenue for each project tracked in DC's internal accounting system. Definition of independent variables can be found in Appendix A.

	Pred. Sign	**DV: *Realized Rev***	
Tenure	N/A	921.06	-
		(0.40)	-
After	N/A	-1034.79	1893.07
		(0.96)	(0.85)
Tenure*After	N/A	-505.94	-746.25
		(0.62)	(0.30)
Age		-2654.92***	-
		(0.00)	-
Log(1+Size)		63232.2***	40510***
		(0.00)	(0.00)
#proj		-14.32***	1.36
		(0.00)	(0.76)
Gov		14851.6	-
		(0.67)	-
Quarter FE		Yes	Yes
Project Manager FE		No	Yes
Client FE		No	Yes
Cluster SE		prj	prj
obs		28551	28549
R^2		0.09	0.37

Table 8 Regression Results Using Alternative DVs

This table presents the empirical results of the following regression(s):

$$DV_{i,t}=\alpha+\beta1*Tenure_{i,t}+\beta2*After_{t}+\beta3*Tenure*After_{i,t}+\gamma*Controls+\epsilon_{i,t}. \quad (2)$$

The DVs are alternative indicators of conversion efficiency other than WC%. Details of the specification is included in each column. "Yes" in Quarter FE indicates the quarter fixed effect is included in the specifications. We cluster standard errors by projects across all specifications. *, **, *** represent results significant at the 90%, 95%, and 99% levels, respectively. In the parentheses, p-values are presented. Standard errors are clustered at project level. Specific definitions of variables can be found in Appendix A.

		Dependent Variables			
	Pred. Sign	**log(1+WC%)**	***WC***	***CCC***	***Pending***
Tenure	-	-0.002*	-512.36*	-0.409***	-0.002**
		(0.06)	(0.07)	(0.00)	(0.01)
After	-	-0.079***	-10351***	-28.82***	-0.324***
		(0.00)	(0.00)	(0.00)	(0.00)
Tenure*After	+	0.005***	541.27*	0.485***	0.010***
		(0.00)	(0.06)	(0.00)	(0.00)
Age		-0.009***	-25.66	-0.977***	-0.003***
		(0.00)	(0.78)	(0.00)	(0.00)
log(1+Size)		-0.003	1750.65	0.294	-0.065***
		(0.36)	(0.14)	(0.47)	(0.00)
#proj		-0.000***	-1.855***	-0.159***	-0.000***
		(0.00)	(0.00)	(0.00)	(0.00)
Gov		0.107**	6524.5*	3.233	0.109
		(0.04)	(0.06)	(0.67)	(0.14)
Quarter FE		Yes	Yes	Yes	Yes
Cluster SE		prj	prj	prj	prj
obs		22301	28551	23313	28551
R^2		0.03	0.01	0.04	0.06

Table 9 Regressions Results by Using Senior as a Dummy Variable

This table presents the empirical results of the following regression:

$$DV_{i,t}=\alpha+\beta1*Senior_{i,t}+\beta2*After_{t}+\beta3*Senior*After_{i,t}+\gamma*Controls+\epsilon_{i,t}. \quad (3)$$

In columns (1) and (2), the DV is *WC%*, in column (3) the DV is *WC* and in column (4) the DV is *CCC*. Details of each specification are included in each column. "Yes" in Quarter FE and Project manager FE indicates the quarter fixed effect, and project manager fixed effect is included in the specification, respectively. We cluster standard errors by projects across all specifications. *, **, *** represent results significant at the 90%, 95%, and 99% levels, respectively. In the parentheses, p-values are presented. The F-stats of testing b1+b3=0 is reported when applicable. *Senior* is a dummy variable that takes the value of 1 if the project is managed by a project manager with tenure larger than 10 years and 0 otherwise. Specific definitions of variables can be found in Appendix A.

		Dependent Variables			
		(1)	(2)	(3)	(4)
	Pred. Sign	***WC%***	***WC%***	***WC***	***CCC***
Senior	-	-19.78***	-	-	-
		(0.00)	-	-	-
After	-	-15.10***	-8.439***	-4931.12***	-23.05***
		(0.00)	(0.00)	(0.00)	(0.00)
Senior*After	+	19.98***	13.84***	7916.0**	10.41***
		(0.00)	(0.00)	(0.03)	(0.00)
Age		-0.040	-	-	-
		(0.59)	-	-	-
log(1+Size)		-2.026***	-2.656***	1474.2	-1.389***
		(0.00)	(0.00)	(0.20)	(0.00)
#proj		-0.007***	-0.000	-1.877	-0.001
		(0.00)	(0.97)	(0.10)	(0.69)
Gov		8.045*	8.282*	579.6	-12.80
		(0.06)	(0.08)	(0.84)	(0.19)
Quarter FE		Yes	Yes	Yes	Yes
Project Manager FE		No	Yes	Yes	Yes
Cluster SE		prj	prj	prj	prj
F-stat (b1+b3)=0		0.03	-	-	-
obs		22696	22696	28551	23313
R^2		0.06	0.11	0.06	0.08

Table 10 Cross-Sectional Evidence Conditional on Type of Client

This table presents the empirical results of the main regression:

$$WC\%_{i,t}=\alpha+\beta1*Tenure_{i,t}+\beta2*After_{t}+\beta3*Tenure*After_{i,t}+\gamma*Controls+\epsilon_{i,t}. \quad (1)$$

Conditional on the type of clients each project is designated. *Gov=0* represents the regression analysis of the subsample where the projects are designated to a non-government client, and *Gov=1* represents the same analysis of the subsample of government projects. *DC Client* represents regression analysis of the subsample of projects that are billed to a different department within the DC corporation, and *NonDC Client* represents the analysis of the rest of projects. We cluster standard errors by projects across all specifications. *, **, *** represent results significant at the 90%, 95%, and 99% levels, respectively. P-values are presented in the parentheses. Specific definitions of variables can be found in Appendix A.

	DV=WC%			
	(1)	(2)	(3)	(4)
	Gov=0	*Gov=1*	*NonDC Client*	*DC Client*
Tenure	-0.261***	0.036	-0.438***	0.281**
	(0.00)	(0.91)	(0.00)	(0.04)
After	-15.641***	-11.037	-13.842***	-5.394**
	(0.00)	(0.37)	(0.00)	(0.03)
Tenure*After	0.617***	-0.071	0.642***	-0.124
	(0.00)	(0.85)	(0.00)	(0.35)
log(1+Size)	-1.782***	-1.071	-2.223***	-1.264***
	(0.00)	(0.58)	(0.00)	(0.00)
Age	-0.388***	-1.576*	0.024	-0.711***
	(0.00)	(0.08)	(0.80)	(0.00)
#proj	-0.005***	-0.004	-0.016***	-0.008***
	(0.00)	(0.87)	(0.00)	(0.00)
Quarter FE	Yes	Yes	Yes	Yes
Project Manager FE	No	No	No	No
Cluster SE	prj	prj	prj	prj
#obs	22,367	329	12,607	10,089
R^2	0.048	0.065	0.044	0.058

Chapter 3 Performance Appraisal on Employee's Future Potential

3.1 Introduction

Gauging employee's talent and identifying promising employees to fit for positions that require more comprehensive set of skills is of crucial importance to organizations. The assessment of employee's potential of adding value to the organization on a continuous basis requires the evaluation system to pick up the forward-looking characteristics in the employee's overall competence, which can significantly diverge from what is implied by the performance outcome achieved on the employee's current job (Bol and Leiby 2018, Deller 2018, Grabner, Künneke and Moers 2018WP, Grabner and Moers 2013, De Pater et al. 2009). In order to identify the underlying potential of each individual employee, firms may resort to metrics that reflect an employee's innate characteristics, such as leadership or teamwork. Appraisals of these "soft qualities" are most commonly provided by a supervisor who interacts with the employee on a regular basis, and are by definition the supervisor's subjective assessment on how strong/weak she perceives the employee to be on certain dimensions of competence. Being a form of subjective assessment in nature, the appraisal can be prone to inaccuracy and biases (e.g., Ittner, Larcker and Meyer 2003, Moers 2005, Bol 2011, Maas, Van Rinsum and Towry 2012, Du, Tang, Young 2012, Demeré, Sedatole and Woods 2018). Biases driven by supervisor's personal incentive or preference set aside (Prendergast 1999, Bol 2011, Grabner, Künneke and Moers 2019), previous literature has documented that supervisor's subjective evaluation tends to echo the information contained in objectively measured performance outcome, which implies that supervisors exert their endowed discretion only to a limited extent, and the presence of standalone performance measure compromises the effective use of subjective evaluation in gauging the soft skills of employees (Ittner et al. 2003, Bailey, Hecht and Towry 2011, Bol and Smith 2011).

Apart from the psychological factor underlying supervisor's evaluation process, I put forward the idea that the observed overlap of information contained in performance appraisal and standalone performance outcomes can also be explained by the purpose these performance appraisals serve. Performance appraisal, in many cases being the only reward or recognition of employee's performance in any form (e.g., Medoff and Abraham 1980, Baker, Gibbons and Holmstrom 1994ab, Gibbs and Hendricks 2004), serves both the function of rewarding the employee for what he has achieved on his current job, as well as representing

supervisor's assessment of the employee's contribution along the upcoming horizon. Since performance appraisal plays both roles at the same time, it only conveys a blurred message of the employee's overall picture on the employee's future potential; and as a consequence, performance appraisal provided by supervisors is not informative for the firm to effectively select promising employees or develop their specific human capital. I therefore predict that if the function of rewarding employees for their past performance is disentangled from the performance appraisal given by the supervisor, the latter will become a more precise metric in gauging the employee's future potential.

I conduct my study with the performance and personnel data from a professional service firm who introduced a performance-based bonus to its employees during the sampling period. The research site in the Netherlands is represented by a managing director, with several functional departments making decisions for the whole area with regard to human resources, marketing, business development, etc. The managing director and the functional departments are collectively referred to as "the firm" in this study. The firm has business sites in three Dutch cities, employees work for one of the three sites and provide financial advice or consulting services for their corporate clients. In addition, within each site there are a couple of business managers (referred to as "supervisors") in charge of general operations with the employees and local clients. The firm keeps track of the revenue each employee brings in every year by conducting service to the clients, a client satisfaction score from a survey filled by their clients, as well as a set of specified dimensions of competence appraised by the supervisor, including communication, client focus, expertise, etc. Before the year 2006, the firm uses a universal pre-set formula to determine each employee's salary raise, in which supervisor's appraisal has the most prominent influence. Beginning in year 2006, the firm introduced an annual bonus to all employees in the firm, which is also determined by a pre-set formula comprises of revenue each employee generates and their customer satisfaction rating. The introduction of a performance-based bonus on top of the existing compensation system now better recognizes and rewards employees for their contemporaneous performance, as a consequence, I expect the appraisal becomes a "cleaner" instrument to more precisely represent the supervisor's assessment of the employee's human capital with a future prospect. I thus examine whether and how supervisors alter their way of providing performance appraisals after the firm introduced the performance-based bonus to the employees.

I find that after the firm introduced the bonus, the supervisor's appraisal across multiple dimensions of each employee becomes more dispersed; moreover, the supervisor's performance appraisal is positively associated with employee's current financial performance before the introduction of bonus, while the positive association is significantly reduced afterwards. These two findings jointly speak to a documented reduction of halo effect: That instead of assigning a similar rating to a series of evaluation based on a general impression (which most likely is the employee's achievement on financial performance outcome), the supervisor now tends to distinguish among the different aspects of an employee's overall competence, and her appraisal indeed reflects something other than what is contained in the employee's current performance measures. I further find that the supervisor's appraisal becomes more predictive of employee's financial performance (i.e., revenue) in the following year, suggesting that supervisor's appraisal now functions more effectively in picking up the forward looking information rather than solely acknowledging what the employee has achieved in the past. In addition, I also find that the outcome of supervisor's appraisal better explains the employee's chance of promotion, which implies that supervisor's appraisal becomes a more reliable source of information in reflecting the employee's underlying human resource in the firm's personnel related decision. Taken together, the findings point towards my prediction that supervisor's appraisal picks up the forward-looking aspects of the employee's overall contribution to the firm than that reflected in the standalone performance measures, making the performance appraisal a precise metric of gauging an employee's future potential. As a result, the collective performance measurement system altogether allows the firm to effectively identify and develop its employees' underlying human resource that looks towards the future, without compromising the need to effectively recognize and reward employees for what they have achieved thus far.

This study primarily contributes to the growing body of literature on the assessment of employee potential over and beyond their current performance. Deller (2018), Bol and Leiby (2018), and Grabner et al. (2018) document the emerging managerial practice of separately examining employee's performance on the current job as well as their prospect of progressing to a higher position, highlighting the importance of identifying employee's potential by the organization. Along the line of assessing employee's potential, my study provides further understanding of the accuracy and usefulness of such assessment. By exploring a change in the research site's pay policy, I identify a condition where the appraisal can serve a "cleaner" function of

delegating supervisor's assessment of employee's soft aspects of performance which is deemed to be evaluated subjectively; and as a consequence, I find the information contained in supervisor's appraisal becomes more diverged from what is reflected by the employee's current performance, and gears more towards the recognition of the employee's competence with a forward-looking nature.

In a similar vein, this study also contributes to research on subjective performance evaluation used for selection purpose. Aside from serving as a performance indicator and motivating employees to exert effort, subjective performance evaluation can also play a crucial role of picking up employee's aspects of innate characteristics for appointment and further human resource development (Grabner and Moers 2013, 2018WP), which has been much less studied and understood. In my study the supervisor's appraisal appears as a form of subjective evaluation, and there has been debates and ambiguity about what elements are evaluated in supervisor's appraisal; Namely, outcome of supervisor's appraisal can be any combination of the employee's ability, potential, and periodic effort (Medoff and Abraham 1980, Gibbs and Hendricks 2004). In this study I explore a condition where the supervisor is granted a better instrument to differentiate the role of providing her true assessment of employee's competence from the need to reward his current effort, and as a consequence, her appraisal can be more effective in picking up the employee's innate characteristics and facilitate the use of their evaluation for selection purpose.

Furthermore, I show in this study that by making it more clear to the supervisors what is expected from them, the firm obtains the intended outcome of having more precise information about its employees' underlying human resource. This is also in line with the finding in Demeré et al. (2018) that supervisors adjust the way they give performance appraisal in response to the firm's implied message on what their appraisal should be. This paper therefore also adds to the scant field evidence that without changing any explicit incentive, the design of management control system can facilitate agents to perform better aligned with the organizational objective (Casas-Arce, Martinez-Jerez and Narayanan 2017).

Remaining of the paper develops as follows. In section 3.2, I explain related theories and develop my hypothesis. In section 3.3, I describe the institutional context of the research setting and data available. I then explain the empirical models to test the hypothesis and corresponding empirical results in section 3.4.

Additional analysis that facilitates robustness and further understanding the result is explained in section 3.5. Section 3.6 concludes the paper.

3.2 Hypothesis Development

Assessing an employee's potential perspective over and above his current performance is of crucial importance to an organization. Bol and Leiby (2018) finds when a manager has the employee's promotion prospect in mind, she tends to diverge her assessment of the employee's chance of promotion as opposed to his current performance. Deller (2018) empirically documents a company's managerial practice of separating assessing employees' current job performance as well as their potential, and finds that employees with higher rating on potential have significantly better chance of being promoted. These studies highlight the importance of a personnel system that effectively documents the employee's underlying human capital for future appointment or training, which is not reflected in the employee's current performance (e.g., Gibbons 1998, Gibbons and Waldman 1999a, Prendergast 1999).

Theory and empirical evidence predict that firms recognize talent by attaching assessment of employees' underlying human capital with a progress in their levels of fixed salary. Prendergast (1993) theoretically predicts that an increase in wage serves as the firm's commitment in recognizing the employee's acquisition of firm-specific human capital. It is reasonable that salary raise rewards and recognizes employee's more permanent characteristic, as downward adjustment in salary is rarely observed (e.g., Gibbs and Hendricks 2004, Frederiksen, Lange, Kriechel 2017), salary raise granted to employees has relatively more on-going consequence to the organization, such that the firm expects the employees to at least make good for the amount of salary raise would entail in their upcoming tenure. While salary raise rewards and recognizes employee's acquisition in human capital with a forward-looking perspective, the presence of a transitory form of reward (i.e., performance based bonus) complements the organizational purpose by recognizing the employee's achievement on his current job (Baker et al. 1988, Ederhof 2011).

To identify and recognize employee's acquisition in human capital, the firm will need some indicator that reveals an employee's innate character that enables him to continuously deliver, such as teamwork, leadership, customer focus, among others. These "soft" aspects of employee characters, if available at all,

usually have to take the form of performance ratings appraised by a supervisor who has reasonably close contact with the regarding employee. To the extent that salary raise effectively recognizes employee's underlying human capital and on-going value to the firm, the firm reasonably expects the supervisor to evaluate her employees based on her judgement on the employee's personnel qualities specified as dimensions of competence, which may not have to be related to the employee's realized objective performance outcome in the current period. In other words, supervisor's appraisal on employee's specified dimensions of competence is expected to serve as selection purpose, which is informative in facilitating firm's overall personnel decisions including training or promotion.

Providing performance appraisal that accurately reflects employee's innate competence may not be an simple and straightforward task though, as information about employee's current performance may obscure the supervisor's objective of assessing the forward-looking performance potential with regard to employees. Ittner et al. (2003) finds that the firm puts the most weight on financial performance outcome when subjectively deciding on the bonus, despite that the financial measure is not the most effective indicator of the employee's action on the balanced score card. Bol and Smith (2011) document in a lab experiment that supervisor's subjective evaluation is highly influenced by the employee's achievement on the objective performance measure, even when the subjective evaluation is about an arguably independent task from the one measured objectively. Bailey et al. (2011) puts forward a behavioral explanation that when coming up with evaluation of certain aspects of employees, the supervisor may nonetheless anchor upon information in the employee's achieved performance outcome.

In a similar vein, Hayes and Schaefer (2000) documents a positive association between the unexplained (by concurrent performance) part of CEO's compensation and their future performance, implying that firms incorporate non-observable, forward-looking information in assessing employee's overall achievement. These studies generally speak to the idea that principals tend to use employee's observed performance outcome as a starting point, and complement this information with an evaluation of how this information reflects the future potential of the employee. This may result in an appraisal with non-trivial information overlap from concurrent performance measures, which hinders the effective identification of employee's talent and future potential. It is particularly so when it is not possible for the principal to differentiate between past performance and performance potential in the performance evaluation system. Such a situation

would exist in the situation where the firm does not provide for variable pay. Employees naturally embraces the aversion of unfair treatment in the sense that their exerted effort is not credibly recognized and rewarded (Lazear 1979, Prendergast 1999). If a supervisor, in that case, wants to compensate an employee for good performance, she has to resort to a salary raise. Under that condition the supervisor must consider past and future performance and attach one number to compensate the past and at the same time express her expectation of the future. Absence of a variable pay based on performance inevitably mixes the selection purpose embraced in the supervisor's appraisal with that of recognizing employee's periodic effort. The supervisor's appraisal therefore, rather than serving the organizational purpose of identifying employee's underlying talent and potential, instead more or less picks up similar aspects of the employee's value as his concurrent performance outcome. The information contained in supervisor's appraisal is largely overlapped with that reflected by employee's performance outcome, and as a consequence, the information desired by the firm about employee's underlying human capital is compromised.

The introduction of a performance based bonus now serves as an effective recognition of employee's contemporaneous performance, and by partitioning the function of rewarding employee's past performance, it grants the supervisors with a better instrument to more precisely represent her assessment of employee's underlying human capital. In order to empirically test that supervisor's appraisal becomes a more precise metric of employee potential, I state the following two parallel hypothesis. In light of Hayes and Schaefer (2000) and Ederhof (2010), if the supervisor's appraisal is indeed effective in capturing the employee's potential in continuous contribution beyond what is reflected in the current performance outcome, then the current period appraisal is expected to be positively associated with future period financial performance. To the extent that after the introduction of a performance-based bonus, the supervisor's performance appraisal will better capture employee's on-going potential rather than the performance he has realized so far, I expect the supervisor's appraisal will be more predictive of the employee's future performance. I state this hypothesis as follows:

H1a: After the firm introduced a performance-based bonus, the supervisor's performance appraisal will be more predictive of the employee's future performance.

Meanwhile, as supervisor's appraisal becomes more accurate in reflecting their assessment of employee's underlying human capital, it will therefore be more informative for the firm to use in its personnel decisions, such as promotion. Aside from serving as a form of incentive mechanism, promotion also plays the role of selecting employees with the required capability to function at a new position (e.g., Gibbs 1995). Campbell (2008) documents in a fast food chain, the store manager's performance on customer satisfaction is positively related to her chance of promotion, the noticeably large association documented in the study provides compelling evidence that the non-financial, qualitative information that may predict the employee's future performance is greatly taken into account in the firm's promotion decision (Gibbs 2008). Grabner and Moers (2013) further provides empirical evidence that when the purpose of selecting employees with the proper ability is more pronounced, performance indicator that reflects the employee's overall capability rather than the observed current performance will be put a relatively heavier weight in making promotion decisions. These findings suggest that if available at all, performance indicators pointing towards an employee's underlying ability to benefit the firm in the long run will be used in the firm's promotion decision, given the promotion serves as sorting purpose. Speaking to my study, before there is a performance-based bonus, supervisor's appraisal contains information more or less similar to that in employee's performance outcome, whereas after the introduction of the bonus, supervisor's appraisal points towards picking up the underlying employee's competence and potential more efficiently than before. As such, taking supervisor's appraisal into account, on top of employee's current performance measures, will better facilitate the firm in making an informed decision on promotion (Deller 2018). In testing the prediction that supervisor's performance appraisal after bonus becomes more precise in picking up employee's underlying human capital, I expect supervisor's appraisal will matter more in explaining employee's chance of promotion after introduction of the bonus. I state the hypothesis as follows:

H1b: After the firm introduced a performance-based bonus, firm's promotion decision will be more sensitive to supervisor's performance appraisal.

3.3 Research Setting and Data Description

The personnel data and operational information used in this study is obtained from an international professional service firm. The headquarter is a listed company with subsidiaries in several west European

countries, the data I use in this study is from its office in the Netherlands, anonymized as FM. The Dutch office is represented by a managing director, with several independent functional departments that take care of specific duties related to human resource management, marketing, project management, etc. For instance, the human resource department is responsible for all professional agents in the Dutch branch, and is in charge of all issues related to personnel management such as hiring, performance evaluation, compensation and promotion, among others. All professional agents at FM are affiliated with one of the three local business sites, each located in one of the Dutch cities and is represented by a site director. In addition to all the professional agents who conduct professional service for the clients, there are also a couple of business managers at each local office, who are appointed to maintain customer relation, manage projects, as well as to monitor and evaluate the agents' performance and personnel development. The organization structure of FM is briefly sketched in Appendix A. In this study, I refer to the professional agents as "employees", the business managers who give the performance appraisals as "supervisors", and everything beyond the supervisor's realm of duty as "the firm".

The data I have available is the personnel characteristics and performance information about all employees over the years 2004 to 2008. The business of FM mainly involves providing financial advice, internal control services, and at times consulting services for their corporate clients, all employees at FM have their college degree (or above) in finance or accounting related majors. In Table 1 Panel A I present the descriptive statistics of the demographic characteristics of the employees. During the sample period, about three quarters (73%) of employees are men, the average age of the employees is 32.29 years old, and on average they stay at FM for 2.82 years. There are in total five levels of employees, the tasks they do and skills needed are more or less comparable among job levels except that at higher levels (e.g., level 4 and 5) employees are also expect to coach lower level colleagues and may serve different kinds of clients. FM hires and assigns employees to different levels according to experience they have accumulated within or outside of FM, but in the same professional service industry. The average years of experience employees in FM have in the relevant industry is 9.25 years.

(insert Table 1 about here)

Because of the nature of the business, employees are the most treasured capital of the firm. The firm thus sets up a comprehensive personnel system that keeps track of multiple indicators of employee performance, based on which the firm also determines its employee compensation. Since the employees mostly work on projects with clients on an individual basis rather than in teams, it is relatively clean to track for each employee how much revenue they bring in for serving their client(s) each year. Upon finishing a project with a client, the firm also sends a survey to the client, and asks them to rate, on a scale of 1 to 6, how satisfied they are with the service provided by the employee who is responsible for them. By the end of each year, the system keeps record of the mean customer satisfaction rating from the client(s) served by each employee. In addition, for the purpose of better understanding and nurturing the value and contribution of each employee, the firm also asks the supervisors to provide the performance appraisal of each employee's competence as well as their involvement with FM. Appraisal of employee competence breaks down to seven dimensions: client focus, communication, result orientation, flexibility, expertise, proactive behavior, and analysis and advise. By the end of the year, the supervisor will have a meeting with each employee individually, talk about their general performance over the past year, and for each dimension of employee competence, evaluate as well on a scale from 1 to 6 of how well they perceive the employee has been doing. The appraisal of supervisors will be recorded in the firm's personnel system, and based on a pre-set formula imposed by the FM headquarter, the supervisor appraisal will be the major determinant (60%) of how much the employee's monthly salary increases in the following year. In addition to determining employee pay raise, the outcome of supervisors' appraisals may also be used for planning the employee's personal development such as study and training. The result of the appraisal, together with other performance indicators as well as employee's payment, are not made public to any other employees in the firm. Detailed definition of the performance indicators and formula of compensation calculation are explained in Appendix B. In this study, I keep the employee-year observations whose year-end revenue, net contribution (revenue minus two times of total salary and other labor costs), customer rating and supervisors' appraisal are all available in the system. In addition, I also manually went through the data and eliminated employee-year observations that are deemed invalid with too much sick leave (more than 300 hours within a year) or where the employee switched from full-time contract to a part-time contract. This resulted in a final sample of 502 employee-year observations in total over the years 2004 to 2008. In Table 1 Panel B, I show the descriptive statistics of the performance indicators in my sample. The correlation matrix is presented in Table 2.

(insert Table 2 about here)

In order to motivate employees to work harder, beginning of year 2006, the headquarter of FM imposed a formula based bonus to all of its subsidiaries. All employees in FM are thus granted a bonus at the end of the year from 2006 and on. The pre-set bonus formula indicates that, based on an employee's monthly salary according to the job level, the employee bonus is determined by a performance score calculated as 70% of the employee's net contribution and 30% of his average customer satisfaction rating. The pre-set bonus formula applies to all employees alike across the firm.

Notice that before the introduction of bonus, salary increase constitutes the only monetary incentive of employees, and according to the pay raise formula, supervisor appraisal has a dominant influence on it (60%). This remains to be true after the bonus comes into place except that on top of salary increase there is also a year-end bonus which, at least at face value, is *not* determined by supervisor appraisal. One may argue however, even though supervisor appraisal does not seem to have a direct influence on bonus, it can still be used to adjust ex-post the bonus each employee gets. I do observe in the data that the actual bonus or salary paid out is, more often than not, different than what is predicted by the formula. However, the adjustment of actual payment is done at the firm rather than the supervisor level, the way of adjustment is largely unknown to the employees and supervisors alike. In additional analysis section 5.1, I discuss and provide empirical evidence that the formula-based bonus is granted to reward employees for their contemporaneous effort, which provides premise for my argument that the introduction of the bonus allows the supervisor to better reflect their true assessment of employee's innate competence in their appraisal.

3.4 Main Tests and Results

In this section, I conduct a series of empirical tests to examine how supervisor appraisal changes in better recognizing employee's underlying human capital. I explain the purpose of the model for each test and its corresponding results.

3.4.1 Change in Patterns of Supervisor Appraisal

In order to understand if the supervisors change their way of providing appraisal to the employees, I first examine the two main patterns of supervisor appraisal, mean and dispersion over all competence dimensions,

to provide evidence of a change after the firm introduced the bonus. I test this with the following empirical model:

$$Pattern_{i,t}=\alpha+\beta 1*After_t+\gamma*Controls+\epsilon_{i,t} \quad (1)$$

I test the model with two dependent variables as *Pattern*: *Meanrating* is the average score across all seven dimensions of competence for each employee-year observation. For observations where not all seven dimensions are given a score, I average across whichever dimensions available for that employee-year. *Dispersion* is the standard deviation of the scores across all dimensions of competence, measured at each employee-year level as well. Since the firm introduced the performance-based bonus at the beginning of 2006, the main variable *After* takes a value of 1 for observations in years 2006, 2007, and 2008, and 0 for observations in years 2004 and 2005. I include the employee personal characters to control for the potential difference in supervisor appraisal among different employees. *Tenure* measures the years since an employee's inception at FM till when he leaves or till present. *Lvduration* measures the number of years an employee has been in the current job level. *Experience* measures the total number of years an employee has been in the same professional service industry. *Age* is measured by the years between an employee's year of born and the current year. Since *Age* and *Experience* are highly correlated (corr=0.9124, p-value=0.00), including both variables may raise concern for multicollinearity. In all the main tests, I exclude the *Experience* variable, and instead include the quadratic term of *Age* (*Age2*) to account for the potential non-linear effect of *Age* on performance and performance appraisal. The result remains consistent across all tests if I replace *Age* and *Age2* with *Experience* and squared experience, respectively. *Education* is a binary variable that takes the value of 1 if an employee has a research university degree or beyond, and 0 if he has an applied college degree. *Gender* is a binary variable that takes a value of 1 for a male employee and 0 for a female employee. I also include a control variable *1stYear* which takes the value of 1 if it is the first year the employee joins the firm, and 0 otherwise. In addition, for tests with control variables included, I also include *Level* and *Location* fixed effect to account for the potential variable across job levels and locations. Since the data I have available presents performance indicators and information at the *employee* level, I cannot identify specific supervisors who provides the appraisal or attribute the appraisal to each supervisor. I account for this limitation by controlling for the location fixed effect (there are about a handful of supervisors in each

location), and for all the following analysis, results remain significant after I cluster standard errors by location.

Result of testing supervisor appraisal pattern is shown in Table 3. Columns (1) and (2) show the regression result with *Meanrating* as the dependent variable, and columns (3) and (4) shows result of the same test with *Dispersion* as the dependent variable. For each dependent variable, I show both results with and without control variables and fixed effects. We see that both the mean (-0.180, p-value=0.00) and dispersion (-0.051, p-value=0.02) of supervisor rating is significantly lower in the first year the employee joins the firm. This is reasonable as the supervisor does not yet have enough information about nor interaction with the employee to provide insightful evaluations. We may see the mean rating is slightly lower after introduction of the bonus, but the effect is far from significant (-0.017, p-value=0.61); while the dispersion among different dimension of competence for each employee significantly increased after introducing the bonus (0.039, p-value=0.04). Both effects remain consistent when control variables and fixed effects are included. The increased dispersion in supervisor appraisal implies that after the employees are granted a performance-based bonus, the supervisors bothers more to differentiate among the multiple dimensions of competence for each employee.

(insert Table 3 about here)

3.4.2 Association of Supervisor Appraisal with Objective Performance Measures

Previous result implies a changed pattern of supervisor appraisal after introducing the bonus, I then investigate *how* supervisors provide their evaluation and if that also changed after the introduction of the bonus. Specifically, I examine what source(s) of information the supervisor refers to when she gives her evaluation of each employee. This test also helps in further understanding if and how the information contained in supervisor appraisal differs from performance outcomes. I therefore conduct the next empirical test with the following model:

$$Meanrating_{i,t}=\alpha+\beta1*log(Rev)_{i,t}+\beta2*After_{t}+\beta3*log(Rev)*After_{i,t}+\gamma*Controls+\epsilon_{i,t} \quad (2)$$

With this model, I examine the association of supervisor appraisal (*Meanrating*) with the employee's current period financial performance, and I expect the coefficient will be different (specifically, the coefficient will decrease) after the introduction of the bonus. I present the result of this test in Table 4.

Table 4 Column (1) shows the empirical result of the above model without control variables or fixed effects. We can see that *Meanrating* is positively associated with *log(Rev)* (0.177, p-value=0.01) before the firm introduces the bonus. This confirms my prediction that without a performance-based bonus, supervisor's appraisal contains information that is non-trevally overlapped with information contained in current performance measures. However, the association between *Meanrating* and *log(Rev)* decreases significantly by 0.173 (p-value=0.03) during the period after a bonus is introduced compared to before. In other words, the association between *Meanrating* and *log(Rev)* during the *After* period is 0.177+(-0.173)=0.04, significantly lower than before, and the joint significant test suggests the association between the two variables after the bonus is insignificant (F-stat=0.01). This result indicates that the supervisor refers to financial performance outcome when she gives performance appraisal during the time when bonus is absent, while she hardly does so after the introduction of the bonus. In other words, supervisor's appraisal points towards picking up different elements after the bonus compared to before. In column (2) I run the regression including another objective performance measure, customer rating (*CS*), and allows the coefficient of *Meanrating* on *CS* to vary (*After*CS*) in order to see if the supervisor reference to customer rating also changes after the bonus. We can see that the association of *Meanrating* on *CS* is highly significantly positive (0.271, p=0.00), and does not show a difference after introduction of the bonus. This is not surprising as both *CS* and *Meanrating* are evaluated on the same scale (1 to 6) and is highly correlated by nature. The crucial finding is that the association of *Meanrating* on *log(Rev)* remains consistent as discussed before, that supervisors refer to financial performance outcome significantly less when the bonus is introduced. The result of the full model with all control variables and fixed effects included is shown in column (3), and again the patterns of coefficients of *log(Rev)* stay consistent. The result in this test suggests that after the firm introduced the bonus, supervisor's performance appraisal is less associated with employee's financial performance, indicating that supervisor instead evaluates aspects of employees beyond the information contained in employee's revenue numbers when giving her performance appraisal.

(insert Table 4 about here)

3.4.3 Supervisor Appraisal in Predicting Future Performance

The two tests discussed before primarily provide evidence that after the firm introduced the bonus, supervisor appraisal picks up different elements of employee's overall value to the firm than what is reflected

by their performance measures. In this sub-section I test H1a, whether it is the case that supervisor's appraisal now instead becomes more accurate in assessing the employee's future potential. I conduct a test of how predictive the current year supervisor appraisal is to revenue that the employee makes in next year, and if it becomes more predictive during the post bonus period. I test this with the following model:

$$Log(Rev)_{i,t+1}=\alpha+\beta1*Meanrating_{i,t}+\beta2*After_t+\beta3*Meanrating*After_{i,t}+\gamma*Controls+\epsilon_{i,t} \quad (3)$$

$Meanrating_{i,t}$ is the average score of supervisor appraisal of the current period, and $Log(Rev)_{i,t+1}$ is the revenue an employee makes in the following year for observations that the employee stays for another year with FM. Result of the above empirical analysis is presented in Table 5. I first test the model without all control variables or fixed effects, but included current period financial performance to account for the persistence of revenue numbers, the regression result is shown in column (1). I find that *Meanrating* is not associated with next year's revenue during the period before bonus is introduced (-0.072, p-value=0.36), but in the post bonus period, *Meanrating* becomes incrementally and significantly more associated with future revenue than it was before (0.188, p=0.06). The combined effect during the after period equals (-0.072)+0.188=0.116, and this positive association between current period appraisal and future revenue during the after bonus period is significant (F-stat=3.44). The result suggests that supervisor appraisal is not predictive of future revenue before the introduction of the bonus, while after a bonus is introduced, an employee with one scale higher supervisor appraisal outcome is likely to show a 11.6% increase in next year's revenue he makes.

Aside from employee's performance appraisal, previous studies provide evidence that non-financial performance measures such as customer satisfaction can also lead to future revenue (e.g., Banker et al. 2000, Campbell et al. 2015). In my study, customer satisfaction rating (*CS*) is by its nature the client's subjective assessment of how satisfied they are with the employee's service, and same as supervisor appraisal (*Meanrating*), it is measured with the same 1 to 6 scale, yet it is provided by a third party that is arguably independent from within in the firm. I therefore expect that current period *CS* also predicts employee's future financial performance, but the pattern of predictability does not change before and after the bonus. I conduct the above analysis including *CS* as an additional control variable, and allow the coefficient on *CS* to vary after the introduction of bonus versus before. The result of this analysis is shown in column (2). The

insignificant coefficient on the interaction term *CS*after* is consistent with my expectation, and further assures that the empirical results regarding to supervisor appraisals are not driven by the structure of the rating (i.e., 1 to 6 scales). Importantly, when controlled for customer satisfaction, we see the pattern that performance appraisal predicts employee's future performance remains consistent and even more pronounced than the result shown in column (1) (coefficient on *meanrating*after* equals 0.256, p-value=0.02, combined effect during the post bonus period is -0.127+0.256=0.129, F-stat=3.23). Further analysis with employee's personal characteristics and fixed effects involved again provides consistent result on the predictability of performance appraisal in employee's future performance, as shown in column (3).

Result of this set of empirical test suggests that after the firm introduced the bonus, supervisor's appraisal becomes more predictive of employee's future financial performance. The result lends support to H1a, and implies that after the introduction of the bonus, supervisor appraisal diverge from confirming the information conveyed by financial performance measure, and instead becomes more effective in reflecting supervisor's assessment of employee's potential in making future contribution to the firm.

(insert Table 5 about here)

3.4.4 Firm's Use of Supervisor Appraisal in Making Promotion Decisions

Previous test provides evidence that after the introduction of the bonus, supervisor's appraisal becomes more accurate in representing her assessment of employee's underlying human capital compared to what it has been before. Following this line of reasoning, taking supervisor's appraisal into account will enable the firm to make more informed decision with regard to personnel decisions such as promotion. I then test H1b, and examine to what extent supervisor appraisal matters when the firm centrally makes promotion decision. I test this prediction with the following model:

$$Promotion_{i,t+1}=\alpha+\beta1*Meanrating_{i,t}+\beta2*After_{t}+\beta3*Meanrating*After_{i,t}+\gamma*Controls+\epsilon_{i,t} \quad (4)$$

$Promotion_{i,t+1}$ takes the value of 1 if in the following year an employee progresses to a higher level of job, and 0 otherwise. With this model, I examine how supervisor appraisal (*Meanrating*), together with other performance measures such as revenue and customer satisfaction, are put into consideration when the firm decides on whom to promote by the end of the year. Specifically, I examine if and how the weight put on supervisor appraisal changes after the bonus is introduced. Table 6 presents the empirical result of the model.

Since the dependent variable *Promotion* is a binary variable with a value of 1 or 0, a reasonable approach is to conduct the analysis using a logit regression. However, the coefficient on interaction terms is not straightforward to interpret with a logit regression model (Ai and Norton 2003, Norton, Wang, and Ai 2004), and including fixed effect in logit regressions will lose observations because of number of clusters. I therefore present the result of the empirical analysis using an OLS regression in Table 6. Result with logit regression model remains consistent and significant with that from the OLS regression (untabulated).

The result of the analysis on determinants of promotion decision is shown in column (1). To better understand the factors taken into account in firm's promotion decision, in addition to *Meanrating*, I also allow for the coefficient on revenue to differ by including the interaction term *after*log(Rev)* in the analysis. We can see that employee's chance of promotion is highly sensitive to the revenue he makes (coefficient on *log(Rev)* equals 0.261, p-value=0.00), and this coefficient does not significantly differ after the introduction of the bonus (coefficient on *after*log(Rev)* equals -0.016, p-value=0.86). Supervisor appraisal *Meanrating* is not taken into account significantly (0.088, p-value=0.36) before the bonus is introduced, this is consistent with my general prediction that before introduction of the bonus, supervisor appraisal contains information about the employee that largely overlaps with those of their objective performance measures, inclusion of this piece of information does not add much for the firm to make promotion decision. While after the introduction of the bonus, *meanrating* is put on significantly more weight for firm's promotion decisions (0.193, p-value=0.10). This implies that once the reward for employee's past performance is present, supervisor appraisal becomes incrementally informative in firm's promotion decision. In other words, during the after bonus period, an employee with one scale higher appraisal will have a 28% (=0.088+0.193) higher chance of getting promoted, and the effect is significant (F-stat=12.18).

With this analysis, I find that the firm will take supervisor appraisal of employee competence into account when making promotion decisions, but only during the after bonus period. The finding lends support for my hypothesis H1b, that supervisor appraisal better explains employee's chance of promotion after introduction of the bonus, which is again consistent with my prediction that supervisor appraisal becomes more accurate in reflecting their true assessment of employee's underlying human capital.

Note that throughout this study, promotion is referred to as the progression to the next level of job (i.e., change of job titles). With the professional nature of FM's business, the primary tasks that employees are expected to perform are similar and at all levels, while employees at higher levels are more likely to be responsible for larger and more complicated clients. Once they reached Level 3 or up, employees are also expected to engage certain hours in coaching their junior colleagues. To have a better understanding of the relation between promotion and performance appraisal conditional on the progressive change of tasks between jobs levels (Grabner and Moers 2013), I conduct the above discussed promotion analysis separately for promotion decision up till Level 3, as opposed to promotions from Level 3 and up. The result of this set of analysis is presented in columns (2) and (3). [1]By allowing the parameter of promotion determinants to alter before and after the introduction of the bonus, we can see that the observed pattern in column (1) is concentrated among the lower level promotion decisions, i.e., supervisor appraisal is not informative on firm's promotion decision before the bonus (-0.047, p-value=0.79), while it becomes incrementally informative in explaining employee's chance of promotion after the bonus (0.387, p-value=0.06) for promotion decision from Level 1 to 3. The change in pattern observed in lower level promotion decision further echoes the message in Bol and Leiby (2018), who finds that supervisors tend to diverge their assessment on employee's potential as opposed to current performance for employees eligible for promotion. Speaking to this study, lower levels employees are more eligible for promotion since they have more room to "climb up the ladder", and thus they are the ones whose prospect of future potential may considerably differ from their current performance compared to the higher level employees. This set of result further sheds light on my prediction that once the performance based bonus is present, supervisor appraisal indeed becomes a more precise metric in gauging employee's future potential.

(insert Table 6 about here)

3.4.5 Discussion of Empirical Results

In this part I discuss the implication of the collective evidence from the previous empirical analysis.

[1] Consistent with the finding in Grabner and Moers (2013), objective performance measure significantly predicts employees' chance of promotion at lower levels but not at higher levels (coefficient on *log(Rev)* equals 0.635, p-value=0.01 for promotion below Level 3, while the same coefficient equals 0.056, p-value=0.53 for promotion above Level 2), and that for higher levels of promotion decisions supervisor's appraisal on employee ability plays a major role (F-stat of *Meanrating* equals 8.11 for Promotion above Level 2, significant at 99%).

After introduction of the bonus, supervisor's evaluation on the multiple dimensions of competence for each employee becomes more dispersed, which suggests supervisors now bother to differentiate across these dimensions that are supposed to capture different aspects of employee's underlying competence. The mean of appraisal scores remains the same as before, however, when taking into account that the employee's financial performance increased after introduction of the bonus, if supervisors keep the way they provide the appraisal, one shall expect an increased rather than unchanged mean rating from the supervisors. The fact that *Meanrating* is the same as before suggests supervisors do not provide their performance appraisal the same way as they used to do after the bonus, nor does the result show a learning effect in supervisor appraisal. This is in line with the reduced association of supervisor appraisal with financial performance after introduction of the bonus. The reduced association as well as the increased dispersion in dimensions of appraisal jointly speaks to a documented reduction of halo effect: That instead of assigning a similar set of rating to a series of evaluation based on a general impression (which most likely is the employee's achievement on financial performance), the supervisor now is able to better distinguish among the different aspects of employee's overall competence (Thorndike 1920, Jacobs and Kozlowski 1985). As a consequence, supervisor appraisal instead picks up employee's overall contribution that resides outside of information contained in financial performance measures. Further analysis finds that supervisor appraisal becomes more predictive of employee's future financial performance, and that the evaluation is put more weight by the firm when making promotion decisions regarding an employee. These findings lend support to the hypothesis I developed before. The empirical evidence collectively speaks to the main message I put forward, that the introduction of a performance based bonus grants the supervisor with a better instrument to represent her true assessment of employee's underlying human capital. Overall, the set of performance measures becomes more precise, which allows the organization to more effectively identify and develop employee's underlying human capital for selection purpose.

I further provide a breakdown of the mean value of performance indicators and performance appraisal over the sample period in Table 7. We can see that performance indicators, such as revenue, employee's net contribution, and customer satisfaction rating generally show an continuous increase over the five years. Since the bonus is introduced universally to all employees in the company, the absence of a control group prevents me from empirically identifying how much of the effect is attributed to employees' increased effort

after having the bonus, or from the natural expansion of the business. Nonetheless, we can see that for performance outcome *CS*, which is rated on a fixed scale and therefore exists no room of inflation, also shows a significant increase after introduction of the bonus (t-stat=2.69). This lends some support that the observed increase in performance outcome is at least partially attributed to the introduction of the bonus. More importantly, we can see that performance appraisal provided by supervisors do not follow the same pattern as standalone performance indicators, there is no significant different between the *Meanrating* before and after introduction of the bonus (t-stat=-0.51). Similarly, employee's chance of promotion does not significantly increase along with the performance indicators (t-stat=-0.63). This joint evidence addresses the potential concern that during prosperous times, the firm provides appraisals more favorably or is generally less restrictive in granting promotions.

(insert Table 7 about here)

3.5 Additional Analysis and Robustness Checks

In this session I describe a set of additional analysis to facilitate further understanding of the institutional context and to validate the empirical result.

3.5.1 Effect of Supervisor Appraisal on Employee Career and Compensation

As discussed along with the research setting, one may argue that on top of reflecting supervisor's true assessment of employee competence, supervisor's appraisal also functions to influence firm's ex post adjustment of the bonus employees get. To address this potential concern, I provide empirical evidence of the determinants of different components of employee compensation and career outcome to facilitate further understanding of the setting. The result is presented in Table 8. We can see that supervisor appraisal has significant and most dominant influence on employee's salary increase, both before and after introduction of the bonus, while it does not have a significant consequence on the bonus each employee gets, neither nominally nor in reality. In a related paper studying the same research setting (Bouwens and Jiang WP), we predict and find that the firm's ex-post adjustment to the formula-based bonus serves to complement the performance measures in the conditions where the objective compensation system is likely to be incomplete in recognizing and rewarding employee's effort, rather than their underlying competence revealed in the appraisal. This descriptive analysis lends further support that the formula-based bonus is granted to reward

employees for their contemporaneous effort. Column (3) provides evidence that employee's turnover in the next year is negatively associated with their current appraisal outcome (-0.10-0.03=-0.13, F-stat=5.39), which lends support to the assumption that absent of a performance-based bonus, supervisors incorporate employee's performance outcomes in their appraisal to account for employee retention.

(insert Table 8 about here)

3.5.2 Alternative Purpose of Supervisor Appraisal

The findings in my study generally points to the idea that objective performance measures pick up employee's effort whereas supervisor appraisal reflects employee's forward looking, persistent characteristics. One may argue that supervisors may use their appraisal to influence ex-post adjustment of employee's bonus in order to better recognize their exerted effort. From interviews with the firm, I was informed that the supervisor, when reviewing the employee's performance and deciding on their bonus, they may make advice to the upper level in the firm their observations more fairly compensate the employees for they have done, and there is no need to signal this piece of information through their appraisal. In a similar vein, if the supervisor appraisal is to adjust what is reflected in objective performance measures so that the employees feel more fairly compensated, we may generally believe that the pay gap between employees will be mitigated after introduction of the bonus. I provide primary descriptive evidence that neither horizontal (standard deviation of total compensation of employees at the same job level) nor vertical (difference in average total compensation across job levels) pay gap in the firm is mitigated, but instead they both exacerbated after introduction of the bonus (result not tabulated). There is thus little evidence to believe that supervisors use their appraisal to mitigate pay gap among employees, and this lends further support that the purpose of appraisal is to provide information incremental to financial performance measures and rating from the clients.

3.5.3 Alternative Measure of Supervisor Appraisal

I use the mean score across all dimensions of competence as supervisor appraisal in all my main tests, alternatively, I construct the latent variable "*Appraisal*" as the underlying factor of all dimensions of competence scores using systematic equation modeling, and this method also allows me to account for the missing values for observations where not all dimensions of appraisal is available. Results using *Appraisal* instead of *Meanrating* are consistent in all analysis.

3.5.4 Effect Over Time and Cross-sectional Evidence

I further examine for all my main findings, the effect over time and cross-sectional variation among different employees. I tabulate results of the test for *Dispersion* as an illustration in Table 9. Similar analysis with all other main tests are generally consistent with the pattern shown in *Dispersion*. Column (1) reproduces the regression result of *After* on *Dispersion* with all control variables, the same as in Table 3 column (4). It indicates supervisor's appraisal on multiple dimensions for each employee becomes more dispersed during the after bonus period. To see more specifically when the treatment starts to take effect, I present the regression result doing the same test but replacing variable *After* by year dummies in column (2). We can find that *Dispersion* does not show any significant change in 2005 compared to 2004 (0.021, p-value=0.51), while it starts to significantly increase immediately since the first year of introducing the bonus (i.e., year 2006) and on. This result addresses the concern that the effect I find during the *After* period is driven by any abnormality taken place in year 2005. Furthermore, one possible explanation for why the supervisor changed her behavior of rating is the supervisor's reciprocal action to spend more effort and give a more fair evaluation when she finds her employees are doing a better job (Maas et al. 2012). If this explanation dominates, one shall expect the observed change in supervisor appraisal materializes over time and occurs at least one year after introducing the bonus, while the result in column (2) that *Dispersion* gets significantly larger starting in year 2006 suggests that reciprocal behavior is unlikely to be the dominant explanation for the observed change in supervisor appraisal. In other words, the result of effect over time lends support to the premise in this study that supervisors change their way of providing performance appraisal in response to the firm's introduction of bonus, rather than a behavior driven from the change in employee performance.

(insert Table 9 about here)

Similar as section 4.4, I also conduct cross-sectional analysis conditional on employee job levels. I separate the sample by observations from Level 3 and up versus those lower than Level 3, and run the main test separately for the two groups. The result of this analysis is presented in columns (3) and (4). We can see that the effect of introducing the bonus on supervisor's evaluation dispersion is mostly concentrated among the lower level employees (0.059, p-value=0.06). This finding is consistent with section 4.4, and is an illustration that the observed pattern among lower level employees are present across different analysis conducted. This cross-sectional evidence further convinces my findings that after the introduction of a performance based

bonus, performance appraisal becomes more informative in picking up the employee's overall contribution with a future prospect rather than employee's past performance.

3.6 Conclusion

In this study, I predict and document a series of evidence that supervisor's performance appraisal becomes more accurate in reflecting their true assessment of employee's underlying human capital after the firm adjusted the compensation system for its employees. I find in a professional service firm in the Netherlands, after the firm introduced a bonus to reward employees for their concurrent performance, supervisors tends to distinguish multiple dimensions of employees' competence beyond what is captured by their performance measures, and their appraisal gets more informative about an employee's on-going potential. I explore a condition where by granting supervisors with a better instrument to represent her true assessment of employee's dimensions of competence, the firm achieved the intended purpose of having more precise metrics to identify and develop human capital. This study provides empirical evidence that the adaption of management control package can facilitate supervisors to better deliver with regard to the organizational purpose, without changing any explicit incentive for them.

My study, as with others, admittedly has limitations. The personnel and performance measure data I have available are exclusively at employee level, I infer supervisor's way of providing appraisal from how it is reflected in the performance and personnel information of individual employees. This does not cause severe concern about the validity in my finding with regard to the research question I examine, however, the absence of personnel information of supervisor (e.g., supervisor's tenure in the firm, experience, etc.) prohibits me from drawing considerable inferences in terms of supervisor's personal preference such as relational concern or cost of gathering information (Bol 2011). Moreover, the main treatment in my study is the firm's introduction of a performance based bonus, while it is reasonable to suspect other changes in the meantime. For instance, the firm I study is in a natural period of business expansion, and the additional variable pay infers a more "competitive" working culture, after which not only the employees but also the supervisors step up and do a better job. I believe though such explanation is not mutually exclusive, and it is likely to have a second-order effect. In my study, I provide theory and empirical evidence to the extent

possible, that the change in supervisor's appraisal is at least partly a response to the adaption of compensation system.

References

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. Economics letters, 80(1), 123-129.

Bailey, W. J., Hecht, G., & Towry, K. L. (2011). Dividing the pie: The influence of managerial discretion extent on bonus pool allocation. Contemporary Accounting Research, 28(5), 1562-1584.

Baker, G., Gibbs, M., & Holmstrom, B. (1994). The internal economics of the firm: Evidence from personnel data. The Quarterly Journal of Economics, 109(4), 881-919.

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

Baker, G. P., Jensen, M. C., & Murphy, K. J. (1988). Compensation and incentives: Practice vs. theory. The journal of Finance, 43(3), 593-616.

Banker, R. D., Potter, G., & Srinivasan, D. (2000). An empirical investigation of an incentive plan that includes nonfinancial performance measures. The accounting review, 75(1), 65-92.

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

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

Bol, J. C., & Leiby, J. (2018). Subjectivity in Professionals' Incentive Systems: Differences between Promotion‐and Performance‐Based Assessments. Contemporary Accounting Research, 35(1), 31-57.

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

Campbell, D. (2008). Nonfinancial performance measures and promotion‐based incentives. Journal of Accounting Research, 46(2), 297-332.

Campbell, D., Datar, S. M., Kulp, S. L., & Narayanan, V. G. (2015). Testing strategy with multiple performance measures: Evidence from a balanced scorecard at Store24. Journal of Management Accounting Research, 27(2), 39-65.

Casas-Arce, P., Martínez-Jerez, F. A., & Narayanan, V. G. (2017). The impact of forward-looking metrics on employee decision-making: The case of customer lifetime value. The Accounting Review, 92(3), 31-56.

Deller, Carolyn. 2018. Identifying, Measuring, and Communicating Employee Fit Through Formal Control Mechanisms: Evidence From the Field. Doctoral dissertation, Harvard Business School.

Demeré, B. W., Sedatole, K. L., & Woods, A. (2018). The role of calibration committees in subjective performance evaluation systems. Management Science.

De Pater, I. E., Van Vianen, A. E., Bechtoldt, M. N., & KLEHE, U. C. (2009). EMPLOYEES'CHALLENGING JOB EXPERIENCES AND SUPERVISORS'EVALUATIONS OF PROMOTABILITY. Personnel Psychology, 62(2), 297-325.

Du, F., Tang, G., & Young, S. M. (2012). Influence activities and favoritism in subjective performance evaluation: Evidence from Chinese state-owned enterprises. The Accounting Review, 87(5), 1555-1588.

Ederhof, M. (2010). Discretion in bonus plans. The Accounting Review, 85(6), 1921-1949.

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

Gibbons, R. (1998). Incentives in organizations. Journal of economic perspectives, 12(4), 115-132.

Gibbons, R., & Waldman, M. (1999). Careers in organizations: Theory and evidence. Handbook of labor economics, 3, 2373-2437.

Gibbons, R., & Waldman, M. (1999). A theory of wage and promotion dynamics inside firms. The Quarterly Journal of Economics, 114(4), 1321-1358.

Gibbs, M. (1995). Incentive compensation in a corporate hierarchy. Journal of Accounting and Economics, 19(2-3), 247-277.

Gibbs, M., & Hendricks, W. (2004). Do formal salary systems really matter?. ILR Review, 58(1), 71-93.

Gibbs, M. (2008). Discussion of nonfinancial performance measures and promotion‐based incentives. Journal of Accounting Research, 46(2), 333-340.

Grabner, I., Künneke, J., & Moers, F. (2016, August). How to mitigate bias in performance evaluations: An analysis of the consequences of supervisors' evaluation behavior. AAA.

Grabner, I., Künneke, J., & Moers, F. (2018, August). Identifying Talent: An Exploratory Investigation of Explicit Potential Assessment. AAA.

Grabner, I., & Moers, F. (2013). Managers' choices of performance measures in promotion decisions: An analysis of alternative job assignments. Journal of Accounting Research, 51(5), 1187-1220.

Hayes, R. M., & Schaefer, S. (2000). Implicit contracts and the explanatory power of top executive compensation for future performance. The RAND Journal of Economics, 273-293.

Ittner, C. D., Larcker, D. F., & Meyer, M. W. (2003). Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. The accounting review, 78(3), 725-758.

Jacobs, R., & Kozlowski, S. W. (1985). A closer look at halo error in performance ratings. Academy of Management Journal, 28(1), 201-212.

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

Maas, V. S., Van Rinsum, M., & Towry, K. L. (2011). In search of informed discretion: An experimental investigation of fairness and trust reciprocity. The Accounting Review, 87(2), 617-644.

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

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

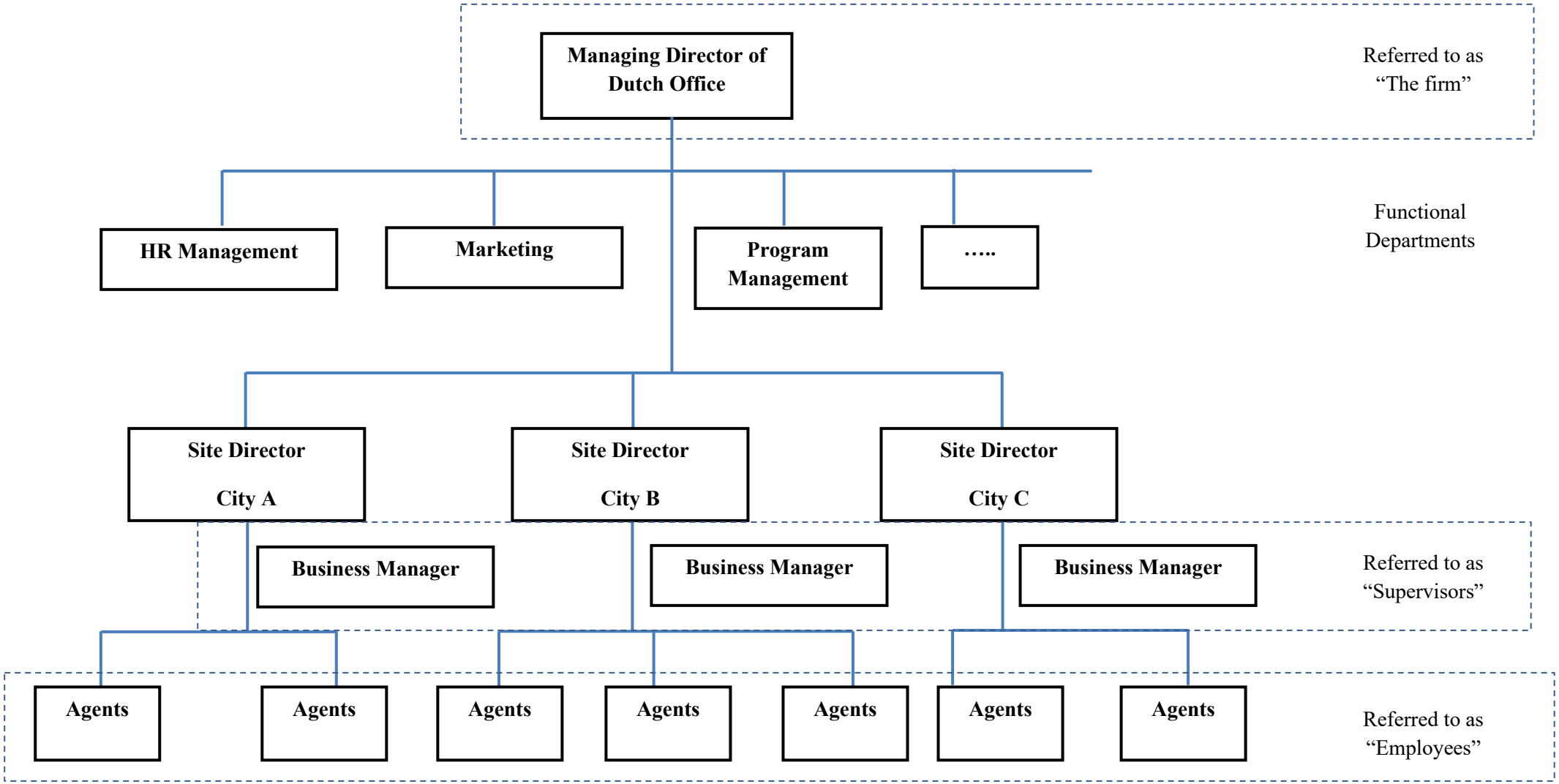
Norton, E. C., Wang, H., & Ai, C. (2004). Computing interaction effects and standard errors in logit and probit models. The Stata Journal, 4(2), 154-167.

Prendergast, C. (1993). The role of promotion in inducing specific human capital acquisition. QUARTERLY JOURNAL OF ECONOMICS-CAMBRIDGE MASSACHUSETTS-, 108, 523-523.

Prendergast, C. (1999). The provision of incentives in firms. Journal of economic literature, 37(1), 7-63.

Thorndike, E. L. (1920). A constant error in psychological ratings. Journal of Applied Psychology, 4(1), 25-29.

Appendix A: Organizational Structure of FM



Appendix B: Description of Pay Determinants in FM

B.1 Definition of Performance Indicators

Objective Performance Indicators:

Revenue:	Total revenue an employee makes for the firm by the end of the year
NetContri:	Net contribution the employee makes for the firm, calculated as (total revenue - 24 months gross salary - cost of travel and other miscellaneous costs)/total revenue
CS:	Customer satisfaction, clients are provided a survey by FM asking about their general satisfaction with the corresponding employee's work, with 1 being least satisfied and 6 most satisfied. If an employee serves more than one clients that year, the CS is the average of all clients satisfaction rating he serves, rounded at each 0.5

Appraisal Provided by Supervisors (all scaled from 1 to 6)

Customer Focus:	Investigate customer wishes and needs and act accordingly. Anticipating customer needs and giving high priority to good service and customer satisfaction.
Communication:	Making things clear to others with the help of clear language, gestures and non-verbal communication. Adjust language and terminology to the target group. To clarify matters in a report or document that has the right set-up and structure.
Result Orientation:	Systematically define goals and priorities and indicate required actions, tasks, time and resources to achieve results. Ensure a structured approach to work. Regularly test the realized progress against the result to be achieved.
Flexibility:	Respond effectively to changing circumstances, as well as to people with different opinions, visions, expectations and questions. Change style and approach if necessary.
Expertise:	Raising trust among others based on their own knowledge and skills, integrity, representativeness and interpersonal sensitivity and convincing others.
Proactive behavior:	Act before anything is asked and before circumstances force action. Seek opportunities to take action.
Analyzing/Advising:	Analyze a problem, a situation or a process and on this basis give an opinion, advice and / or own vision.

B.2 Uniform Formula in Determining Salary Increase

Salary increase in the Netherlands is reflected as increase in number of salary steps. Within each level of job, there are twenty (20) steps of gross salary. FM uses a pre-set, uniform formula to determine number of salary steps each employee progresses, and the corresponding actual amount of salary is consistent with firms in the same industry across the Netherlands.

In FM, the increase in salary step is determined with a score denoted as **Sscore**:

Sscore = 0.6 * *Meanrating* + 0.15 * **NCscore** + 0.15 * *CS* + 0.1 * *Other*

Where NCscore is the score determined with net contribution (*NetContri*), with the following correspondence:

NetContri	**NCscore**
NC>40%	6
30%<NC<=40%	5
20%<NC<=30%	4
10%<NC<=20%	3
0<NC<=10%	2
NC<=0	1

Once the Sscore is calculated, the predicted steps of salary increase is determined as:

Total Sscore	Steps of Salary Increase
S<=3.4	0
3.4<S<=4.1	1
4.1<S<=4.4	2
4.4<X<=4.7	3
S>4.7	4

B.3 Uniform Formula in Determining Annual Bonus

Since 2006, FM introduced an annual bonus for all employees, determined by a uniform, pre-set formula. Similar as salary increase, the bonus is also determined with a score denoted as **Bscore**, times the bonus base which varies among different levels of employees.

Bscore = 0.7 * **NCscore** + 0.3 * *CS*

Where NCscore is calculated the same as discussed in B.2, and employee's bonus base is as follows:

Level	Bonus base
1	0.5 monthly salary
2	0.5 monthly salary
3	1 monthly salary
4	1.5 monthly salary
5	2 monthly salary

Appendix C Variable Definition

Age	Age is measured by the years between an employee's year of born and the current year.
Experience	Experience is the total number of years an employee has been stayed in the financial service industry.
Tenure	Tenure measures the years since an employee's inception at FM till when he leaves or till present.
Level	Level represents different job levels in FM, with 1 being the lowest and 5 the highest.
Lvduration	Lvduration measures the number of years an employee has been in the current job level.
Gender	Binary variable that takes a value of 1 for a male employee and 0 for a female employee.
Education	Binary variable that takes the value of 1 if an employee has a research university degree or beyond, and 0 if he has an applied college degree.
Log(Rev)	The natural logarithm of annual revenue the employee makes scaled by number of working months.
NetContri	Employee's net contribution calculated as annual revenue minus two times total salary and other labor costs scaled by revenue.
CS	The mean customer satisfaction rating from all client(s) the employee serves each year.
Promotion	Promotion takes the value of 1 if an employee progresses to a higher level of job, and 0 otherwise.
Meanrating	The average score across all seven dimensions of competence for each employee-year observation.
Dispersion	The standard deviation of the scores across all dimensions of competence for each employee-year observation.
After	After takes a value of 1 for observations in years 2006, 2007, and 2008, and 0 for observations in years 2004 and 2005.
1stYear	Binary variable equals 1 if it is the first year the employee in the firm, and 0 otherwise.

Table 1 Descriptive Statistics of Main Variables

This table presents the descriptive statistics of variables used in the main analysis. The data is tracked annually in FM's personnel system, ranges from year 2004 to 2008. Panel A shows summary statistics of employee demographic characteristics. Panel B shows summary statistics of employee's performance indicators. Detailed variable definition is presented in Appendix C.

Panel A: Employee Demographic Characteristics

	N	mean	min	p25	median	p75	max	sd
age	502	32.29	21	26	31	37	54	7.15
experience	498	9.25	1	4	8	12	34	7.07
tenure	502	2.82	0.05	0.92	1.95	4.25	10.68	2.42
level	502	2.73	1	2	3	4	5	1.14
lvduration	502	2.43	0.05	0.92	1.58	3.08	10.67	2.29
gender	502	0.73	0	0	1	1	1	0.45
education	502	0.26	0	0	0	1	1	0.44

Panel B: Performance Indicators

	N	mean	min	p25	median	p75	max	sd
log(Rev)	502	9.04	6.59	8.76	9.06	9.35	11.13	0.43
NetContri	502	0.21	-3.54	0.16	0.26	0.33	0.62	0.26
CS	502	4.53	3	4	4.5	5	6	0.54
Promotion	502	0.09	0	0	0	0	1	0.29
meanrating	502	4.31	2.86	4.07	4.29	4.5	5.57	0.33
dispersion	502	0.34	0	0.22	0.36	0.45	0.88	0.18

Table 2 Correlation Table of Main Variables

This table presents the correlation metrics of variables summarized in Table 1. Variables are defined in Appendix C. P-values are below each respective correlation.

	log(rev)	*NetContri*	*CS*	*meanrating*	*dispersion*	*age*	*experience*	*tenure*	*level*	*lvduration*	*gender*	*education*
log(rev)	1											
NetContri	0.3903	1										
	0.0000											
CS	0.0712	0.0417	1									
	0.1112	0.3517										
meanrating	0.0540	0.1485	0.4597	1								
	0.2274	0.0008	0.0000									
dispersion	0.0426	0.0256	0.2146	0.2693	1							
	0.3410	0.5678	0.0000	0.0000								
age	0.4990	-0.1620	0.0274	-0.0919	0.0258	1						
	0.0000	0.0003	0.5401	0.0397	0.5641							
experience	0.4626	-0.1708	0.0597	-0.0527	0.0256	0.9124	1					
	0.0000	0.0001	0.1835	0.2408	0.5684	0.0000						
tenure	0.1028	-0.0976	0.1356	0.0486	0.0909	0.3534	0.3294	1				
	0.0213	0.0288	0.0023	0.2774	0.0419	0.0000	0.0000					
level	0.6530	-0.1359	0.0885	0.0104	0.0701	0.8009	0.7571	0.3191	1			
	0.0000	0.0023	0.0474	0.8162	0.1166	0.0000	0.0000	0.0000				
lvduration	0.1133	-0.0995	0.0618	0.0049	0.0598	0.4204	0.3935	0.8967	0.2931	1		
	0.0000	0.0259	0.1667	0.9134	0.1812	0.0000	0.0000	0.0000	0.0000			
gender	0.3004	0.0205	-0.0047	0.0296	-0.0036	0.2169	0.1895	0.1022	0.2195	0.0978	1	
	0.0000	0.6470	0.9162	0.5083	0.9361	0.0000	0.0000	0.0220	0.0000	0.0284		
education	0.1642	0.0665	-0.0723	0.0824	0.1405	-0.0096	-0.0289	-0.0757	0.1226	-0.0852	-0.0129	1
	0.0002	0.1370	0.1057	0.0649	0.0016	0.8303	0.5201	0.0901	0.0059	0.0564	0.7737	

Table 3 Mean and Dispersion of Supervisor Appraisal

This table presents the empirical result of the following regression:

$$Pattern_{i,t}=\alpha+\beta 1*After_t+\gamma*Controls+\epsilon_{i,t} \quad (1)$$

Where *Pattern* is represented by *Meanrating* and *Dispersion*, *Meanrating* is the average score across all seven dimensions of competence for each employee-year observation. *Dispersion* is the standard deviation of the scores across all dimensions of competence for each employee-year observation. *After* takes a value of 1 for observations in years 2006, 2007, and 2008, and 0 for observations in years 2004 and 2005. Control variables are defined in Appendix C. *Level* and *Location* fixed effects are included in columns (2) and (4). P-value is shown in parenthesis, *, **, *** represents result significant at 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	**Meanrating**	**Meanrating**	**Dispersion**	**Dispersion**
after	-0.017	-0.003	0.039**	0.037**
	(0.61)	(0.92)	(0.04)	(0.05)
tenure		-0.007		0.007
		(0.64)		(0.43)
lvduration		0.003		-0.004
		(0.84)		(0.68)
age		-0.085***		0.020
		(0.00)		(0.15)
age2		0.001***		-0.000
		(0.00)		(0.12)
education		0.068*		0.049**
		(0.05)		(0.01)
gender		0.047		-0.011
		(0.15)		(0.55)
1stYear		-0.180***		-0.051**
		(0.00)		(0.02)
Constant	4.322***	5.852***	0.310***	0.023
	(0.00)	(0.00)	(0.00)	(0.92)
Level FE	No	Yes	No	Yes
Location FE	No	Yes	No	Yes
Observations	502	502	502	502
R-squared	0.001	0.136	0.009	0.070

Table 4 Association of Supervisor Appraisal with Objective Performance Measures

This table presents the empirical result of the following regression:

$$Meanrating_{i,t}=\alpha+\beta1*log(Rev)_{i,t}+\beta2*After_t+\beta3*log(Rev)*After_{i,t}+\gamma*Controls+\epsilon_{i,t} \quad (2)$$

Meanrating is the average score across all seven dimensions of competence for each employee-year observation. *After* takes a value of 1 for observations in years 2006, 2007, and 2008, and 0 for observations in years 2004 and 2005. Variables are defined as Appendix C. *Level* and *Location* fixed effects are included in columns (3). The row F-stat presents the test of joint significance of b1 and b3. P-value is shown in parenthesis, *, **, *** represents result significant at 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)
VARIABLES	**Meanrating**	**Meanrating**	**Meanrating**
log(Rev)	0.177**	0.140**	0.154**
	(0.01)	(0.03)	(0.03)
after	1.520**	1.166*	1.263*
	(0.04)	(0.08)	(0.06)
*after*log(Rev)*	-0.173**	-0.147**	-0.146**
	(0.03)	(0.04)	(0.04)
CS		0.271***	0.258***
		(0.00)	(0.00)
*after*CS*		0.018	-0.003
		(0.76)	(0.96)
tenure			-0.017
			(0.24)
lvduration			0.012
			(0.39)
age			-0.049**
			(0.03)
age2			0.001*
			(0.08)
education			0.083***
			(0.01)
gender			0.044
			(0.15)
1stYear			-0.127***
			(0.00)
Constant	2.744***	1.877***	2.739***
	(0.00)	(0.00)	(0.00)
Level FE	No	No	Yes
Location FE	No	No	Yes
F-stat (b1+b3)	0.01	0.03	0.04
F-stat (b4+b5)	-	180.33***	80.71***
Observations	502	502	502
R-squared	0.013	0.225	0.297

Table 5 Supervisor Appraisal in Predicting Future Performance

This table presents the empirical result of the following regression:

$$Log(Rev)_{i,t+1}=\alpha+\beta1*Meanrating_{i,t}+\beta2*After_t+\beta3*Meanrating*After_{i,t}+\gamma*Controls+\epsilon_{i,t} \quad (3)$$

Log(Rev) is the natural logarithm of annual revenue the employee makes scaled by number of working months, and $Log(Rev)_{i,t+1}$ is the revenue the employee makes in the following year if an employee stays in the firm. Variables are defined in Appendix C. *Level* and *Location* fixed effects are included in columns (2). The row F-stat presents the test of joint significance of b1 and b3. P-value is shown in parenthesis, *, **, *** represents result significant at 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)
VARIABLES	**logRev(t+1)**	**logRev(t+1)**	**logRev(t+1)**
logRev(t)	0.606***	0.605***	0.277***
	(0.00)	(0.00)	(0.00)
meanrating(t)	-0.072	-0.127	-0.038
	(0.36)	(0.14)	(0.64)
after	-0.743*	-0.545	-0.489
	(0.09)	(0.24)	(0.24)
*after*meanrating(t)*	0.188*	0.256**	0.209**
	(0.06)	(0.02)	(0.04)
CS(t)		0.092	0.028
		(0.12)	(0.60)
*after*CS(t)*		-0.112	-0.067
		(0.15)	(0.34)
tenure			0.014
			(0.43)
lvduration			-0.013
			(0.47)
age			-0.017
			(0.52)
age2			0.000
			(0.62)
education			0.050
			(0.20)
gender			0.172***
			(0.00)
1stYear			0.101**
			(0.03)
Constant	3.901***	3.740***	6.468***
	(0.00)	(0.00)	(0.00)
Level FE	No	No	Yes
Location FE	No	No	Yes
F-stat (b1+b3)	3.44*	3.23*	6.48**
Observations	291	291	291
R-squared	0.484	0.489	0.622

Table 6 Firm's Use of Supervisor Appraisal in Promotion Decision

This table presents the empirical result of the following regression:

$$Promotion_{i,t+1}=\alpha+\beta1*Meanrating_{i,t}+\beta2*After_t+\beta3*Meanrating*After_{i,t}+\gamma*Controls+\epsilon_{i,t} \quad (4)$$

$Promotion_{i,t+1}$ takes the value of 1 if in the following year an employee progresses to a higher level of job indicated by change of job title, and 0 otherwise. Variables are defined in Appendix C. All columns present result of conducting OLS regressions. *Level* and *Location* fixed effects are included in all columns. The row F-stat presents the test of joint significance of b1 and b3 in column (2). P-value is shown in parenthesis, *, **, *** represents result significant at 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)
	All	*Level<3*	*Level>2*
VARIABLES	**Promotion**(t+1)	**Promotion**(t+1)	**Promotion**(t+1)
meanrating	0.088	-0.047	0.157
	(0.36)	(0.79)	(0.11)
after	-0.722	0.676	-1.094
	(0.42)	(0.77)	(0.29)
*after*meanrating*	0.193*	0.387*	0.084
	(0.10)	(0.06)	(0.48)
log(Rev)	0.261***	0.635***	0.056
	(0.00)	(0.01)	(0.53)
*after*log(Rev)*	-0.016	-0.274	0.077
	(0.86)	(0.29)	(0.47)
CS	0.029	0.014	0.013
	(0.52)	(0.88)	(0.77)
tenure	-0.041*	-0.038	-0.024
	(0.08)	(0.62)	(0.23)
lvduration	0.016	0.006	0.007
	(0.48)	(0.94)	(0.72)
age	0.093***	0.358***	0.089**
	(0.01)	(0.00)	(0.01)
age2	-0.001**	-0.005***	-0.001**
	(0.01)	(0.00)	(0.02)
gender	-0.037	-0.077	0.023
	(0.42)	(0.32)	(0.66)
education	0.055	-0.065	0.086*
	(0.26)	(0.57)	(0.06)
1stYear	-0.149***	-0.426***	0.007
	(0.01)	(0.00)	(0.91)
Constant	-3.817***	-9.883***	-2.930***
	(0.00)	(0.00)	(0.01)
Level FE	Yes	Yes	Yes
Location FE	Yes	Yes	Yes
F-stat (b1+b3)	12.18***	5.75**	8.11***
Observations	291	121	170
R-squared	0.285	0.386	0.250

Table 7 Breakdown of Performance Indicators Overtime

This table presents the mean value of performance indicators and appraisals separately for every year from 2004 to 2008. Column "diff" in the right presents the t-stats of the difference between each variable in the before period (year 2004 and 2005) and in the after period (year 2006, 2007, and 2008), *, **, *** represents result significant at 10%, 5%, and 1% level, respectively. Variables are defined in Appendix C.

	After=0		*After=1*			diff (t-stat)
	2004	**2005**	**2006**	**2007**	**2008**	
n	60	70	100	126	146	
log(Rev)	8.893	8.903	9.041	9.055	9.148	4.36***
NetContri	0.120	0.167	0.218	0.245	0.235	3.31***
CS	4.408	4.432	4.556	4.607	4.542	2.69***
Meanrating	4.354	4.295	4.349	4.390	4.200	-0.51
Promotion	-	0.2	0.02	0.103	0.123	-0.63

Table 8 Effect of Supervisor Appraisal on Employee Career and Compensation

This table presents the consequence of supervisor appraisal on employee's annual bonus and monthly salary increase, and turnover in the following year. *PayInc(t+1)* is the difference of employee's monthly salary in January in the following year compared with salary in the current January or the first month available in the current year, for observations where the employee stays with the firm. *Bonus(t)* is the employee's annual bonus at the end of year, observations with bonus are only available after 2006 (i.e., when *After*=1). The binary variable *Turnover (t+1)* equals 1 if the employee leaves the firm in the following year, and 0 if he stays. Independent variables as defined as before. *Level* and *Location* fixed effects are included in both columns. The row F-stat presents the test of joint significance of b1 and b3. P-value is shown in parenthesis, *, **, *** represents result significant at 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)
VARIABLES	**PayInc**(t+1)	**Bonus**(t)	**Turnover**(t+1)
meanrating	73.57**	371.05	-0.10
	(0.05)	(0.19)	(0.14)
after	26.26	-	0.18
	(0.89)	-	(0.62)
*after*meanrating*	-2.92	-	-0.03
	(0.95)	-	(0.69)
NetContri	42.34	1,465.48***	0.00
	(0.11)	(0.00)	(0.98)
CS	18.84	230.75	0.02
	(0.28)	(0.17)	(0.47)
tenure	12.91	22.19	0.01
	(0.14)	(0.78)	(0.73)
lvduration	-12.47	82.69	-0.00
	(0.16)	(0.31)	(0.94)
age	14.80	207.31	-0.03
	(0.25)	(0.12)	(0.27)
age2	-0.28	-2.75	0.00
	(0.10)	(0.12)	(0.21)
gender	3.84	265.53	-0.02
	(0.83)	(0.14)	(0.49)
education	34.87*	-30.97	0.00
	(0.06)	(0.87)	(0.94)
1stYear	-47.07**	-1,361.41***	0.01
	(0.03)	(0.00)	(0.89)
Constant	-417.31	-4,232.35	0.90*
	(0.14)	(0.11)	(0.09)
Level FE	Yes	Yes	Yes
Location FE	Yes	Yes	Yes
F-stat (b1+b3)	5.23**	-	5.39**
Observations	291	369	291
R-squared	0.250	0.765	0.098

Table 9 Effect Over Time and Cross-sectional Evidence

This table presents the treatment effect of introducing the bonus over time as well as cross-sectional evidence conditional on employee level. Variables are defined as before. *Level* and *Location* fixed effects are included in all columns. Result in column (1) is the same as in Table 3 column (4). Column (2) presents result of replacing *After* with year dummies, where *year2004* dropped because of multicollinearity. Column (3) presents result of test in (1) for the sub-group of observations below level 3, and column (4) presents result of same test for the remaining observations. P-value is shown in parenthesis, *, **, *** represents result significant at 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)
	Main test	*Year dummies*	*Level<3*	*Level>2*
VARIABLES	**Dispersion**	**Dispersion**	**Dispersion**	**Dispersion**
after	0.037**		0.059*	0.018
	(0.05)		(0.06)	(0.45)
year2005		0.021		
		(0.51)		
year2006		0.054*		
		(0.07)		
year2007		0.054*		
		(0.06)		
year2008		0.039		
		(0.17)		
tenure	0.007	0.007	0.012	0.005
	(0.43)	(0.41)	(0.65)	(0.59)
lvduration	-0.004	-0.004	-0.011	-0.004
	(0.68)	(0.64)	(0.68)	(0.66)
age	0.020	0.020	0.048	0.016
	(0.15)	(0.14)	(0.21)	(0.39)
age2	-0.000	-0.000	-0.001	-0.000
	(0.12)	(0.11)	(0.24)	(0.34)
education	0.049**	0.049**	-0.004	0.080***
	(0.01)	(0.01)	(0.90)	(0.00)
gender	-0.011	-0.012	-0.030	0.012
	(0.55)	(0.53)	(0.29)	(0.64)
1stYear	-0.051**	-0.051**	-0.012	-0.092***
	(0.02)	(0.02)	(0.75)	(0.00)
Constant	0.023	0.005	-0.429	0.088
	(0.92)	(0.98)	(0.45)	(0.81)
Level FE	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes
Observations	502	502	213	289
R-squared	0.070	0.072	0.052	0.114

Chapter 4 Seniority and Use of Subjectivity in Performance Evaluation

4.1 Introduction

In this paper, we study whether the principal uses different degrees of ex-post discretion in her bonus decision among employees with different seniority in the firm to account for effort that insufficiently surfaces in the objective performance measure in a setting where the same performance measurement system applies to all employees. By providing empirical evidence that the firm makes ex post adjustment to formulated bonus to a larger extent for more senior employees, we document that firms exert discretion in incentive contract when the pre-set, universal contract is more likely deemed imperfect in picking up employee's overall contribution to the company.

Firms typically rely on performance measures specified in a formally communicated (i.e., written) contract to assess the extent to which the agent exerts effort in a desired way, and reward the agent accordingly (e.g., Holmstrom 1979, Milgrom and Roberts 1992 p.221-p.224, Feltham and Xie 1994). Such performance measure however, can fall short in the sense that it incompletely picks up the agent's overall effort choice that is in the best interest of the firm, when it is insensitive to the agent's effort, provides a noisy signal of effort (Feltham and Xie 1994, Datar et al. 2001), or when the specified performance measure does not induce the agent to exert effort congruent with the firm's goal (Baker 2002). Subjective incentive compensation, on the other hand, when incorporated with objective performance measures, may mitigate the limitations with the objective and explicit compensation contract (Baker et al. 1994, Gibbs et al. 2004, Bol 2008, Hoppe and Moers 2011). Principal's use of subjectivity is beneficial when it is difficult or inefficient to contract on all dimensions of the agent's effort choice, and such a condition becomes prominent as an agent gets more senior in the company. Seniority is often interpreted as moving up in the hierarchy of an organization (e.g., from sales representative to director of sales and marketing). However, degree of seniority also varies within the same job and position as the employee stays longer in the firm, to the extent that more senior employees coach their junior peers, or that they take care of essential coordinative tasks as they develop more oversight given the many years of accumulated experience in the respective job. Note that these senior activities are fairly difficult, if possible at all, to be specified ex ante in an incentive contract. In other words, such coaching or coordinative activities benefit the firm but are not likely to be reflected in an employee's own

current or future performance given the uniform performance measurement system applied to all employees doing the formally same job. This would mean that senior employees are not evaluated or rewarded for their more senior execution of the same job, despite the fact that it is in the firm's best interest that senior employees conduct their job in this manner. Given that it is not possible for the firm to always fine-tune their (written) bonus contract, we put forward the idea that firms instead exert ex post discretion to adjust the compensation for more senior employees to account for their overall contribution to the firm that resides outside of their explicit performance measures.

Meanwhile, as an agent serves longer time in the company, the mutual trust between the agent and the principal has been built, it makes it possible and effective to deviate from the explicit contract based on ex ante specified performance measures. According to relational contract theory (Baker et al. 2002, Gibbons and Henderson 2012), when trust exists between two parties, the principal trusts that the agent's interest is better aligned with the firm and that he will act in the best interest of the firm; in turn, the agent would trust that the principal can fully observe and appreciate his devotion, and he therefore can choose the action that does not give very attractive outcome in short-term performance but is good for the firm's overall development. We therefore predict that the amount of subjective compensation the principal pays to the agent is an increasing function of the agent's seniority.

We conduct our study using personnel data and compensation information obtained from a professional service company in the Netherlands. The respective firm offers a formula-based incentive contract to their employees, where their annual bonus is determined by a pre-set formula comprised of a fixed percentage of revenue the employee generates and their customer satisfaction rating. We observe that however, more often than not, the firm deviates from this formula when offering bonus to the employees. We aim to explain why and when the firm decides to do so, and we believe that the formula provides a suitable benchmark to study the phenomenon of subjectivity in that we consider the formula outcome as the objective performance outcome, and the deviation resembles the principal's subjective evaluation of performance. We use this measure to assess whether the levels of subjectivity are identical over the career of the agent as represented by the time they have stayed with the company. We find compelling evidence that principals reward their senior agents subjectively to a larger extent than they reward junior agents subjectively. In addition, we also find that subjective evaluations as reflected by the firm's ex post adjustment to formula-based bonus are

more predictive of future performance of junior agents than of future performance of senior agents. The findings jointly speak to our prediction, that tasks performed by the senior agents span more performance areas than captured by the pre-set performance measure. The fact that we find the subjective compensation paid to a junior agent is more predictive to future performance than that paid to a senior agent indicates that the incompleteness of performance measure is the main driving force of the result. The adjusted portion of compensation is mostly paid to reward the non-contracted effort in the current period, and it is likely that the effort rewarded by subjective compensation is still not reflected in explicitly measured performance outcomes. By rewarding senior employees' performance with a larger extent of subjectivity, the firm encourages its senior employees to continuously make decisions in such a way that benefits the firm's overall performance. This finding suggests firms treasure the 'softer' capacity of agents gained from accumulated experience by compensating them with subjective adjustment to the contracted incentive plan.

Our study primarily contributes to literature on matters of subjective performance evaluation. Extant literature in subjective evaluation identifies, among others, an important role (ex post) subjective performance evaluation plays in incentive contracting, that it can shield the agents from risk by filtering out external factors beyond their control (e.g., Bol 2008, Hoppe and Moers 2011). Our study on the other hand, provides theoretical and empirical evidence that ex post adjustment to formulated bonus contract can also increase the congruence of incentive contract. We establish the idea we put forward that senior agents are arguably more likely to engage in value-adding activities beyond the explicit performance measures, and in the meantime the principal effectively recognizes the agent's effort that resides out of what is specified in the formal contract. By making greater ex post adjustment to the pre-set bonus formula for employees with longer tenure in the firm, the firm encourages senior employees to exert effort in the best interest of the firm, and thus enhances firm performance.

In a similar vein, the extant literature has investigated *why* subjectivity is used – in other words, they looked at the issue from the perspective of the economic determinants of the relative use of subjectivity versus objectivity in incentive compensation (e.g., Gibbs et al. 2004). Our study, however, focuses on *when* subjectivity comes into play. In this study we emphasize that the use of subjective compensation can be a dynamic process over an agent's career. Subjective compensation plays different roles to constitute the better incentive for the agent at different stages of his career, as well as the continuously developing objective of

the firm. Specifically, we establish that for principals and agents to feel comfortable with subjectivity, it requires that the parties have built a level of trust that can only come about over multiple periods of interaction between the principal and the agent.

In addition, existing empirical studies on this matter, such as Hayes and Schaefer (2000) and Ederhof (2010), rely on publically available data (e.g., proxy statements). Such data provides crude measures of subjectivity and the measures focus on the top executives of the firm only. However, our firm-specific dataset allows us to define a more accurate measure of subjectivity. This measure increases the validity in the light of testing our proposed theory. Adding to validity our data offers specific information about all levels of agents in the organization.

The remaining of the paper develops as follows: In section 4.2 we provide a literature review of related studies, and develop the theory and hypothesis of our study. In section 4.3 we introduce the data we have obtained and the research setting that is helpful to understand the empirical test of the theory. Empirical models and results are presented in section 4.4, where we provided evidence and discussions to empirically support the theories and hypothesis raised in the second section. In section 4.5, we presented a comprehensive discussion of the collective findings, and suggested how to properly interpret the results. Finally, in section 4.6 we conclude the paper and discuss a few issues that can be of interest to investigate in the future based on this study.

4.2 Theory and Literature Review

Firms design contracts in order to incentivize the agent to act in the best interest of the firm. The incentive contract typically entails performance measures that are reflective of the effort choice by the agents (e.g., Holmstrom 1979, Milgrom and Roberts 1992 p.221-p.224, Feltham and Xie 1994, Baker et al. 1988). A common practice is for firms to contract on standalone, objective performance outcome. Agents are in this setting compensated conditional on them achieving the targets set by the firm. Such a contract requires that the principal ex ante identifies the relevant performance measures and to set targets. The challenge the firm faces is to decide on the set of performances measures and the weighting across them. The composition of the set depends on how well the measure reflects the agent's achievement: (1) how sensitive and accurate the

set of performance measures reflect the achievements of the agent (e.g. Holmstrom, 1979; Banker and Datar 1989), and (2) whether the set of measures reflects the value created by the agent's achievement, i.e. is undistorted (Baker 2002).

However, it is highly unlikely that a set of performance measures fully meets all criteria at the same time, and the use of subjective assessment offers the opportunity in complementing the imperfections in performance measures (e.g., Baker et al. 1994; Baiman and Rajan 1995; Bol 2008; Hoppe and Moers, 2011). For instance, lack of sensitivity of a profit measure may lead to the situation where the agent did put in effort while his profit did not improve because of the dire condition the agent faced. In that case a standard that does not account for the economic conditions would make the measure insensitive to effort. The principal, however, may through direct observation of effort correct for substandard performance. In addition, the measure may move independently of the agent's effort, i.e. the measure is noisy (e.g. Bouwens and Van Lent, 2007). In that case the principal may also resort to direct observation to correct for measurement error. Finally, the measure may turn out to be distorted, i.e. not any improvement in the measure is reflective of value being created (Baker, 2002). In that case the principal may also decide to assess to what extend the agent actually created value with improving in the measure. Such a condition is likely to occur when the agent becomes more senior as he spends more time in the company, and that he may engage in value creating activities that benefits the firm but such contribution is not reflected in his own performance outcome.

Compared to a junior agent, a senior agent in the firm has accumulated knowledge and experience that is valuable to the firm. This experience extends beyond the tasks the agent is directly responsible for. That is, the whole firm may benefit from his extended experience, and the principal may for that reason expect the senior agent to take on responsibilities that extend beyond his direct influence such as coaching the junior agents or coordination tasks that are consistent with the firm's strategy. Therefore, the actions of senior agent may generate positive externalities in the firm that are likely to spill over to also affect the other agent's performance (Bushman et al. 1995, Bouwens et al. 2013). Firms may resort to using aggregate performance measures so that these externalities are reflected in the performance measure (Keating 1997, Bouwens and van Lent 2007). However, when measures are held constant, and the agent gets more senior as he stays longer time with the company, it becomes increasingly likely that measures become less sensitive to the effort of the agent. For instance, when the agent's performance is measured by his sales number, such

performance measure fails to reflect the extent to which the agent coaches his junior colleagues. If any, the time spent on coaching will impact on the attention that the senior agent can direct to his client, i.e. his performance that surfaces in the measures may suffer. His effort directed to activities that benefit the whole firm can, however, be picked up by the principal through direct observation of the agent or from company records like the minutes of meetings. While these observations are subjective, they may be necessary to assure that the (senior) agent contributes to firm performance in areas that go beyond his own client base. Moreover, as an agent's tenure increases, he is more aware of the overall strategic picture of the firm rather than simply his assigned task at hand. In this case, the actions the agent takes to boost the firm's future development may not be reflected in his performance measure at face value, and thus is reasonable to be compensated subjectively.

Following the above argument, we suggest that firms value the agent's experience that is not fully captured in the compensation formula, and we further argue that firms are able to induce such desired performance by using ex post compensation adjustment properly. Over the years an agent spent in a firm, mutual trust between the agent and the principal has been built, and according to relational contract theory, neither party has incentive to sacrifice the mutual trust and long-term profit for the short-sighted benefits (Baker et al. 2002, Gibbons and Henderson 2012). In the subjective compensation scenario, the principal believes the senior agent's interest is better aligned with the firm, and she can trust that the agent is able and willing to do good for the long-term development of the firm. The senior agent, in turn, can trust the principal to honor his performance in word and action. That is, he can be sure that contributions not picked up in the objective measures will be reflected in his performance assessment. We therefore expect that subjectivity will play a role in the assessment of senior agents.

In sum, we argue that the performance measures for senior agents do not resemble the set of effort they exert as much as those of the junior agents. The extent to which the principal would have to resort to subjective assessment in order to include all relevant performance areas is therefore smaller for juniors than for seniors. This suggests that subjectivity is less prominently present for juniors than for senior agents.

We thus summarize the above arguments in the following hypothesis:

H1: The extent to which the principal use subjectivity in agent's compensation increases with the agent's seniority in the firm.

4.3 Data and Research Settings

The personnel data and operational information used in this study is obtained from an international professional service firm. The headquarter is a listed company with subsidiaries in several west European countries, the data I use in this study is from its office in the Netherlands, anonymized as FM. The Dutch office is represented by a managing director, with several independent functional departments that take care of specific duties related to human resource management, marketing, project management, etc. For instance, the human resource department is responsible for all professional agents in the Dutch branch, and is in charge of all issues related to personnel management such as hiring, performance evaluation, compensation and promotion, among others. The business of FM mainly involves providing financial advice, internal control services, and at times consulting services for their corporate clients. All employees at FM have their college degree (or above) in finance or accounting related majors, and are affiliated with one of the three local business sites, each located in one of the Dutch cities.

In year 2006, FM introduced an annual bonus to all of its employees, on top of the monthly salary they get as compensation. The bonus is determined by a bonus formula universal to all the employees. Depending on the position the employee is currently holding at FM, a maximum individual bonus (usually a function of the employee's salary) is set as the base for his expected bonus. In addition, the firm has a function to determine the bonus score – the percentage as a multiplier to the maximal individual bonus – for each employee. The bonus score is jointly determined by the consultant's performance in customer satisfaction (with a weighting of 0.3) and his net contribution to the firm's profit (with a weighting of 0.7). In addition, the employees get extra bonus for successfully introducing new employees or clients to join FM. Specific description of the bonus formula is explained in Appendix B. The annual bonus that each employee gets has to refer to this identical bonus formula, as it is requested by the global headquarter. In the dataset we are able to calculate the predicted bonus with respect to the bonus formula based on employee's performance outcome. However, the actual bonus paid out to the employee is, more often than not, not equivalent to the predicted one. The difference between the actual reported bonus by FM and our predicted bonus data can be seen as the

subjective compensation, and this is how our dataset allows us to fully disentangle and examine the function of the ex post compensation adjustment. We put forward the idea that this ex post adjustment to predicted bonus is the alternative instrument the company resort to when recognizing an employee's effort beyond the specified performance measure.

In this study, we have a sample period starting from 2006 when bonus is first introduced, to year 2008. We keep the employee-year observations whose performance indicators and appraisals are available in the system. In addition, we also manually went through the data and eliminated employee-year observations that are deemed invalid with too much sick leave (more than 300 hours within a year) or where the employee switched from full-time contract to a part-time contract. This resulted in a final sample of 201 identical employees and in total 369 employee-year observations over the years 2006 to 2008.

Table 1 Panel A we present the descriptive statistics of the demographic characteristics of the employees. During the sample period, about three quarters (73%) of employees are men, the average age of the employees is 32.34 years old, and on average they stay at FM for 2.82 years. There are in total five levels of employees, the tasks they do and skills needed are more or less comparable among job levels except that at higher levels (e.g., level 4 and 5) employees are also explicitly expected to coach lower level colleagues and may serve different kinds of clients. FM hires and assigns employees to different levels according to experience they have accumulated within or outside of FM, but in the same professional service industry. The average years of experience employees in FM have in the relevant industry is 9.28 years.

In Panel B we show the descriptive statistics of different compositions of employee's annual bonus. We see that on average, employees get an annual bonus of 4,208 euros, while based on the bonus formula, the average predicted bonus is 2,859 euros. We can see that positive adjustment of bonus is much more common than negative adjustment (p25=385>0), and the average ex post adjustment to bonus equals 1348 euros, which is roughly 17% of the actual bonus an employee receives.

To conduct the follow up analysis, we also provide information regarding employee's performance, measured by revenue an employee brings to FM each year, as well as the change in revenue in the next year compared to the current one. Descriptive statistics are shown in Table 1 Panel D.

(Insert Table 1 about here)

In Table 2, we present the correlation coefficient table and corresponding p-values for the variables above. We can see the correlation between ex post bonus adjustment and employee's tenure in the firm is 0.39, significant at 99% confidence level (p-value=0.00), which lends primary support to our prediction.

(Insert Table 2 about here)

Moreover, we provide a univariate analysis to offer primary insight in our main idea put forward in the study, that the ex post adjustment to bonus granted to senior agents is to account for their overall contribution to the company that spans beyond the specified performance measure. We categorize observations by five quantiles for the revenue each employee makes in years 2006, 2007, and 2008, respectively, and provided the mean tenure and ex post adjustment of bonus for observations that fall in each revenue quantiles, the result is shown in Table 3. We can see that in general, longer tenured agents do not show better performance reflected in revenue than their junior colleagues. Nonetheless, they receive relatively higher level of subjective bonus in general, which indicates that the firm adjusts their bonus to compensate for effort that has not been reflected in performance outcomes specified as revenue. The fact that certain non-documented performance outcomes are rewarded entails that the firm treasures and wants to continuously incentivize the senior agents' capabilities. Note also that in year 2008, the performance ranking as resembled by generating revenue is more in line with the ranking of employee's tenure in firm, compared with the evidence from previous years. This lends support to the idea we put forward that, since senior employees are better aware of the operational environment of the company, they are more likely to find ways to sustain their business given the 2008 financial crisis than their junior colleagues, and the ability to survive the crisis is among those that are treasured and continuously rewarded by the firm. In the next session, we further substantiate our findings with evidence from regression analysis.

(Insert Table 3 about here)

4.4 Empirical Models and Results

In this section, we discuss our empirical model and the results related to the theory and hypothesis put forward in the previous sections.

4.4.1 Regression Methods

Given that our dataset is an unbalanced panel dataset, one applicable empirical method is to use pooled OLS regression. When using this method, we assume that each person-year observation has i.i.d distribution. This assumption might be violated, as there can be individual effect associated with each particular person. Potential ways to overcome this problem usually fall into two methods: fixed effect regression and random effect regression. Although both two approaches are useful panel data methods, fixed effect regression assumes that the individual effect is fixed over time for each individual, and is usually applied when the time span is relatively long compared with the number of individuals. The random effect regression assumes the individual effect is randomly distributed, and usually gives more accurate estimation when the time span is rather short relative to the number of individuals (Gelman and Hill 2006). As in our dataset an individual has a maximum of three years' observations, fixed effect model is likely to give rather inaccurate estimations.

Therefore, for all the following empirical models, we applied both pooled OLS regression as well as the random effect method. Note that the magnitude of the coefficients and corresponding significance levels are similar applying both regressions, as a result, we present the result of pooled OLS regression in the following subsection.

4.4.2 Analysis and Result

We use the following empirical model to test our hypothesis:

$$Subjectivity_{i,t} = \beta * tenure_{i,t} + \gamma * control_{i,t} + \varepsilon_{i,t}$$

In this model, the variable *Subjectivity* accounts for the ex post adjustment of bonus when employee's annual performance is realized, which equals the difference between actual bonus and the predicted bonus based on the specified pre-set formula. The major variable of interest *Tenure* represents the years since an employee's inception at FM till when he leaves or till present, and is the proxy for employee's seniority in the firm. In addition, we also include demographic information of the employees as control variables. *Age* is measured by the years between an employee's year of born and the current year. *Experience* is the total number of years an employee has been stayed in the financial service industry. *Gender* is a binary variable that takes a value of 1 for a male employee and 0 for a female employee. *Education* is a binary variable that takes the value of 1 if an employee has a research university degree or beyond, and 0 if he has an applied college

degree. We also include fixed effects where necessary. With respect to *H1*, we expect that *Subjectivity* is positively associated with *Tenure*.

The empirical results of the pooled OLS regression is shown in Table 4.The main effect of tenure on subjectivity is presented in column (1), where control variables and fixed effects are not included in the regression. We can observe that for the subjective adjustment in the final bonus, the tenure of the employee is significantly and positively related to subjective bonus, i.e., one year increase in tenure is associated with 276 euros in ex post adjustment of the bonus an employee gets. The positive effect of tenure on ex post adjustment of bonus holds true when control variables and fixed effects are included, as shown in column (2). Since *Age* and *Experience* are highly correlated (corr=0.91, p-value=0.00), including both variables may raise concern for multicollinearity. In all the main tests, we exclude the *Experience* variable, and instead include the quadratic term of *Age* (*Age2*) to account for the potential non-linear effect of *Age* on the subjective part of bonus an employee gets. Replacing *Age* and *Age2* with *Experience* and quadratic term of *Experience* does not qualitatively change the coefficients and significance for all analysis. In addition, since ex post adjustment to the predicted bonus can also arise when there is the need to filter out the uncontrollable events occurred for employees of the same operational environment (Bol 2008, Hoppe and Moers 2011), we also include the Location and Year fixed effect to hold constant in case of such situations. From column (2) we can see that the inclusion of control variables and fixed effects explains a larger portion of the variation in ex post bonus adjustment (R^2 increase from 0.157 when only *Tenure* is included, to 0.193), while the positive effect of tenure on ex post adjustment of bonus remains significant (coef=243.3, p-value=0.00). This result of the main test offers a primary support to our hypothesis.

We also use alternative form of bonus adjustment to proxy for subjectivity in compensation, namely, we use the percentage of ex post bonus adjustment scaled by the actual bonus, *(B_adj)%* as the dependent variable and conduct the same analysis, the result is presented in columns (3) and (4). From column (3) we can see that one year increase in tenure will result in a roughly 7.6% increase in the adjustment of bonus an employee gets as the portion of his actual bonus (coef=0.076, p-value=0.00). The positive relationship holds when control variables and fixed effects are included (coef=0.084, p-value=0.01 in column 4). The result of this analysis is again consistent with the our prediction that the level of subjectivity in the employee's bonus decision is positively associated with the employee's tenure in the firm.

Overall, this section provides empirical evidence to support our hypothesis. We may come to the conclusion that in general, the subjectivity in an employee's compensation is positively associated with the employee's tenure, the extent to which the principal uses subjectivity in compensating the employee's overall contribution to the firm increases with the employee's seniority in the firm.

(Insert Table 4 about here)

4.5 Additional Analysis and Discussion

In the previous section, we provide empirical evidence that the firm uses subjectivity in compensating employee's overall contribution to a larger extent as employee's seniority in the firm progresses. In this section, we provide further insight in establishing the underlying rationale of our finding.

4.5.1 Subjective compensation in predicting future performance

We find primary evidence that the company makes ex post adjustment to the pre-set bonus formula, and the extent to which the company does so differs between junior employee and senior employee. Our first follow up analysis is to examine whether the ex post adjustment to the employee's bonus is predictive of their future performance, and whether the employee's tenure moderates such relationship.

We follow the theory and analysis in Hayes and Schaefer (2000) and Ederhof (2010), they put forward the idea that if subjectivity in employee's compensation is effectively granted, then the unexplained part of employee's current period bonus will translate into his future performance. Conducting an analysis on whether current period subjective bonus is predictive of employee's future performance offers insight in firm's effective use of decision on subjectivity. Moreover, the potential dispersion of such effect between seasoned employees versus their junior peers provides further insight in understanding the rationale underlying our main empirical finding. This analysis is conducted using the following empirical model:

$$Performance_{i,t+1} = \beta_1 * subjectivity_{i,t} + \beta_2 * tenure_{i,t} + \beta_3 * subjectivity * tenure_{i,t} + \gamma * control_{i,t} + \varepsilon_{i,t}$$

Subjectivity and *Tenure* are defined as before. Since we are interested in the interaction effect of *Subjectivity*Tenure*, in this analysis we standardized *Tenure* by subtracting mean and scaling by standard deviation. Definitions of all control variables correspond with the ones used in the previous tests. Location, and year fixed effects are included where needed. Note that both the subjectivity measure and the controls

pertain to the current period (t). The dependent variable *Performance* is the individual employee's performance outcome in the next period (t+1). In our empirical test, as the bonus formula and the explicit performance measure remains the same over the employee's career, the future performance measure is specified as the increase of revenue (in euro) generated by each employee, calculated as the difference between next year's revenue and revenue in the current year.

The empirical results of the above regression analysis are presented in Table 5. We can see from column (1), current period subjectivity in bonus *B_adj* is positively associated with future performance (coef=2.06, p-value=0.07), which indicates that subjectivity in bonus is predictive of future performance, but only for junior employees. Longer tenure however, negatively affects the relation between ex-post adjustment of bonus and the employee's future performance (coef=-1.46, p-value=0.10). The dispersion of the effect that employee's current subjective bonus has on their future performance between junior versus senior employees are consistent when location and year fixed effects are included (results presented in column 2), and when both control variables and fixed effects are included (results presented in column 3). Including control variables and fixed effects explains a larger portion of variations in the dependent variable (R^2 increase from 0.029 to 0.115). In column (3), we can see that for junior employees, the ex post adjustment on their bonus in the current period leads to increase in their revenue made in the next year (coef=2.26, p-value=0.05), while for seasoned employees, their subjective portion of bonus does not translate into more revenue they make (coef=-1.53, p-value=0.08).

The collective evidence suggests that senior employees are on average granted a greater amount of subjective compensation, while the portion of subjectively granted compensation is less predictive to their future performance than that of their junior peers. It is reasonable for the company to make ex post adjustment to its junior employees, as the same rationale (incompleteness of performance measure in capturing effort) can apply to senior employees as well as less to senior ones. Meanwhile, when faced with a junior agent, the principal is in a arguably better position to assess his ability to perform the respective job, and the subjective adjustment to what the explicit performance formula entails represents the principal's signal to the agent about his potential to continuously deliver on the job (Benabou and Tirole 2003). In other words, the junior employee's effort or potential observed by the principal that is not yet reflected in his own current

performance is likely to translate into his future performance, and this justifies our finding that for junior agents, the ex post adjustment to his formula based bonus is predictive to his own future performance.

For a senior agent however, as he has spent more time in the firm, the principal gradually gets to know about the agent's personality and ability, and the principal can therefore rely more on her own judgment other than the ex-ante contracted performance measures. Moreover, the knowledge the agent gained over his accumulated years in the firm allows him to contribute to the overall development of the firm apart from the revenue generated from serving each client, which is likely to have positive externalities to other (especially junior) agent's observed performance. In this case, the senior agent's effort exerted to boost the overall development of the firm might even be detrimental to his own client-based performance outcome, and the expected bonus accordingly. The subjectively adjusted portion of bonus by the principal is therefore to compensate the senior agent's effort input that is not fully captured by the explicit performance measures.

Overall, the collective evidence provided thus far offers further insight that the firm's ex post adjustment to employees' bonus are granted to account for the imperfection in the ex-ante specified performance measures.

(Insert Table 5 about here)

4.5.2 Explicit performance measure to in picking up agent's effort

One might argue that the seemingly worse performance of higher tenured employees compared with their less tenured colleagues may indicate that the measurement of future performance in our study, i.e. revenue, is incomplete even for the explicitly recorded and ex-ante contracted performance measures. This can make sense to some extent, however, the other (non-financial) performance measure dimension, i.e., a survey score of customer satisfaction, generally has a rather high score for almost all employees (mostly ranged from 4 to 5) in the firm, with relatively little variation. Moreover, non-financial performance measures such as customer satisfaction can be more prone to employee's manipulation. The firm in our study also records a set of performance appraisal on employee's multiple dimensions of competence given by their local manager, these appraisals however, captures employee's underlying characteristics rather than their effort exerted in each period. We find the firm's decision of ex post adjustment on bonus is not sensitive to employee's appraisal outcome (result untabulated). In other words, to the extent that bonus rewards employee's periodic

effort rather than underlying *ability*, we believe that revenue is arguably the most direct and indicative among all available performance measures.

4.5.3 Nonlinear relationship between subjectivity and tenure

In addition, we also provided an empirical test for the potential non-linearity in the relation between tenure and subjectivity. The empirical model used to test for this non-linearity is as follows:

$$Subjectivity_{i,t} = \beta_1 * tenure_{i,t} + \beta_2 * tenure^2_{i,t} + \gamma * control_{i,t} + \varepsilon_{i,t}$$

If the relationship between tenure and subjectivity is represented with a U-shaped curve, we expect the coefficient on the first-order term to be positive, and the one on the second-order term to be negative. The empirical result of the above regression with control variables and fixed effects is shown in Table 6. Column (1) presents the main result of non-linearity when control variables and fixed effects are not included. We can clearly observe the estimated coefficients correspond to our expectation for subjectivity in ex post adjusted bonus, that is, the coefficient on *Tenure* equals 669.6 and significant at 99% confidence level (p-value=0.00) and the coefficient on *Tenure^2* equals -45.6 and also significant at 99% confidence level (p-value=0.00). The same relationship remains when we include control variables and fixed effects, as shown in column (2). Applying the estimated coefficients from column (2), we can calculate at which level of tenure the model reaches the expected point of inflection (626.3 + 2*(-43.6)**tenure*=0), we have an estimated tenure of around 7.2 years. Compared with the descriptive statistics in Table 1, we know the estimated tenure level at the inflection point is among the top percentiles of tenure distribution (p75=4.1<7.2). More specifically, about 10% of observations in our data has a tenure longer than 7.2 years. This result implies that as an agent's tenure reaches a relatively high level, tenure does not speak much to the subjective compensation adjustment. It could be the case that the most senior employees are those who are not the best "management material" but are experts in their own field so the firm keeps them. For these employees we may reasonably expect that their effort can be better captured by the ex-ante specified performance measures and thus are less spilled over to the performance of others. The result of this analysis lends support that the relation between tenure and subjectivity in bonus is a U-shaped form, while for the majority of employees in the company, the extent to which their annual bonus gets adjusted increases rather monotonically in their tenure in the company.

4.6 Conclusion and Extensions

In this paper, we employed a unique dataset to provide insights regarding ex post adjustment of incentive compensation. We study the situation when the firm applies a universal performance formula to all its employees, what alternative it resorts to in order to better recognize each employee's overall contribution. We put forward the idea that use of subjectivity in bonus decision is the instrument to serve the purpose. We find that subjectivity is more likely to be applied to an agent when he has longer tenure in the firm, and the extent that subjectivity in bonus is predictive of employee's future performance is more prominent for junior agents relative to senior ones. The main reason for the above collective results is that the performance measures set in an explicit compensation contract are often far from optimal, and the likelihood of flaw is especially high for senior agents rather than junior ones.

As with all studies ours is subject to limitations. As all our data and incentive system information are obtained from one firm in a specific industry (a financial consulting firm), we have to be cautions in terms of generalizability. Moreover, with only one firm in the dataset we are unable to control for the general macroeconomic environment or exogenous shock that prevails to the whole industry. Absence of the control for the overall economic environment might hinder the credibility of the explanatory power of subjective incentive compensation over future performance. In addition, limited time span of sample period – mostly fell into the time of financial crisis, thus the performance measures can be quite noisy and unrepresentative – has also restricted us to provide further supporting evidence derived from the main arguments of our study.

There are a number of further issues that may be of interest to look at in the following studies. In the current study we examined subjectivity in bonus, the short-term incentive. For further extensions, it can be insightful to incorporate the subjectivity in long-term incentive (salary) as well as the interplay between the two types of incentive. It is reasonable to believe that subjectivity might be applied to different extent in long-term (salary increase) and short-term (annual bonus) incentives, as some studies has identified that subjectivity as an incentive for employee's development is mainly executed through bonus pool allocation (Bailey et al. 2011) other than through salary.

Moreover, the mechanism we propose that subjectivity is used in gauging future performance requires an accurate and sufficient monitoring system. Studying the interplay between subjective compensation and the firm's monitoring system can provide further and deeper insights into firm operation and organizational design. We may try to obtain further information from the firm about how their employee performance is monitored, and investigate such as the relation between level of subjectivity and monitoring intensity.

References

Bailey, W. J., Hecht, G., & Towry, K. L. (2011). Dividing the pie: The influence of managerial discretion extent on bonus pool allocation. Contemporary Accounting Research, 28(5), 1562-1584.

Baiman, S., & Rajan, M. V. (1995). The informational advantages of discretionary bonus schemes. Accounting Review, 557-579.

Baker, G. (2002). Distortion and risk in optimal incentive contracts. Journal of human resources, 728-751.

Baker, G., Gibbons, R., & Murphy, K. J. (1994). Subjective performance measures in optimal incentive contracts. The quarterly journal of economics, 109(4), 1125-1156.

Baker, G., Gibbons, R., & Murphy, K. J. (2002). Relational Contracts and the Theory of the Firm. The Quarterly Journal of Economics, 117(1), 39-84.

Baker, G. P., Jensen, M. C., & Murphy, K. J. (1988). Compensation and incentives: Practice vs. theory. The journal of Finance, 43(3), 593-616.

Banker, R. D., & Datar, S. M. (1989). Sensitivity, precision, and linear aggregation of signals for performance evaluation. Journal of Accounting Research, 27(1), 21-39.

Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. The review of economic studies, 70(3), 489-520.

Bol, J. C. (2008, December). Subjectivity in compensation contracting. In AAA Management Accounting Section (MAS) 2006 Meeting Paper.

Bouwens, J., Hofmann, C., & Van Lent, L. (2018). Performance measures and intra-firm spillovers: Theory and evidence. Journal of Management Accounting Research, 30(3), 117-144.

Bouwens, J., & Van Lent, L. (2007). Assessing the performance of business unit managers. Journal of Accounting research, 45(4), 667-697.

Bushman, R. M., Indjejikian, R. J., & Smith, A. (1995). Aggregate performance measures in business unit manager compensation: The role of intrafirm interdependencies. Journal of Accounting research, 33, 101-128.

Datar, S., Kulp, S. C., & Lambert, R. A. (2001). Balancing performance measures. Journal of Accounting Research, 39(1), 75-92.

Ederhof, M. (2010). Discretion in bonus plans. The Accounting Review, 85(6), 1921-1949.

Feltham, G. A., & Xie, J. (1994). Performance measure congruity and diversity in multi-task principal/agent relations. Accounting review, 429-453.

Gelman, A., & Hill, J. (2006). Data analysis using regression and multilevel/hierarchical models. Cambridge university press.

Gibbons, R., & Henderson, R. (2012). Relational contracts and organizational capabilities. Organization science, 23(5), 1350-1364.

Gibbs, M., Merchant, K. A., Stede, W. A. V. D., & Vargus, M. E. (2004). Determinants and effects of subjectivity in incentives. The Accounting Review, 79(2), 409-436.

Hayes, R. M., & Schaefer, S. (2000). Implicit contracts and the explanatory power of top executive compensation for future performance. The RAND Journal of Economics, 273-293.

Hölmstrom, B. (1979). Moral hazard and observability. The Bell journal of economics, 74-91.

Höppe, F., & Moers, F. (2011). The choice of different types of subjectivity in CEO annual bonus contracts. The Accounting Review, 86(6), 2023-2046.

Keating, A. S. (1997). Determinants of divisional performance evaluation practices. Journal of Accounting and Economics, 24(3), 243-273.

Paul, M., & Roberts, J. (1992). Economics, organization and management.

Appendix A Variable Definition

Level	Level represents different job levels in FM, with 1 being the lowest and 5 the highest
Tenure	Tenure measures the years since an employee's inception at FM till when he leaves or till present.
Age	Age is measured by the years between an employee's year of born and the current year.
Experience	Experience is the total number of years an employee has been stayed in the financial service industry.
Gender	Gender is a binary variable that takes a value of 1 for a male employee and 0 for a female employee.
Education	Education is a binary variable that takes the value of 1 if an employee has a research university degree or beyond, and 0 if he has an applied college degree.
Real_bonus	The actual bonus an employee receives at the end of each year
Pred_bonus	The expected bonus calculated with the specified bonus formula for each year
B_adj	Expost bonus adjustment, the difference between an employee's actual bonus and predicted bonus, =(Real_bonus-Pred_bonus)
(B_adj)%	Percentage of ex post bonus adjustment scaled by the employee's actual bonus, =(B_adj/Real_bonus)
Rev	Annual revenue the employ makes of each year
Rev_inc	Increase of annual revenue in the following year compared with the current year, =(f.Rev-Rev)

Appendix B Uniform Formula in Determining Annual Bonus

Since 2006, FM introduced an annual bonus for all employees, determined by a uniform, pre-set formula. The bonus is determined with a score denoted as **Bscore**, times the bonus base which varies among different levels of employees.

Bscore = 0.7 * **NCscore** + 0.3 * *CS*

Where NCscore is the score determined with net contribution (*NetContri*), with the following correspondence:

NetContri	**NCscore**
NC>40%	6
30%<NC<=40%	5
20%<NC<=30%	4
10%<NC<=20%	3
0<NC<=10%	2
NC<=0	1

NetContri is the revenue the employee makes at the end of the year minus all labor cost.

CS is the customer satisfaction score collected from client's opinion the employee serves.

Once the Bscore is calculated with the above formula, it is timed by employee's bonus base as follows:

Level	Bonus base
1	0.5 monthly salary
2	0.5 monthly salary
3	1 monthly salary
4	1.5 monthly salary
5	2 monthly salary

In addition, employees get additional 500 euros for successfully referring one person to join FM, and 1000 euros for successfully introducing one new client to the company.

Table 1. Descriptive Statistics

This table presents the descriptive statistics of variables used in the main analysis. The data is tracked annually in FM's personnel system, ranges from year 2006 to 2008. Panel A shows summary statistics of employee demographic characteristics. Panel B shows summary statistics of employee's composition of bonus. Panel C shows summary statistics of performance indicators of employees. Variable definitions can be found in Appendix A.

	N	mean	min	p25	p50	p75	max	std.dev
Panel A: Demographic characteristic								
Level	369	2.70	1	2	3	4	5	1.15
Tenure	369	2.82	0.16	0.90	1.91	4.10	10.67	2.55
Age	369	32.34	21	26	31	37	54	7.27
Experience	366	9.28	1	4	8	12	34	7.10
Gender	369	0.73	0	0	1	1	1	0.45
Education	369	0.26	0	0	0	1	1	0.44
Panel B: Compensation								
Real_bonus	369	4208	0	2264	3443	5456	16185	2944
Pred_bonus	369	2859	290	947	2408	4167	12097	2234
B_adj	369	1348	-6673	385	1526	2259	11010	1774
(B_adj)%	367	0.17	-17.6	0.13	0.39	0.57	0.85	1.29
Panel C: Performance								
Rev	369	115719	8791	80230	108215	142824	815805	56180
Rev_inc	181	10981	-92225	1096	10981	24150	88856	24991

Table 2. Correlation Statistics of Main Variables

	Level	***Tenure***	***Age***	***Experience***	***Gender***	***Education***	***B_adj***	***Rev_inc***
Level	1							
Tenure	0.34***	1						
	(0.00)							
Age	0.81***	0.37***	1					
	(0.00)	(0.00)						
Experience	0.76***	0.35***	0.91***	1				
	(0.00)	(0.00)	(0.00)					
Gender	0.23***	0.11**	0.21***	0.19***	1			
	(0.00)	(0.04)	(0.00)	(0.00)				
Education	0.09*	-0.05	-0.01	-0.03	0.00	1		
	(0.06)	(0.27)	(0.78)	(0.58)	(0.99)			
B_adj	0.27***	0.39***	0.26***	0.26***	0.11**	0.07	1	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)	(0.17)		
Rev_inc	-0.005	0.01	-0.03	-0.03	-0.02	0.14*	0.11	1
	(0.94)	(0.87)	(0.69)	(0.69)	(0.76)	(0.06)	(0.14)	

* represents level of significance. *** - significant at 99% confidence level. ** - significant at 95% confidence level. * - significant at 90 confidence level. P-values are presented in the parenthesis. Variable definitions can be found in Appendix A.

Table 3 Mean Tenure and Subjective Bonus of Each Revenue Quantiles

This table presents mean value of tenure and ex post adjustment on bonus for different percentiles of performance in each year, respectively. Sample is categorized by revenue quantile each employee fall in each year. Percentile Revenue 5 represents observations whose revenue is above p80, percentile revenue 4 represents observations whose revenue is between p60 and p80, percentile revenue 3 represents observations with revenue between p40 and p60, percentile revenue 2 represents observations with revenue between p20 and p40, and lastly, percentile revenue 1 represents the rest of observations.

	2006		*2007*		*2008*	
Percentile Revenue	***Tenure***	***B_adj***	***Tenure***	***B_adj***	***Tenure***	***B_adj***
5	2.84	2120	2.83	930	3.11	1347
4	3.15	1851	4.12	2007	3.56	1901
3	3.08	1783	2.74	1365	3.07	1120
2	2.85	1096	2.31	915	1.98	1285
1	2.86	1083	2.16	748	1.77	907

Table 4 Relationship Between Tenure and Subjectivity in Bonus

This table presents the empirical result of the following regression:

$$Subjectivity_{i,t}=\alpha + \beta 1 * Tenure_{i,t} + \gamma * Controls + \epsilon_{i,t} \quad (1)$$

Where *Subjectivity* is represented by ex post adjustment on bonus *B_adj*, variables are defined in Appendix A. *Location* and *Year* fixed effects are included in columns (2) and (4). P-value is shown in parenthesis, *, **, *** represents result significant at 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	**B_adj**	**B_adj**	**(B_adj)%**	**(B_adj)%**
tenure	276.0***	243.3***	0.076***	0.084***
	(0.00)	(0.00)	(0.00)	(0.01)
age		40.4		-0.123
		(0.69)		(0.13)
age2		-0.1		0.002
		(0.92)		(0.19)
gender		191.9		-0.074
		(0.33)		(0.63)
education		411.5**		0.002
		(0.04)		(0.99)
Constant	570.2***	-623.8	-0.046	2.269*
	(0.00)	(0.71)	(0.65)	(0.10)
Location FE	No	Yes	No	Yes
Year FE	No	Yes	No	Yes
Observations	369	369	367	367
R-squared	0.157	0.193	0.022	0.050

Table 5 Predictability of Bonus Subjectivity in Future Performance

This table presents the empirical result of the following regression:

$$Performance_{i,t+1} = \alpha + \beta 1 * Subjectivity_{i,t} + \beta 2 * Tenure_{i,t} + \beta 3 * Subjectivity_{i,t} * Tenure_{i,t} + \gamma * Controls + \epsilon_{i,t} \quad (2)$$

Where *Subjectivity* is represented by ex post adjustment on bonus *B_adj*, variables are defined in Appendix A. *Tenure* is standardized by subtracting mean and scaled by std.dev. *Performance* is represented by the increase in employee's annual revenue *Rev_inc*. *Location* and *Year* fixed effects are included in columns (2) and (3). F-stat of coefficients b1+b3 is shown in each column respectively. P-value is shown in parenthesis, *, **, *** represents result significant at 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)
VARIABLES	**Rev_inc**	**Rev_inc**	**Rev_inc**
B_adj	2.06*	2.47**	2.26**
	(0.07)	(0.03)	(0.05)
Tenure	1,286.52	1,237.64	1,728.19
	(0.63)	(0.64)	(0.53)
*B_adj*Tenure*	-1.46*	-1.61*	-1.53*
	(0.10)	(0.06)	(0.08)
age			-300.54
			(0.89)
age2			3.03
			(0.92)
gender			-1,105.35
			(0.80)
education			6,427.10
			(0.14)
Constant	9,458.53***	4,137.72	9,773.71
	(0.00)	(0.33)	(0.79)
Location FE	No	Yes	Yes
Year FE	No	Yes	Yes
F(b1+b3)	0.20	0.45	0.31
Observations	181	181	181
R-squared	0.029	0.102	0.115

Table 6 Non-linear Relationship between Tenure and Subjectivity

This table presents the empirical result of the following regression:

$$Subjectivity_{i,t} = \beta_1 * tenure_{i,t} + \beta_2 * tenure^2_{i,t} + \gamma * control_{i,t} + \varepsilon_{i,t}$$

Where *Subjectivity* is represented by ex post adjustment on bonus *B_adj*, variables are defined in Appendix A. *tenure^2* is the quadratic form of *tenure*. *Location* and *Year* fixed effects are included in column (2). P-value is shown in parenthesis, *, **, *** represents result significant at 10%, 5%, and 1% level, respectively.

	(1)	(2)
VARIABLES	**B_adj**	**B_adj**
tenure	669.6***	626.3***
	(0.00)	(0.00)
tenure2	-45.6***	-43.6***
	(0.00)	(0.00)
age		18.0
		(0.86)
age2		0.1
		(0.93)
gender		200.1
		(0.30)
education		394.3**
		(0.04)
Constant	118.3	-707.8
	(0.50)	(0.67)
Location FE	No	Yes
Year FE	No	Yes
Observations	369	369
R-squared	0.186	0.218

Summary in English

This thesis contains three essays, where I conduct my research using data and managerial practice from the field, and through the lens of economic theory to shed light on a broader understanding of the information contained in performance measures.

Chapter 2 puts forward the idea that company's use of performance measures can contain information about the desired direction of effort that the principal expects the agent to exert. In this chapter we provide theoretical and empirical evidence that the introduction of an additional performance measure works effectively as communicating the principal's desired direction of effort to the agents, while the effect is mainly driven by the less experienced agents. For more experienced agents however, the introduction of this performance measure infringes their authority of doing things "in their own way". Consistent with this message, we are aware that conditional on the agent's seniority in the firm, the universal performance measure can be imperfect in picking up the agent's overall contribution to the firm. We provide evidence in Chapter 4 that the principal resorts to making ex-post adjustment in (senior) agent's bonus to account for the incompleteness of the pre-set performance measure. Apart from directing and recognizing the agent's effort, performance measures also serve the role of identifying and gauging agent's overall value of human capital. In Chapter 3 I communicate the idea that performance measures, if used properly, can contain information about the extent to which principal distinguishes her assessment of the agents' current performance from their future potential.

Summary in Dutch

Dit proefschrift bevat drie essays, waarin ik mijn onderzoek doe aan de hand van gegevens en de managementpraktijk uit het veld, en door de lens van de economische theorie licht werpt op een breder begrip van de informatie die in prestatiemetingen is vervat.

In hoofdstuk 2 wordt het idee geopperd dat het gebruik van prestatiemetingen door het bedrijf informatie kan bevatten over de gewenste richting van de inspanning die het bedrijf van de agent verwacht. In dit hoofdstuk geven we theoretisch en empirisch bewijs dat de invoering van een extra prestatiemaatstaf effectief werkt als het communiceren van de door het bedrijf gewenste inspanningsrichting naar de agenten, terwijl het effect vooral wordt gedreven door de minder ervaren agenten. Voor de meer ervaren agenten is de introductie van deze prestatiemaatstaf echter in strijd met hun bevoegdheid om dingen "op hun eigen manier" te doen. In overeenstemming met deze boodschap zijn we ons ervan bewust dat, afhankelijk van de anciënniteit van de agent in het bedrijf, de universele prestatiemaatstaf onvolmaakt kan zijn in het oppakken van de totale bijdrage van de agent aan het bedrijf. We leveren bewijs in hoofdstuk 4 dat het bedrijf zijn toevlucht neemt tot het maken van ex-post aanpassing in de (senior) agent bonus om rekening te houden met de onvolledigheid van de vooraf ingestelde prestatie-indicator. Naast het sturen en herkennen van de inspanning van de agent, dienen de prestatiemetingen ook de rol van het identificeren en meten van de totale waarde van het menselijk kapitaal van de agent. In hoofdstuk 3 communiceer ik het idee dat prestatiemetingen, indien goed gebruikt, informatie kunnen bevatten over de mate waarin de opdrachtgever haar beoordeling van de huidige prestaties van de agenten onderscheidt van hun toekomstige potentieel.