HARVARD | BUSINESS | SCHOOL

DOCTORAL PROGRAMS |

DISSERTATION ACCEPTANCE CERTIFICATE

(To be placed in First Library Copy)

The undersigned, approved by the Chair of the Doctoral Programs, have examined a dissertation entitled

Overcoming Limits in the Design of Formal Management Control Systems: Evidence from the Field

presented by **Jee-Eun Shin**

candidate for the degree of Doctor of Business Administration and hereby certify that it is worthy of acceptance.

Signature ....

Dennis Campbell, Chair

Signature Srikant Datar

Srikant M. Datar

Signature Brian J. Hall

Brian J. Hall

Signature Paul Healy

Paul M. Healy

Signature ....

Charles C. Y. Wang

Date August 20, 2018

WYSS HOUSE | BOSTON, MA 02163 | Ph 617.495.6101 | Fx 617.496.2203 | doctoralprograms@hbs.edu | GEORGE F. BAKER FOUNDATION

Overcoming Limits in the Design of Formal Management Control Systems:

Evidence from the Field

A dissertation presented

by

Jee-Eun Shin

in partial fulfillment of the requirements

for the degree of

Doctor of Business Administration

Harvard University

Graduate School of Business Administration

Cambridge, Massachusetts

August 2018

Dennis Campbell **Jee-Eun Shin**

OVERCOMING LIMITS IN THE DESIGN OF FORMAL MANAGEMENT CONTROL SYSTEMS: EVIDENCE FROM THE FIELD

ABSTRACT

The management accounting literature emphasizes the importance of management control systems as means to monitor and influence activities in a firm. Whereas prior works in this area primarily focus on the design of explicit mechanisms such as the design of formal incentive contracts or performance evaluations systems, a growing stream of research suggests the importance of considering informal means as part of an effective management control system. However, the difficulty in observing such informal controls has given rise to abundant analytical research with only scant empirical evidence. In my dissertation, I intend to shed light on how organizations implement and maintain different forms of informal control mechanisms by providing empirical evidence. In my first two essays, I explore informal controls at two organizations, and demonstrate how such mechanisms can mitigate limitations in explicit contracting. My third essay explores the efficacy of relative performance benchmarks at public firms, and thereby sheds light on whether the use of relative performance benchmarks provides a successful means to follow the informative principle in standard incentive contracts.

ACKNOWLEDGEMENTS

This dissertation would not have been possible if it were not for my committee. I am deeply thankful to Dennis Campbell, my dissertation chair, who always had faith in my ideas, and provided me with immense support and encouragement in pursing them. I am sincerely grateful to Srikant Datar for his wisdom and guidance on my research and my career; Brian Hall for his time and support to pursue my research agenda; Paul Healy for his invaluable help and advice that always pushed myself and my research to a higher level; and Charles Wang for sharing his resources, and teaching and guiding me throughout our collaboration that allowed me to grow into a more thorough researcher.

I would like to thank all the faculty at the Accounting and Management unit, and would like to express my sincere gratitude to three faculty members in particular: Tatiana Sandino, Bob Simons, and Bob Kaplan. I would not have been able to take on this journey if it were not for Tatiana who believed in my research ideas before even joining the program. Her passion for research and her warmth with which she sympathizes with me make her the greatest mentor throughout my doctoral studies. I am very grateful to Bob Simons for having faith in me, and giving me access to all his resources to complete this dissertation. I would like to express my appreciation to Bob Kaplan for giving me the inspiration to pursue my research path – his true devotion to research allowed me to believe in myself throughout my doctoral studies.

I appreciate the opportunity for being able to collaborate on projects with Jan Bouwens, Paul Ma, Susanna Gallani, and Ohchan Kwon during the course of my studies. This dissertation has benefited from significant positive spillover effects by working with them. I thank Jan for not only hand-holding me through the process of my very first joint project, but also for having become a good friend with whom I can share personal opinions on my European-influenced taste for good bread; Paul for the incredible learning experience by working together on our joint project, and for the generosity and understanding to sympathize with the status as a job-searching doctoral candidate; Susanna for her endless research-related and mental support allowing me to re-gain my confidence at critical moments on the job market; and Ohchan for the birth of our interdisciplinary research project which is part of this dissertation, and allowing me to re-define “doing research” as a process that is more fun than work-related.

I would also like to thank all other accounting DBA students (or by now, faculty) for being a great support as colleagues and friends – especially, Shelley Li whose care and support made me feel like I have a bigger sister; Carolyn Deller who lit up my otherwise dull doctoral life by being the most supportive and warmest office mate; Akash Chattopadhyay and Rajesh Vijayaraghavan who have commiserated with me about doctoral life and encouraged me to keep moving on; and Wei Cai for sharing unforgettable memories with me that gave Boston a special place in my heart. I am thankful for the support from all my friends at HBS – especially, Yo-Jud Cheng who was always there for me to listen to all my worries, and celebrate good occasions with me by conquering the Boston food scene; and Cheng Gao who infused my doctoral life with fun and laughter. Also, I am grateful to the members of the Doctoral Programs office – especially, Jennifer Mucciarone, Marais Young, and LuAnn Langan who were always ready to help me with any matter since my very first day of arrival on the North American continent and see me graduate successfully. This dissertation would not exist if it were not for Jae Yong Shin, my Master’s thesis advisor at Seoul National University, whose research inspired me to pursue a doctorate degree in management accounting. I am indebted to the support from the Korea Foundation for Advanced Studies during my doctoral studies.

Last, but not least, I am most thankful to my parents – my father On-Myung Shin, and my mother Kyoung-Won Paik – for raising me as a person that I am today. Their care and love provide me with the strength to overcome hard times, and the energy to keep going on from cherishing the good times.

To all my family and friends, I dedicate this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Limitations in the design of formal systems have spawned an abundant stream of research exploring the determinants and sustainability of informal controls. Nevertheless, there is only scant research regarding the implementation of such informal controls in organizations. Thus, in my dissertation, I use archival data to examine the prevalence and effectiveness of informal control mechanisms. My first two essays explore specific forms of informal control mechanisms that have been implemented at two different organizations. My third essay evaluates the use of relative performance benchmarks at public firms as a means to mitigate limitations in the design of optimal incentive contracts.

My first essay examines the role of relationships as informal control mechanisms. I use data from a company that outsources its core business operations to a nationwide vendor network, and that encourages its network managers to cultivate relationships with high-performing vendors as a means to induce better vendor performance instead of relying on explicit performance-based contracts. In particular, I examine the effect of disruptions in management control relationships on performance by examining how network manager changes affect subsequent vendor performance. I find that network manager changes are used as a control mechanism to sustain vendor performance such that changes initiated by the organization reverse vendor under-performance. My findings also document the costs to relational incentives when manager changes happen due to reasons unintended by the organization.

My second essay, co-authored with Ohchan Kwon, examines how, under periods of organizational change, alignment between management and employees on organizational objectives can act as informal control mechanisms to induce more desirable employee behaviors.

We use data from a company that made a company-wide merger announcement to its employees, and examine the performance effects of store locations that exhibit high versus low management-employee alignment. We show that following a merger, employees at store locations with low management-employee alignment focus effort levels more on relatively short term-oriented performance measures (e.g. sales) and less on more goal-congruent performance measures (e.g. customer satisfaction). Collectively, our findings suggest that management-employee alignment can mitigate career concern-related myopic employee behaviors under periods of organizational change.

My third essay, co-authored with Paul Ma and Charles Wang, examines whether and to what extent managers at public firms are evaluated, in their relative performance incentive contracts, on the bases of systematic performance. Theoretically, the use of relative performance benchmarks intends to filter out common noise underlying the chosen performance measure in the incentive contract that is beyond the manager's control – i.e. evaluate the manager based on systematic performance. Focusing on relative total shareholder returns (rTSR), the predominant metric specified in these contracts and used by market participants to evaluate managers, we document that 60% of firms that choose specific peers do a remarkable job of capturing the systematic component of returns. However, 40% of firms that choose index-based benchmarks retain substantial systematic noise in their rTSR metrics, which they could have substantially corrected by using their self-chosen compensation benchmarking peers. We show that selection of noisy benchmarks is economically important, is driven by compensation consultants' tendencies and by firms' governance-related frictions, and is associated with lower ROA.

CHAPTER 2

RELATIONAL DISRUPRIONS AS A CONTROL MECHANISM

2.1. Introduction

This study examines the effect of changes in relationships (i.e. relational disruptions) on performance. The literature provides ambiguous predictions regarding performance consequences subsequent to relational disruptions. On the one hand, the relational contracts literature[1] stipulates the advantages of relationship-building such that development of trust over an on-going, long-term relationship results in productivity improvements. If so, relational disruptions could hamper these benefits. On the other hand, literatures in accounting and management highlight potential drawbacks of relationships. The trust and familiarity from on-going, long-term relationships may trigger adverse behaviors such as collusion or negligence in monitoring. These literatures propose that knowingly implementing changes in relationships (for example via auditor, employee, and job rotation) can have positive consequences, and suggest that relational disruptions may result in productivity improvements.[2]

I address this debate in the literature by examining the performance consequences of relational disruptions using field data from a roadside assistance company that outsources its towing services to a nationwide vendor network. The organization's vendor network is divided into multiple different geographic zones to which network managers are assigned. Within their zone, network managers have complete authority over selection of vendors, assignment of services to its vendors, and monitoring of overall vendor performance. To study the effects of relational

[1] See, e.g., Bull (1987); MacLeod and Malcomson (1989); Gibbons (1997); Bernheim and Whinston (1998); Levin (2003); Board (2011); Halac (2012); and Malcomson (2013)
[2] For literatures on auditor rotation see e.g. Dopuch et al. 2011; Myers et al. 2003; Gipper et al. 2017; etc. For literatures on employee and job rotation see e.g. Ortega 2001; Arya and Mittendorf 2004; Eriksson and Ortega 2006; etc.

disruptions on performance, I examine how network manager changes affect subsequent vendor performance. Vendor performance is measured using two performance measures in this setting: (1) Call Acceptance – the extent by which vendors commit to the services that have been assigned to them, and (2) ETA Accuracy – representing service quality measured by the fraction of services that have been delivered on time.

This research site provides an attractive setting to examine the performance consequences of relational disruptions. First, the institutional details of the setting allow me to empirically test for the performance consequences of relational disruptions. Network manager changes in each zone happen at different points in time which enables me to employ a difference-in-differences research design to compare vendor performance in zones with and without a network manager change. Second, the organization primarily relies on relational incentives to motivate better vendor performance which allows me to isolate the performance consequences of changes in manager-vendor relationships from changes in explicit contracts. Vendor relations with the organization are formalized via explicit cost contracts that follow industry-wide standards. Instead of relying on explicit performance-based contracts to motivate better vendor performance, network managers are encouraged to maintain relationships with high-performing vendors. I provide empirical evidence to validate the prevalence of relational incentives in this setting.

My findings show that performance subsequent to relational disruptions is affected differently depending on the nature of the change in the relationship. Specifically, I proxy for disruptions in manager-vendor relationships using vendor network manager changes, and examine vendor performance subsequent to such relational disruptions. In doing so, I distinguish between two types of relational disruptions. First, relational disruptions that are due to employee-initiated reasons such as self-development, new job opportunities, or family issues. In my setting, these relational disruptions constitute network manager changes not intended by the organization (hereafter, "unintended relational disruptions"). Second, relational disruptions that are company-initiated. In my setting, these relational disruptions constitute network manager changes that the organization knowingly implemented as part of its management control system to reverse declined vendor performance (hereafter, "intended relational disruptions").

My finding related to unintended relational disruptions documents the potential costs of using relational incentives in that employee-initiated network manager changes exhibit performance declines relative to vendors that did not experience a network manager change. Consistent with the "unintended" nature of such employee-initiated relational disruptions, there are no significant performance differences prior to the unintended network manager change between vendors in zones subject to such relational disruptions and vendors in zones that did not experience any relational disruption. The empirical results show that unintended relational disruptions are associated with subsequent vendor performance deteriorations immediately after the employee-initiated network manager change which gradually starts to improve starting the fourth month subsequent to the change. Collectively, these findings suggest that the reliance on relational incentives to motivate performance is vulnerable to adjustment costs from unintended relational disruptions.

My finding related to intended relational disruptions shows that network manager changes are used as a control mechanism by the organization as a response to vendor performance declines. Specifically, a one percentage decrease of zone Call Acceptance (ETA Accuracy) in the prior month is associated with a 40% (57%) increase in the probability that the organization will implement a network manager change. The results examining the performance consequences of such intended relational disruptions suggest that vendors exhibit subsequent performance improvements relative to vendors that did not experience a network manager change. In particular, vendors exhibit underperformance prior to the intended relational disruption which reverses subsequent to the *announcement* of the company-initiated network manager change. Collectively, these findings suggest that the organization purposely uses network manager changes in response to deteriorations in vendor performance, and that the vendors respond to such intended disruptions with efforts to maintain their relationship with the organization.

I conduct additional cross-sectional tests to examine by what kinds of relationships the performance consequences from relational disruptions are driven. The results from additional cross-sectional tests provide evidence that relative bargaining power of vendors in the relationship with network managers is a primary determining factor. In particular, my findings show that vendors not facing any competition from other within-network vendors for service assignments in their corresponding market area exhibit ETA Accuracy declines subsequent to intended *and* unintended relational disruptions. Such vendors face relatively less pressure in having to gain trust from network managers, and are subject to less performance incentives and presumably have less incentives to adjust to changes in network managers.

The study contributes to the broader literature on management control and incentive systems in several ways. First, much of the extant literature has focused on explicit contract choices that resolve incentive problems – for example, explicit incentive compensation choices, or the choice of performance measures. It does not address the challenges of implementing optimal explicit contracts in the presence of high negotiation and adaptation costs. The findings in this paper demonstrate that relational incentive systems serve as an alternative means to mitigate such contracting difficulties. Using field data from an organization that encourages relationship-building to motivate performance, this study demonstrates that renewal of relationships can act as a viable management control system to sustain performance in relational incentive systems.

Second, the study contributes to the relational contract literature by investigating potential costs of relational incentives. Prior theoretical and empirical studies document the benefits of sustainable, long-term relationships, but pay limited attention to their potential costs when disruptions occur that require new relationships to be formed. This study documents the costs of unexpected relationship renewals and the types of relationships that are more/less susceptible to such costs. These findings are also likely to be of interest to practitioners for mitigating adverse performance consequences from unexpected employee turnover within their organization.

Finally, this study sheds light on what kinds of relationships matter in sustaining relational incentives. The relational contracting literature has emphasized the role of on-going, long-term relationships without considering heterogeneity in the decision-making authority of the involved contractual parties. However, Aghion and Tirole (1997) suggest that organizations exhibit varying degrees of allocation between formal (the right to decide) and real authority (the effective control over decisions), and it remains an open question as to whether the distinction between these two types of authorities matters in the formation of relational contracts. This study is based on a single field site where formal authority is held constant, and relationship changes are solely driven by changes in real authority (i.e. network manager changes). The findings that, both, intended and unintended, manager changes are associated with significant performance disruptions demonstrate that relational contracts are influenced by relationships among contractual parties with real authority.

The remainder of the paper is organized as follows. In Section 2.2, I review the prior literature. Section 2.3 describes my research setting. In Section 2.4, I develop my hypotheses.

Section 2.5 describes my data and provides summary statistics. I explain my empirical research design and discuss the corresponding results in Section 2.6, and conduct additional tests in Section 2.7. Finally, I conclude with Section 2.8.

2.2. Prior Literature

The subject of considerable research in economics revolves around how the optimal contract should link the agent's pay to information relevant to his or her performance, carefully weighing risks against incentives. The informativeness principle (Ross 1973; Holmstrom 1979) argues that agent performance should be evaluated using performance metrics that are informative about managerial effort or talent, and that any additional information about the agent's action, however imperfect, can be used to improve the welfare of both the principal and the agent. However, information asymmetry, incomplete performance measures, and environmental unpredictability constitute impediments to the design of optimal explicit contracts. For example, a classic problem articulated in the principal-agent models is that outcomes conditional on the agent's action are uncertain, and that the agent's behaviors are, therefore, unobservable. In fact, any observable performance measure included in the explicit contract is subject to varying levels of desirable performance measure attributes: sensitivity, precision, and goal-congruence (Banker and Datar 1989). Moreover, ongoing employment relationships often give rise to dynamic incentives that are difficult to predict ex-ante. For example, economic recessions, the opening of new markets, the arrival of new technologies, the introduction of new legislation, and industry consolidation may constitute shocks that were unforeseen at the time of contracting. Under such circumstances, the optimal contract that was designed to incentivize employee effort may become obsolete. The hardship lies in explaining the details of performance due to such environmental unpredictability.

Literatures in economics and accounting propose several alternative management control mechanisms to circumvent such drawbacks from explicit contracting. Within the design of explicit contracts, research shows that the use of subjectivity and relative performance evaluation can mitigate some incentive problems. For example, analytical models show that incentive distortions created by incomplete observable performance measures can be complemented by using subjective performance measures and predict that compensation contracts based on both type of measures can result in better outcomes (Baker et al. 1994; Baiman and Rajan 1995, Rajan and Reichelstein 2006, 2009). Another way to mitigate incentive problems in the design of optimal explicit contracts is the use of relative performance evaluation. Since performance measures are incomplete and contain common noise that is beyond the agent's control, it can be desirable to filter out such common noise by comparing performance relative to a benchmark. In theory, such practice is well known to elicit costly unobservable effort and maximize performance (Ross 1973; Holmstrom 1979, 1982).

Alternatively, a growing body of research suggests the use of organizational design choices as a viable means for mitigating limitations from explicit contracting. For example, information will be more efficiently used if more decision rights are delegated to managers (Baiman et al. 1995; Jensen and Meckling 1992; Nagar 2002), and complemented with incentives to ensure that they do not abuse authority (Nagar 2002; Prendergast 2002). The reason is that managers will have better job-specific or market-related knowledge, such that they are more able to set appropriate performance benchmarks to induce greater agent effort, and better monitor agent behaviors (Datar et al. 2009; Aghion et al. 2017). Moreover, Campbell (2012) proposes employee selection as an important element of organizational control systems to mitigate incentive problems due to incomplete explicit contracts. He documents that employees selected via channels that are likely to sort on the alignment of their preferences with organizational objectives are more likely to use decision-making authority in an overall desirable manner for the organization.

Other than explicit contracts or organizational design choices, research increasingly shows that relationship dynamics between economic agents are significant drivers of firms' competitive advantage. For example, Henderson and Gibbons (2012) attribute productivity differences in firms with similar observable characteristics to relational contracts. A relational contract is defined as "informal agreements and unwritten codes of conduct" (Gibbons 1997) between two or more parties that is based on subjective measures which are enforced by the shadow of the future rather than by the threat of legal action. The literature on relational contracts[3] shows that on-going relationships can motivate desirable behaviors among economic agents absent an explicit formal contract that specifies detailed performance-based conditions. Relational contracts eliminate the difficulties of having to specify detailed performance incentives or the need to adapt contractual terms as environmental conditions change. Even though non-binding, game theoretic models demonstrate that relational contracts are sustainable because they are self-enforcing – i.e. the value of the future relationship is sufficiently large that neither party wishes to renege. Research also shows that relational contracts serve as important complements to the explicit formal contract, especially in inter-firm or supply relationships (Poppo and Zenger 2002; Goo et al. 2009).

Whereas on-going, long-term relationships may substitute for the implementation of costly alternative incentive or control mechanisms as argued in the relational contracts literature, they also lead to the development of social relationships between economic agents. Due to repeated

[3] For analytical research on relational contracts, see Bull (1987); MacLeod and Malcomson (1989); Gibbons (1997); Bernheim and Whinston (1998); Levin (2003); Board (2011); Halac (2012); and Malcomson (2013) etc. For empirical research on relational contracts, especially in inter-firm relationships, see McMillan and Woodruff (1999); Banerjee and Duflo (2000); Johnson et al. (2002); Gil (2013); Gil and Marion (2013); Zanarone (2013); Barron et al. (2015); Calzolari et al. (2015); etc.

transacting experience, contractual parties may learn about each other's preferences, and acquire a deeper understanding and specific knowledge about each other. Accordingly, relationship-building may act as a double-edged sword. On the one hand, it can result in favorable performance consequences via synergies in productivity and cost management as mutual understanding between the contractual parties develops. On the other hand, the development of mutual trust may facilitate negligent monitoring behavior. Contractual parties may become numb to each other's performance such that performance deteriorations may be tolerated more easily. If the costs of mutual trust outweigh its benefits, the reliance on relationships can result in unfavorable performance consequences.

In fact, some literatures highlight the adverse consequences of on-going, long-term relationships, and propose relational disruptions as a means to enhance productivity. For example, relational disruptions via rotating employees in assigning jobs have been shown to be effective in sustaining performance within organizations. Eriksson and Ortega (2006) articulate three explanations for organizations engaging in job rotation. First, rotation makes employees more versatile and facilitates employee learning. Second, job rotation serves as a mechanism by which the organization can learn about the employees' productivities and the profitability of different jobs or activities (e.g. Ortega 2001, Arya and Mittendorf 2004). Third, as rotation mitigates boredom, rotating employees to different jobs may enhance employee motivation. Similarly, relational disruptions via rotating auditors have been proposed to improve auditor quality by preempting the possibility that management and auditors renege or collude, and thus may result in biased reports.[4] As described later, this study uses data from an organization that outsources its core business

[4] Yet, the empirical evidence as to whether mandatory auditor rotation and audit partner tenure affect audit quality remain mixed. (U.S. Senate 1977; AICPA 1978, 1992; SEC 1994; Gipper et al. 2017; Dopuch et al. 2011; Myers et al. 2003)

operations to a nationwide vendor network, and relies on changes of vendor network managers as a proxy for relational disruptions to examine the associated performance consequences.

2.3. Research Setting

The research site for this study is a US roadside assistance company (hereafter, ROAD). The company's primary business is to arrange roadside assistance for drivers covered by automobile original equipment manufacturers (OEM). New car sales of OEMs typically include three- to five-year warranties covering towing and repairs for vehicles disabled by mechanical problems, dead batteries, lockouts, etc. Its revenue stream consists primarily of the contracts with the OEMs. Motorists calling for roadside assistance assume that the OEM is in charge of the dispatch service, and not ROAD. Accordingly, the quality of roadside assistance services that ROAD provides to the motorists directly affects the OEM's brand reputation, and maintaining a sufficient quality for its roadside assistance services is crucial to secure future business with the OEMs.

For the execution of the requested roadside assistance services, ROAD completely outsources them to outside vendors (or service providers).[5] However, such service providers are free to accept or reject service requests from ROAD based on other available service requests from competitors, distance to be traveled, and availability of equipment. In fact, service providers obtain a large number of roadside assistance requests from other non-ROAD companies, the local police, and individuals. Accordingly, to ensure timely delivery of requested roadside assistance services to its clients, ROAD is reliant on strong relationships with its service provider network. ROAD's major objective constitutes of inducing service providers to prioritize their service requests at a

[5] I.e. in this research setting, towing businesses constitute the vendors.

service quality that can maximize motorist satisfaction. To do so, ROAD maintains a nation-wide service provider network which is divided into 14 different geographic zones (see Figure 2.1 for a mapping of the 14 geographic zones). A network manager is in charge of managing the service provider network in each zone. ROAD's unique service provider network management system provides an ideal research setting to examine the performance consequences of relational disruptions for two reasons. First, the institutional details of the research setting allow me to exploit network manager changes to empirically examine the performance consequences of a disruption to the relationship between network managers and their assigned service providers. Second, ROAD primarily relies on a relational incentives to motivate better service provider performance which allows me to isolate the effects of changes in relationships from changes in explicit contracts. I will elaborate on both reasons in the following subsections 2.3.1 and 2.3.2, respectively.

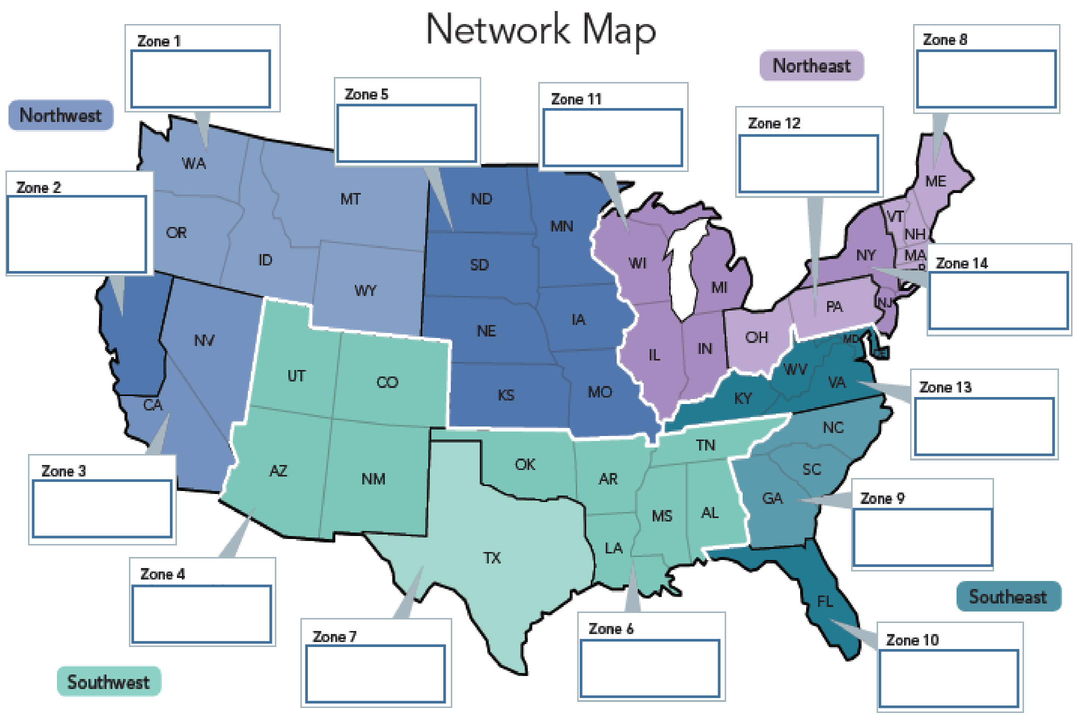
2.3.1. Service Provider Network and Network Managers

There are two crucial institutional elements that make this research setting suitable to empirically examine the performance consequences of relational disruptions. First, authority and responsibility regarding service provider performance management are largely delegated to the network managers. Network managers are responsible for delivering the most cost-efficient and high-quality roadside assistance services for their geographic zone. They engage in the following activities to fulfil their responsibility: selection of service providers, service assignment to service providers, and service provider performance monitoring. To facilitate the monitoring of service provider performance, ROAD heavily invests in data analytics to keep track of various performance measures that are used to evaluate service provider performance.[6] Service provider

[6] A detailed explanation of key service provider performance measures is provided in Section 2.5.1.1.

Figure 2.1. Service Provider Network

This figure provides a graphical illustration of the geographic partitions for the 14 different zones. Zone network managers are in charge of the service provider network within each of their geographical area



network performance on these measures constitutes a considerable element in the network managers' bonus and promotion determinations.

More importantly, network managers are encouraged to strengthen the service provider network by building interpersonal relationships. Relationship-building is particularly important as service providers can always choose to prioritize service requests from other non-ROAD sources. Specifically, service requests from the local police and individuals usually offer higher reimbursement rates than ROAD such that the service providers do not have incentives to accept service requests from ROAD if not for relational incentives. Details regarding the relational incentives embedded in the relationship between the network manager and service providers will be illustrated in more detail in Section 2.3.2.

Second, as network manager changes in each zone occur at different points in time, I am able to compare service provider performance of zones that were subject to network manager changes to zones that were not subject to network manager changes at all. By employing a difference-in-differences research design that includes service provider-fixed effects, network manager-fixed effects, and time-fixed effects, I am able to isolate changes in service provider performance due to the network manager change from time-invariant service provider characteristics (such as location, equipment endowment etc.), manager-specific characteristics (such as managerial ability etc.), and time effects (such as weather conditions, seasonal effects etc.).

2.3.2. Relational Incentives

ROAD considers relationship-building between the network manager and service providers as a crucial element for success. There are several reasons why relational incentives are more effective to motivate service providers than explicit incentive contracts. First, it is costly for ROAD to offer an attractive explicit contract that the service providers will accept. ROAD's ability to compete on price with the rates offered by the local police or individuals for their service requests is highly limited. Accordingly, network managers have to persuade service providers that prioritizing ROAD's service requests at lower rates will still be profitable for them. They do so by promising larger volume of services that the service providers can execute in a smaller radius such that they can reduce their travel time and boost their equipment utilization rate. Second, efforts to cater the explicit contracts for each service provider to induce optimal incentives are very costly. Network managers would have to manage explicit contracts for a large number of service providers – on average, with more than 220 different service providers that are dispersed over a wide geographic area. Third, motorist satisfaction is largely dependent on how service providers respond to unpredictable events which are inadequately captured by the performance measures that ROAD primarily relies on for monitoring purposes. For example, an initial service request may suddenly require a different set of equipment onsite. Another example is when service providers are willing to execute service requests in more remote areas without service providers within ROAD's network. Finally, even if network managers would be able to design explicit contracts that can specify price and volume terms together (thus, lowering the price per service with a guarantee of volume), and would also include performance-contingent terms (thus, with a guarantee of performance), maintaining relationships based on such explicit contracts would come at a cost of having to commit to a smaller number of service providers to whom they choose to make the contract offer. However, due to the unpredictability for the demand of roadside assistance services, network managers' utmost priority is to maintain and expand the service provider network to as many service providers as they can. Maintaining a simple cost contract that specifies price terms per service, and incenting better performance with the promise for high volume is the most cost-efficient and effective way to sustain their network.

The formal relationship between ROAD and its within-network-service providers starts by agreeing on a cost contract that specifies terms for each services rendered, and follows industry-wide formats.[7] These cost contracts do not embed any performance-based incentives or volume-related terms. To motivate service provider performance, network managers build relationships with high-performing service providers by promising them high volume of services in exchange of low price. Service providers that attain such status are referred to as "Preferred" service providers.[8] In addition to "Preferred" service provider status assignment, network managers engage in several important activities to strengthen the service provider network. First, network managers invite top service providers to three-hour meetings where they review the overall status of the market, preview new technology, and discuss with them how they can be better partners. Second, to gain their allegiance, ROAD also sponsors golf tournaments for top providers, invites them to share best practices at trade shows, and offer free, one-day training session to teach service providers how to manage their businesses more profitably. Third, ROAD also sponsors a blog where service providers share best practices, and monitor service provider sentiment through formal surveys twice per year. Lastly, to enhance the value offered to service providers, ROAD also negotiates with various vendors to offer promotions and discounts to its service providers on key equipment, such as vehicles, uniforms, mobile phones, and fuel.

---

[7] Details of the cost contracts are illustrated in Figure 2.2. Services rendered to clients comprise a pre-and a post-client-facing stage; costs associated with each stage are reimbursed in accordance with the terms of formal contracts. These typically specify a flat fee for each service rendered, and insures the vendor against potential service cancellations.

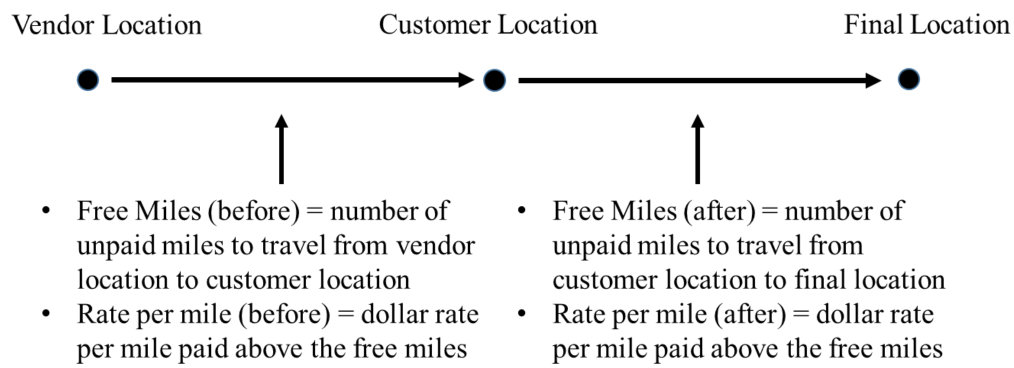
[8] The assignment of "Preferred" status occurs at the market area level which ROAD defines at the five-digit zip code level.

Figure 2.2. Cost Reimbursement Contract

This figure provides a graphical illustration of the detailed contractual terms of the explicit contract between the service provider and the network managers.

For any service:

- Flat fee = flat dollar fee paid to do service
- No show fee = payment issued to vendor if customer was not at available
- Rate schedule based on miles traveled:



2.4. Hypotheses Development

2.4.1. The Relational Contract between Service Providers and Network Managers

To illustrate the relational incentives between the service provider (SP) and network manager (NM), I use the theoretical model on relational contracts developed by Gibbons and Henderson (2012). The value of an on-going, long-term relationship between SP and NM can be illustrated in a repeated game that is equally likely to end after any period. In every period, both agents have the option to play either one of the following three strategies: "Cooperate", "Defect", or "Punish". The probability that the game will end influences the interest rate r per period that both contractual parties use in discounting their multi-period payoffs. Assuming that the payoffs per period are C from "Cooperate", D from "Defect", and P from "Punish", where D>C>P, the decision of whether to cooperate or defect then amounts to comparing two time paths of payoffs: (C, C, C, …) versus (D, P, P, P, …). The time path of cooperation yields a higher present value than the time path of defection if

$$(1+\frac{1}{r})C > D + \frac{1}{r}P\,, \tag{1}$$

where 1/r is the present value of a dollar to be received every period (until the game ends) starting next period. Rearranging (1) yields

$$r < \frac{C-P}{D-C} \tag{2}$$

Put differently, (2) indicates that if SP and NM are sufficiently patient (i.e., if r is sufficiently close to zero), then it is optimal to cooperate, forgoing the short-run temptation for the long-term gain.

Figure 2.3a illustrates a timeline of an on-going relationship between NM and SP when both play "Cooperate". In addition, the remaining strategies pertaining to "Defect", and "Punish" are illustrated in Figure 2.3b. The relational contract between NM and SP is sustainable at the cooperative outcome (highlighted in grey) because (1) SP defecting on service commitment and

Figure 2.3. Relational Contract between NM and SP

These figures illustrate the relational contract between the service provider (SP) and network manager (NM). Figure 2.3a illustrates the timeline of the relationship between a network manager (NM) and a service provider (SP) if they both cooperate. The relationship begins with the network manager offering a cost contract to the service provider and continues if SP agrees to the cost contract terms. The additional remaining strategies for "Defection", and "Punishment" for NM and SP in the multi-period game are illustrated in Figure 2.3b.

Figure 2.3a: Timeline of On-going Relationship between NM and SP

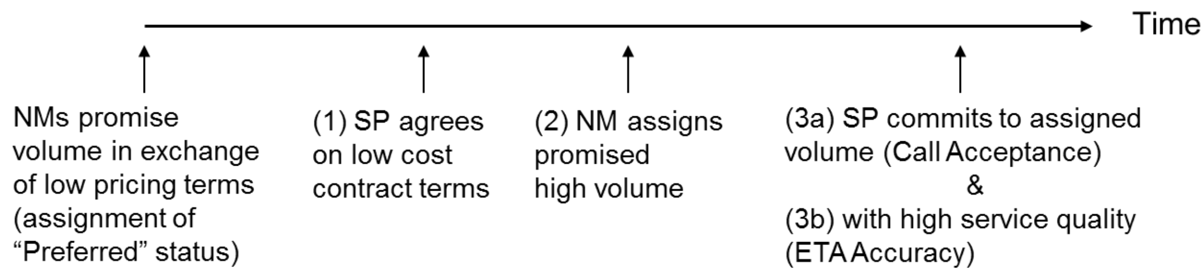


Figure 2.3b: NM and SP Strategies in Multi-period Game

	Action		
Actor	Cooperate (**C**)	Defect (**D**)	Punish (**P**)
Service Provider (**SP**)	**(1)** Agree on low cost contract terms **(3a)** Commit to assigned volume (Call Acceptance) **(3b)** High service quality (ETA Accuracy)	- Prioritize competitors' service requests - Shirk on service quality	- Demand higher price - Don't accept NM service requests
Network Manager (**NM**)	**(2)** Assign promised volume	- Prioritize service assignment to other SPs	- Don't assign services to SP - Seek out new SP as replacement

quality will trigger NM to stop assigning service volume in the following period, and (2) NM defecting on assigning volume to SP will likely trigger SP to prioritize alternative service requests. One might argue that NM's threat of defection is not credible in that the reason they keep on assigning service volume to SP is because they have agreed on low-cost, and thus they constitute the most cost-efficient option. However, expanding the SP network to fill coverage gaps constitutes another important objective of the NM in order to decrease the number of service assignments that go to costly out-of-network service providers without any pre-established cost rates.[9] To reach that objective ROAD maintains service provider recruitment specialists who are in charge of identifying coverage gaps in market areas, and cooperate with network managers in onboarding new service providers into the existing SP network base. However, NM's efforts to build relationships with as many SPs as possible to cover potential coverage gaps is associated with costs of having to decrease the allocated share of assignment to each individual SP, and thereby threatening the existing relational contract with incumbent service providers in ROAD's network.

My first set of hypotheses aims to test whether the relationship between network managers and service providers is at the cooperative outcome of the relational contract as illustrated in Figure 3b. In particular, the characterization of the relational contract yields two empirically testable predictions. First, if the negotiated price terms embed a trust element that lower price will be reciprocated by larger volume, the share of service assignments should be negatively associated with the negotiated cost terms. Second, better performance by the service providers will be reciprocated with a larger share of service assignments by the network managers. Accordingly, if ROAD relies on relational incentives to motivate its service providers, service provider

[9] The flat fee for services rendered to out-of-network service providers can increase up to $180. This is, on average, more than 5 times costly than the flat fee ROAD charges to its within-network service providers.

performance should be positively associated with the share of assigned services in the subsequent performance period, even in the absence of explicit performance-related explicit incentives. Accordingly, I test the following two hypotheses to test whether the relationship between NM and SP are at the cooperative outcome:

H1a: The share of service assignments is negatively associated with negotiated cost terms.
H1b: The share of service assignments is positively associated with SP performance.

2.4.2. Performance Consequences of Relational Disruptions

Reliance on relational incentives may make organizations more vulnerable to changes in relationships that were not intended by the organization (i.e. "unintended relational disruptions"). For example, Drexler and Schoar (2014) show that in the commercial banking industry which is heavily reliant on managers' decentralized knowledge, a disruption in the relationship between borrowers and loan officers (due to loan officer turnover) is associated with adverse consequences such as lower probability of borrowers to receive new loans from the bank, and a higher probability to miss payments or go into default. Similarly, ROAD finds itself reaping benefits from on-going, long-term supplier relationships that its manager has cultivated over time. Such built-up relational capital with the suppliers is individual-specific such that the manager becomes a valuable human asset to the organization. Even though the organization would like to retain such managers indefinitely, managers may decide to leave the organization for personal reasons such as for other job opportunities, self-development, or family-related reasons. Explicit contracts that specify detailed performance-based conditions may mitigate incentive problems arising from such unintended turnovers. However, in the absence of such explicit contracts, organizations that critically rely on the value of on-going, long-term relationships to sustain performance may be particularly vulnerable to the costs of unexpected changes in relationships. Accordingly, for organizations in which relationship-building is an integral part of driving agent performance, such unintended relational disruptions may be associated with significant adjustment costs resulting in temporary performance declines. Thus, I hypothesize:

H2a (Unintended Relational Disruptions): Employee-initiated network manager changes are associated with performance deteriorations after the change.

Despite the benefits of on-going relationships in sustaining performance in the absence of explicit performance-contingent contracts, increasing familiarity may compromise performance quality as both contractual parties are reluctant to update their monitoring and production processes with evolving business patterns. For example, in organizations such as ROAD that are heavily dependent on outside service providers, managers' may be prone to decreasing monitoring incentives of service providers with whom they have established relationships. Under such circumstances, knowingly changing incumbent relationships (i.e. "intended relational disruptions") by changing network managers may force the newly-assigned managers to take on a "fresh look" and reevaluate the performance of existing service provider relationships. Moreover, a change in the manager has an incentive effect for incumbent service providers to perform better. Incumbent suppliers expect that the newly-assigned manager will reevaluate supplier relationships, and therefore, will exert greater effort to maintain their position in the supply chain. Accordingly, the use of relational disruptions may constitute a viable control mechanism to reverse declining service provider performance. Thus, I hypothesize:

H2a (Intended Relational Disruptions): Company-initiated network manager changes are associated with performance deterioration before the change, and improvement after.

2.5. Data and Descriptive Statistics

2.5.1. Data

I obtain data on the number of service assignments to each service provider, service provider performance, and service provider cost reimbursement contracts (as illustrated in Figure 2.2) for the time period January 2011 to May 2016. In addition, I also obtain data on network manager-zone assignments, and the reasons for network manager changes for the corresponding sample period.

2.5.1.1. Service Provider Performance

Even though ROAD emphasizes different performance measures from time to time depending on their business strategy, the following two performance measures are most important, and are persistently used in evaluating service provider performance and determining network managers' variable bonus compensation: call acceptance, and estimated time of arrival (ETA) accuracy.

Call Acceptance: After receiving a service request from ROAD, service providers have the option to either accept or decline the service request. For timely service execution, the rate by which the service provider agrees to deliver the service when requested by ROAD is the most significant factor in determining overall network performance. The reason is that searching for other service providers once a service request has been declined consumes much time, resulting in delayed services. Such service delays are badly reflected in customer reviews and ultimately result in a lower likelihood of re-contracting with ROAD's OEM clients which constitutes its primary revenue stream. Call acceptance is measured as the fraction of service requests that were accepted by the service provider in a given month. This performance measure captures the commitment of service providers to execute a service request assigned by the network manager, and thus, does not reflect the ex-post quality of the service provided.

Estimated Time of Arrival (ETA) Accuracy: ETA accuracy is measured as the fraction of service requests that were rendered on time out of all services that the service provider accepted. When a service is requested, and the service provider agrees to execute the assigned service, the service provider is required to provide an estimate for the time frame by which he/she will be able to deliver the service. The accuracy of the estimate is the most crucial factor in determining the motorist's satisfaction with the service. Unlike Call Acceptance, this performance measure captures the ex-post quality of the service provided.

2.5.1.2. Relational Disruptions

To conduct my hypotheses tests related to the performance consequences of relational disruptions, I distinguish between two types of network manager changes. I proxy for intended relational disruptions (unintended relational disruptions) using network manager changes that are company-initiated (employee-initiated). [10] Data pertaining to the two reasons of manager changes are directly obtained from ROAD. As depicted in Table 2.1 Panel A, there are a total of 43 network manager changes in my sample, or approximately 3 manager changes per zone. Out of the 43 network manager changes, 15 are company-initiated and 28 are employee-initiated. [11]

---

[10] Similarly, the incumbent vendor-manager relationship may be affected by manager changes at the vendor firms. However, vendor firms in this research setting constitute small (local, family-run) towing businesses with relatively low levels of decentralization. Accordingly, firm manager changes at the vendor firm are unlikely to be predominant factors to influence the incumbent vendor-manager relationship.

[11] The reasons for intended relational disruptions are twofold: (1) network manager changes driven by "strategic initiatives" reasons (i.e. these changes are primarily driven by performance concerns); and (2) network manager changes driven by the incumbent network managers' promotion reasons (i.e. these changes are not driven by performance concerns). In additional robustness tests, I further distinguish between these two reasons for intended relational disruptions, and examine whether the hypothesized performance improvements subsequent to intended relational disruptions (as stated in H2b) are driven by network manager changes due to "strategic initiatives" and not due to promotion reasons.

Table 2.1. Descriptive Statistics

This table reports the descriptive statistics for the data used in this study. The data are monthly and range from January 2011 to May 2016. Panel A summarizes the total number of network manager changes in the sample period (column 1), and the mean number of changes in each zone (column 2). Panel B summarizes statistics at the zone-month level for all 14 zones. Panel C summarizes statistics for the 37 unique network managers in my sample period. Panel D summarizes statistics at the market areal-month level. Market areas are defined at the 5-digit zip code level. Panel E summarizes statistics at the service provider-month level for all service providers that were assigned services in my sample period. Panel F summarizes statistics at the service provider-month level only confined to service providers that serve in markets with more than one service provider. Rank denotes the rank of the service providers in terms of the number of total assignments. Accordingly, by definition, "Preferred" status service providers are denoted with rank equal to 1. Panel G reports the pairwise correlation matrix for the variables. * indicate significance at the 5% level. All variables are defined in Panel H.

Panel A: Network Manager Change

	(1) N	(2) mean	(3) min	(7) max
# NM Change	43	3.071	1	5
# of Employee-initiated NM Change	28	2	1	5
# of Company-initiated NM Change	15	1.071	1	3

Panel B: Zone

	(1) N	(2) mean	(3) min	(4) p25	(5) p50	(6) p75	(7) max
# of Zip5 areas	910	185.4	37	139	178	222	387
# of SPs	910	218.7	38	166	215	268	450
Unregistered %	910	0.01	0.00	0.01	0.01	0.02	0.06
# of Assignment to Unregistered SPs	910	501.0	75	262	415	667	2,534
Call Acceptance of Unregistered SPs	910	99.15	95.71	98.84	99.30	99.60	100
ETA Accuracy of Unregistered SPs	910	72.62	54.55	69.37	72.79	76.19	87.69
Call Acceptance	910	40.49	19.42	35.06	40.31	45.77	62.53
ETA Accuracy	910	78.12	66.22	75.53	78.08	80.56	88.46

Panel C: Network Manager

	(1) N	(2) mean	(3) min	(4) p25	(5) p50	(6) p75	(7) max
NM Tenure (Zone)	37	15.12	1	8	14	18.50	41
NM Tenure (Total)	37	24.51	1	8	19	33	65
# of Managed Zones	37	1.595	1	1	1	2	4

Panel D: Market Area (Zip5)

	(1) N	(2) mean	(3) min	(4) p25	(5) p50	(6) p75	(7) max
Preferred SP Tenure	168,511	17.83	1	5	13	27	65
Change of Preferred SP	168,511	0.030	0	0	0	0	1
# of Assignments in Zip5	168,511	153.6	1	28	66	157	6,578
# of SPs in Zip5	168,511	1.370	1	1	1	1	8
# of potential SPs in Zip5	168,511	2.087	0	1	2	3	15
# of New SPs in Zip5	168,511	0.040	0	0	0	0	2

Table 2.1. Descriptive Statistics (Continued)

Panel E: Service Provider (All)

	(1) N	(2) mean	(3) min	(4) p25	(5) p50	(6) p75	(7) max
% of Assignments	199,035	84.75	0.0391	82	100	100	100
Single SP	199,035	0.720	0	0	1	1	1
Preferred SP	199,035	0.848	0	1	1	1	1
# of Assignments	199,035	130.1	1	26	59	134	5,753
Free Miles (b)	199,035	7.089	0	5	7.750	10	30
Rate (b)	199,035	1.457	0	1.250	1.500	1.750	10
Flat Fee	199,035	37.14	17.50	32	35	40	150
Free Miles (a)	199,035	2.379	0	2.500	2.500	2.500	20
Rate (a)	199,035	1.209	0	1	1.125	1.250	5
No Show Fee	199,035	16.46	0	15	15	15	150
Call Acceptance	199,035	39.45	0.114	17.65	35.29	58.70	100
ETA Accuracy	190,127	78.38	0	66.67	83.33	100	100

Panel F: Service Providers (Confined to Market Areas with Multiple Service Providers)

	Rank=1 N=25,313	Rank=2 N=25,313	Rank=3 N=4,031	Rank=4 N=713	Rank=5 N=198	Rank=6 N=68	Rank=7 N=17	Rank=8 N=5
% of Assignments	69.63	28.05	12.86	7.862	5.172	2.500	1.523	0.912
Preferred SP	1	0	0	0	0	0	0	0
# of Assignments	234.1	79.16	68.74	67.02	70.35	55.96	38.94	20.60
Free Miles (b)	7.289	6.986	7.277	7.139	7.717	9.085	8.765	10
Rate (b)	1.361	1.502	1.433	1.441	1.506	1.434	1.478	1.300
Flat Fee	35.09	38.08	36.81	36.95	36.27	33.86	34	32.50
Free Miles (a)	2.479	2.343	2.333	2.464	2.538	2.500	2.500	2.500
Rate (a)	1.157	1.230	1.220	1.207	1.182	1.105	1.199	1.150
No Show Fee	15.90	16.66	16.06	16.26	16.20	15.15	15	15
Call Acceptance	46.42	32.60	30.23	24.26	23.93	30.46	31.43	23.27
ETA Accuracy	79.32	76.69	74.64	73.85	69.70	64.38	69.39	88.89

Table 2.1. Descriptive Statistics (Continued)

Panel G: Correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) # of Assignments	1													
(2) Free Miles (b)	-0.06*	1												
(3) Rate (b)	-0.25*	0.28*	1											
(4) Flat Fee	-0.18*	-0.17*	0.35*	1										
(5) Free Miles (a)	0.18*	0.25*	-0.2*	-0.24*	1									
(6) Rate (a)	-0.14*	-0.19*	0.46*	0.49*	-0.18*	1								
(7) No Show Fee	-0.1*	-0.13*	0.2*	0.47*	-0.21*	0.31*	1							
(8) Single SP	-0.05*	-0.01*	0.02*	0.04*	-0.02*	0.03*	0.03*	1						
(9) # of Assignments in Zip5	0.76*	-0.03*	-0.21*	-0.15*	0.15*	-0.12*	-0.09*	-0.31*	1					
(10) # of SPs in Zip5	0.07*	0.02*	-0.04*	-0.06*	0.03*	-0.05*	-0.04*	-0.85*	0.44*	1				
(11) # of New SPs in Zip5	-0.03*	0.03*	-0.01*	-0.01*	0.01*	-0.02*	-0.02*	-0.1*	0.01*	0.13*	1			
(12) # of potential SPs in Zip5	0.02*	-0.01*	-0.13*	-0.10*	0.04*	-0.11*	-0.05*	-0.17*	0.10*	0.25*	0.04*	1		
(13) Call Acceptance	0.18*	-0.05*	0.06*	0.02*	0	0.03*	0.03*	0.02*	0.08*	-0.03*	0.02*	0.04*	1	
(14) ETA Accuracy	-0.07*	-0.03*	0.07*	0.05*	-0.04*	0.04*	0.05*	0.02*	-0.08*	-0.03*	-0.05*	-0.03*	0.06*	1

Table 2.1. Descriptive Statistics (Continued)

Panel H: Variable Definitions

Zone-Level	
# of Zip5 areas	Total number of market areas in a zone
# of NM Change	Total number of network manager changes in a zone
# of SPs	Total number of service providers serving in a zone
Unregistered %	Percentage of services that were executed by out-of-network service providers
# of Assignments to Unregistered SPs	Total number of service assignments by out-of-network service providers
Call Acceptance of Unregistered SPs	Percentage of services that were accepted by out-of-network service providers
ETA Accuracy of Unregistered SPs	Percentage of services that were executed within the estimated time frame provided by out-of-network service providers
Network Manager-Level	
NM Tenure (Zone)	Total number of months that a network manager served in a particular zone
NM Tenure (Total)	Total number of months that a network manager served as network manager regardless of different zone assignments
# of Managed Zones	Total number of zones that a network manager has served as network manager
Market Areal-Level	
Advantage SP Tenure	Number of consecutive months a “Preferred” status service provider maintains its status in a market area
# of Assignments in Zip5	Total number of service assignments in a market area
# of SPs in Zip5	Total number of service providers serving in a market area
# of potential SPs in Zip5	Total number of establishments as reported in the County Business Patterns (CBP) data from the Census Bureau
# of New SPs in Zip5	Total number of service providers with newly-registered cost contracts in a market area
SP-Level	
% of Assignments	Share of services assigned of the total number of service assignments to all service providers within the service provider’s market area
Single SP	Dummy equal to 1 if the service provider is the only service provider in its market area
Preferred SP	Dummy equal to 1 if the service provider is assigned the largest share of services in its market area
# of Assignments	Total number of services assigned
Free Miles (b)	Cost contract-related variables as defined as in Figure 2.2
Rate (b)	Cost contract-related variables as defined as in Figure 2.2
Flat Fee	Cost contract-related variables as defined as in Figure 2.2
Free Miles (a)	Cost contract-related variables as defined as in Figure 2.2
Rate (a)	Cost contract-related variables as defined as in Figure 2.2
No Show Fee	Cost contract-related variables as defined as in Figure 2.2
Call Acceptance	Percentage of services that were accepted by the service provider
ETA Accuracy	Percentage of services that were executed within the estimated time frame provided by the service provider

To empirically validate the nature of the company-initiated network manager changes as a viable control mechanism used by ROAD to reverse declining service provider performance, I estimate logit regressions on the panel of zone-year-month observations as follows:

$$\textit{Network Manager Change}_{z,t} = \alpha + \beta_1 * \textit{Call Acceptance}_{z,t-1} + \beta_2 * \textit{ETA Accuracy}_{z,t-1} + u_t + v_z + \varepsilon_{z,t} \quad (3)$$

The dependent variable *Network Manager Change* is either one of the following three indicator variables: *NM Change, Intended NM Change*, and *Unintended NM Change*. *NM Change* is coded as one for periods in which a zone underwent a network manager change, and zero otherwise. *Intended NM Change* (*Unintended NM Change*) is coded as one for periods in which a zone experienced a company-initiated (employee-initiated) network manager changes as indicated in the data obtained directly from ROAD. The key variables of interest are the lagged performance-related variables at the zone-level, *Call Acceptance* and *ETA Accuracy*.[12] As illustrated in the descriptive statistics in Table 2.1, each zone covers on average about 185 different market areas. To account for heterogeneity in the importance of each market area in evaluating overall zone performance, *Call Acceptance* and *ETA Accuracy* are weighted by the share of total assignments in each market area within the zone. The year-month fixed effects, u_t, control for any unobservable time-aggregate effect or time trend in the dependent variables. The zone fixed effects, v_z, control for any omitted zone characteristics that are time invariant. Standard errors are clustered at the zone- and year-month level. Accordingly, β_1 and β_2 capture whether prior zone performance is a significant determinant in predicting network manager changes.

[12] The recording of zone performance takes place monthly. In evaluating zone performance, ROAD primarily benchmarks against prior month's performance. ROAD's electronic recording system of zone performance-related variables highlights the performance differential between the current and previous months. Accordingly, prior month's performance constitutes the primary variable of interest.

The results are tabulated in Table 2.2. Columns 1, 2, and 3 report the results when the corresponding dependent variables are *NM Change, Intended NM Change,* and *Unintended NM Change,* respectively. Prior zone performance is not a significant determinant for overall network manager changes (column 1), and unintended network manager changes (column 3) as evidenced by the insignificant coefficients. However, the coefficients on the zone performance variables are significantly negative when the dependent variable is *Intended NM Change* (column 2). In particular, this implies that an additional unit increase in the weighted Call Acceptance (ETA Accuracy) measure decreases the probability of an intended network manager change by about 40% (57%). This result provides corroborating evidence that intended network manager changes by ROAD are driven by zone performance concerns, and that changing network managers constitutes a control mechanism used by ROAD to induce zone performance improvements.

2.5.2. Descriptive Statistics

As illustrated in Panel B of Table 2.1, on average, each zone covers approximately 185 market areas, and services are assigned to about 219 different service providers. On average, only 1% of all service requests are executed by out-of-network service providers that do not have a pre-established cost contract with ROAD. Panel C of Table 2.1 summarizes network manager characteristics. Over the sample period, there are a total of 37 unique network managers in charge of the 14 different zones. Some network managers rotate to different zones such that one network manager serves, on average, more than a single zone. The average tenure for a network manager serving in a single zone (in all zones combined) is about 15 months (25 months).

Network managers primarily focus on service provider performance at the market area level. Market areas at ROAD are defined at the 5-digit zip code level. Panel D of Table 2.1 provides

Table 2.2. Determinants of Network Manager Changes

This table provides the results from tests that empirically validate the nature of intended and unintended relational disruptions. In particular, the tables tabulate the results from logit regressions of equation (3) in Section 2.5.1.2. The sample constitutes monthly panel data for all zones in the time period between January 2011 and May 2016. The dependent variable in column 1 is *NM Change*, and is coded as one for periods in which a zone underwent a network manager change, and zero otherwise. The dependent variable in column 2 is *Intended NM Change*, and is coded as one for periods in which a zone experienced a company-initiated network manager change. The dependent variable in column 3 is *Unintended NM Change*, and is coded as one for periods in which a zone experienced an employee-initiated network manager change. *Call Acceptance* is the percentage of services that were accepted by service providers within ROAD's network in each zone. ETA Accuracy is the percentage of services that were executed on time by service providers within ROAD's network in each zone. Zone- fixed and year-month fixed effects are included throughout. Robust standard errors are reported in brackets and are clustered at the zone- and year-month level. ***, **, and * indicate significance at the 1, 5, and 10% level, respectively.

	(1) NM Change	(2) Intended NM Change	(3) Unintended NM Change
Call Acceptance$_{t-1}$	0.05	-0.50*	-0.05
	(0.090)	(0.272)	(0.111)
ETA Accuracy$_{t-1}$	0.01	-0.85**	0.19
	(0.153)	(0.423)	(0.175)
Observations	896	512	832
Zone FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Pseudo R-squared	0.469	0.648	0.465

summary statistics on the service providers' market area characteristics. "Preferred" status is defined as an indicator variable equal to 1 if a service provider is assigned the largest share of services within its corresponding market area. The average tenure as consecutive "Preferred" service provider is about 18 months. On average, about 1.4 service providers execute services in a market area (*# of SPs in Zip5*) which ranges up to 8. In each month, about 0.04 service providers in a market area newly join the network (*# of New SPs in Zip5*). Moreover, in each market area, there are about 2 towing businesses that qualify as potential service providers (*# of Potential SPs in Zip5*), but are not registered as service providers with ROAD. The maximum number of potential service providers in a market area ranges up to 15. These data are obtained from the US Census Bureau Country Business Patterns (CBP) which is an annual series that provides subnational economic data by industry. In particular, I obtain the number of establishments registered for the NAICS industry code 48841 ("Motor vehicle towing") at the five-digit zip code level.[13]

There are a total of 199,035 unique service provider-month observations. Descriptive statistics on service assignments, cost reimbursement terms, and service provider performance are summarized in Panel E of Table 2.1. The average service provider gets assigned about 130 services in a month. The summary statistics also suggest that there is some variation in the determination of contract details. From interviews with network managers, I learned that the elements in the cost reimbursement contract that largely determine the variability of total cost for each service provided are flat fee and rate (b). Whereas the average flat fee (rate (b)) for a single service dispatch is about

---

[13] The anonymous nature of the CBP data do not allow me to match the establishments with service providers in ROAD's network. Moreover, the CBP data are not recommended for use as a time-series, but rather represent a snap shot in time. Therefore, I only use the CBP data as a control for the pool of potential service providers that are registered as "businesses" as defined by the Census Bureau. ROAD's majority of service providers constitute small owner/operator towing establishments which may not be adequately captured by the CBP data. All my empirical results remain unchanged when I exclude *# of Potential SPs in Zip5* in my regression models.

$37 ($1.5), the maximum flat fee that ROAD pays can be up to $150 ($10). Summary statistics related to service provider performance indicate that, on average, service providers accept about 39% of all service request assignments, and execute about 78% of their accepted service request assignments within their estimated time frame. The average service provider gets assigned about 85% of all services within a market area (*% of Assignments*). *Single SP* is an indicator variable that is defined as one if a service provider is responsible for all service assignments in the market area (i.e. if *% of Assignments* is equal to 100%). The summary statistics indicate that more than 72% of all service providers are solely in charge of a market area.

Taken together, the summary characteristics suggest that ROAD's service provider network is largely reliant on one service provider for each market area. This could arise if (a) there are no other service providers in the corresponding market area that could have been assigned a service or (b) ROAD is comfortable allocating all its service assignments to a single service provider. ROAD's main objective to minimize coverage gaps, and expand the service provider network suggests the former. Moreover, the positive correlation between # of SPs in Zip5 and # of Potential SPs in Zip5 in Panel G of Table 2.1 suggests that # of SPs in Zip5 provides a good reflection of the availability of potential alternative out-of-network service providers, and represents a proxy to measure the extent by which within-network service providers face competition for ROAD's service assignments within each market area. [14]

[14] I expect the strength of relational incentives to be most predominant in the subsample of market areas where service providers face competition for ROAD's service assignments – i.e. market areas with more than one within-network service provider (*Multiple SP* subsample). Panel F of Table 2.1 provides summary statistics on the *Multiple SP* subsample. More specifically, it summarizes the same characteristics as in Panel E based on their rank in terms of the number of services they got assigned – rank 1 corresponds to the "Preferred" status service provider, rank 2 corresponds to the service provider with the second highest number of service assignments in a market area, and so on. There are a total of 25,313 observations in the *Multiple SP* subsample.

2.6. Empirical Tests and Results

2.6.1. The Relational Contract between Service Providers and Network Managers

To test the relational contract between service providers and network managers in terms of the negotiated cost terms (H1a), and whether network managers adjust their service assignment patterns taking into consideration service provider performance (H1b), I estimate the following regression model on the panel of service provider-year-month observations:

$$\begin{aligned} \textit{\% of Assignments}_{i,t} = \alpha &+ \beta_1 * \textit{\# of SPs in Zip5}_{i,t} + \beta_2 * \textit{\# of Potential SPs in Zip5}_{i,t} \\ &+ \beta_3 * \textit{Cost Contract Terms}_{i,t} + \beta_4 * \textit{SP Performance}_{i,t} \\ &+ u_t + v_i + w_m + \varepsilon_{i,t} \end{aligned} \tag{4}$$

The results are tabulated in Panel A of Table 2.3. The dependent variable is the share of services that were assigned to a service provider in its market area (*% of Assignments*). Determinants of the share of service assignments to each service provider include the availability of within-network service providers in the market area (*# of SPs in Zip5*), potential service providers outside of ROAD's network (*# of Potential SPs in Zip5*), the cost contract terms (as illustrated in Figure 2.2), lagged service provider performance (*Call Acceptance* and *ETA Accuracy*), service provider fixed effects (v_i), year-month fixed effects (u_t), and network manager fixed effects (w_m). Standard errors are clustered at the service provider- and year-month level. The predominant reliance on a single service provider for a market area may impede the network managers' ability to motivate better performance by allocating a larger share of service assignments. Accordingly, I will focus on the subsample of market areas with multiple service providers – i.e. the *Multiple SP* subsample in columns 3 and 4 to interpret my empirical findings.[15]

[15] The *Multiple SP* subsample is identified based on the indicator variable *Single SP* which is equal to 1 if the service provider is the only service provider in its market area, and zero otherwise. Market areas with only a single service provider in ROAD's network is referred to as the *Single SP* subsample.

Table 2.3. Service Assignment and SP Performance

This table examines the empirical results from testing H1a and H1b. Panel A reports the results from OLS regressions of equation (4) in Section 2.6.1. The sample constitutes monthly panel data for all service providers in the time period between January 2011 and May 2016. Columns 1 and 2 estimate equation (4) on the total sample, and columns 3 and 4 estimate equation (4) on the *Multiple SP* subsample only. The dependent variable is the percentage of services assigned to a service provider within the service provider's market area. All independent variables are defined in the Panel H of Table 2.1. Service provider-, network manager-, and year-month fixed effects are included throughout. Robust standard errors are reported in brackets and are clustered at the service provider- and year-month level. Panel B reports the results from logit regression of equation (5) in Section 2.6.1. The sample constitutes monthly panel data at the five-digit zip code (market) level in the same time period. The dependent variable is *Change of "Preferred" SP* which is an indicator that is defined as one if the service provider attaining the largest share of services in the market area has changed from the previous month, and zero otherwise. *Call Acceptance* is the percentage of services that were accepted by service providers within ROAD's network in each market area. ETA Accuracy is the percentage of services that were executed on time by service providers within ROAD's network in each market area. Market area-, network manager-, and year-month fixed effects are included throughout. Robust standard errors are reported in brackets and are clustered at the market area- and year-month level. In both panels, ***, **, and * indicate significance at the 1, 5, and 10% level, respectively.

Panel A: Service Provider Level

	Total Sample		Multiple SP	
	(1)	(2)	(3)	(4)
VARIABLES	% of Assignments in Zip5		% of Assignments in Zip5	
# of SPs in Zip5	-27.87***	-27.30***	-8.92***	-8.88***
	(0.738)	(0.762)	(0.561)	(0.587)
# of potential SPs in Zip5	-0.07	-0.05	-0.20	-0.22
	(0.123)	(0.126)	(0.166)	(0.171)
Free Miles (b)	0.15	0.13	0.17	0.11
	(0.120)	(0.120)	(0.253)	(0.237)
Rate (b)	-0.91	-0.95	-6.48***	-6.61***
	(0.701)	(0.729)	(1.909)	(1.967)
Flat Fee	-0.12***	-0.12***	-0.14*	-0.12
	(0.044)	(0.045)	(0.086)	(0.083)
Free Miles (a)	0.53	0.47	1.54*	1.49*
	(0.430)	(0.442)	(0.920)	(0.906)
Rate (a)	-1.42	-1.50	-2.86	-2.96
	(1.663)	(1.699)	(2.963)	(2.944)
No Show Fee	-0.19	-0.19	-0.11	-0.19
	(0.161)	(0.168)	(0.453)	(0.460)
Call Acceptance$_{t-1}$	0.04***	0.04***	0.14***	0.12***
	(0.004)	(0.004)	(0.012)	(0.010)
ETA Accuracy$_{t-1}$	0.01***	0.01***	0.02***	0.02***
	(0.002)	(0.002)	(0.004)	(0.004)
Call Acceptance$_{t-2}$		0.02***		0.07***
		(0.003)		(0.008)
ETA Accuracy$_{t-2}$		0.01***		0.02***
		(0.002)		(0.004)
Constant	129.11***	128.03***	75.64***	74.57***
	(3.401)	(3.444)	(7.404)	(7.333)

Table 2.3. Service Assignment and SP Performance (Continued)

Observations	176,245	160,782	49,480	45,376
NM FE	YES	YES	YES	YES
SP FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Adjusted R-squared	0.880	0.881	0.799	0.806

Panel B: Market Level

	Total Sample		Multiple SP	
	(1)	(2)	(3)	(4)
VARIABLES	Change of "Preferred" SP		Change of "Preferred" SP	
# of SPs in Zip5	1.259***	1.239***	0.385***	0.406***
	(0.034)	(0.035)	(0.058)	(0.058)
# of Potential SPs in Zip5	0.009	0.009	-0.022	-0.023
	(0.019)	(0.019)	(0.022)	(0.023)
Call Acceptance$_{t-1}$	-0.027***	-0.023***	-0.022***	-0.016***
	(0.001)	(0.002)	(0.002)	(0.002)
ETA Accuracy$_{t-1}$	-0.005***	-0.005***	-0.005***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Call Acceptance$_{t-2}$		-0.008***		-0.011***
		(0.002)		(0.002)
ETA Accuracy$_{t-2}$		-0.003***		-0.004***
		(0.001)		(0.001)
Observations	50,227	47,085	19,652	19,159
NM FE	YES	YES	YES	YES
Market FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Pseudo R-squared	0.0965	0.0967	0.0278	0.0316

Columns 1 and 3 only include lagged service provider performance variables by one month. Columns 2 and 4 additionally include lagged service provider performance variable by two months.

The results provide supporting evidence for H1a. For example, the coefficients on *Rate (b)* suggests that a one dollar decrease in rate (b) is associated with 6 % higher market share, and corroborates the trust element for service providers of entering into a low price contractual relationship in exchange for larger volume from the network managers. The results also provide supporting evidence for H1b. We observe positive coefficients on the lagged performance-related variables (*Call Acceptance* and *ETA Accuracy*) in column 3. In particular, a 1% increase in accepted calls (services rendered within the estimated time frame) is associated with 0.14% (0.02%) higher share of assigned services within the service provider's market area in the subsequent months. Even though the economic magnitudes are relatively small, this result suggests that network managers take service provider performance into consideration to adjust for service assignment patterns.

Other than the share of service assignments, the attainment or loss of "Preferred" status comprises a more powerful incentive mechanism even in market areas only with a single service provider. Losing "Preferred" status in market areas served by a single service provider implies losing market share by 100%. Even in market areas with multiple providers, losing "Preferred" status implies a loss in market share of about on average 40% as illustrated in Panel F of Table 2.1. Therefore, the threat of losing "Preferred" status is likely associated with more performance incentives for service providers than simply measuring the change in the share of service assignments. To corroborate my empirical findings for H1b that network managers take into consideration service provider performance in allocating service assignments, I estimate the following logit regressions on the panel of market area-year-month observations as follows:

$$\textit{Change of Preferred SP}_{a,t} = \alpha + \beta_1 * \textit{\# of SPs in Zip5}_{a,t} + \beta_2 * \textit{\# of Potential SPs in Zip5}_{a,t} + \beta_3 * \textit{SP Performance}_{a,t} + u_t + v_a + w_m + \varepsilon_{a,t} \quad (5)$$

The dependent variable *Change of Preferred SP* is an indicator variable that is defined as one if the service provider attaining the largest share of service assignments in a market area changed from the previous month, and zero otherwise. The results are tabulated in Table 2.3 Panel B. Similar as before, columns 1 and 2 (columns 3 and 4) estimate equation (5) on the total sample (only on the *Multiple SP* subsample). Columns 1 and 3 only include lagged service provider performance variables by one month. Columns 2 and 4 additionally include lagged service provider performance variable by two months. I control for the total number of potential service providers (*# of SPs in Zip5*), the total number of potential service providers outside of ROAD's network in the market area (*# of Potential SPs in Zip5*), and include market area fixed effects (v_a), year-month fixed effects (u_t), and network manager fixed effects (w_m). The key variables of interest are the lagged service provider performance-related variables (*Call Acceptance* and *ETA Accuracy*). Standard errors are clustered at the market area- and year-month level. The results corroborate the findings in Panel A, and suggest that a one percentage decrease in prior month's *Call Acceptance* (*ETA Accuracy*) increases the probability of losing "Preferred" status by about 3% (2%).

2.6.2. Performance Consequences of Relational Disruptions

To examine the performance consequences of relational disruptions, I exploit the institutional feature of this research setting that network manager changes in each zone happen at different points in time. This allows me to employ a difference-in-differences research design to compare the performance consequences of zones with and without a change in the network manager. As illustrated in Section 2.5.1.2, I distinguish between network manager changes that are company-initiated (i.e. intended relational disruptions) and those that are employee-initiated and are due to network managers' personal reasons (i.e. unintended relational disruptions). Specifically, for intended and unintended relational disruptions, I estimate the following model separately:

$$\begin{aligned} SP\ Performance_{i,t} = {} & \alpha + \beta_1 POST_{i,t} + \beta_2 TREAT_{i,t} * POST_{i,t} \\ & + Controls_{i,t} + u_t + v_i + w_m + \varepsilon_{i,t} \end{aligned} \tag{6}$$

To circumvent the lack of a control group without any network manager change, the sample for the difference-in-differences estimation is confined to the 12-month period surrounding each manager change event.[16] For example, if zone 1 has a new manager in July 2012, vendor performance data from January to December 2012 are included in the sample. For this unique network manage change event date, *POST* is defined as one for the months of July to December 2012 (with the new manager), and zero for the months of January to June 2012 (with the previous manager). Since zone 1 experienced a network manager change, *TREAT* is defined as one for all service providers in zone 1, and service providers in any other zones that experienced a network manager change in July 2012. The associated control group (i.e. defined as zero for *TREAT*) are the service providers in zones that did not experience any network manager change from January to December 2012. The sample on which I run the respective analyses to estimate the effects of intended (unintended) relational disruptions stacks all the service provider-months observations pertaining to each of the network manager change event dates. The dependent variable is service provider performance, and is measured as either *Call Acceptance* or *ETA Accuracy*. Control

[16] The choice of a 12-month period allows me to maximize the number of observations for the post-network manager change period as the fastest network manager turnover happens within 6 months. In untabulated robustness tests, I vary the length of the window ranging from 10 months to 14 months. The results remain unchanged.

variables include *# of SPs in Zip5*, *# of Potential SPs in Zip5,* all cost contract terms, service provider fixed effects (v_i), year-month fixed effects (u_t), and network manager fixed effects (w_m). Note that I do not include the variable *TREAT* by itself, because it drops out due to the inclusion of the service provider fixed effects. Standard errors are clustered at the service provider- and year-month- level.

The empirical model specified in equation (6) allows for a difference-in-differences test, where the coefficient β_2 captures the change in service provider performance between the pre- and post- periods for the service providers in zones that experienced a manager change, benchmarked against the change for service providers that did not experience a manager change during the same period. If unintended relational disruptions are associated with performance deteriorations as hypothesized in H2a, I expect $\beta_2 < 0$ in the sample of employee-initiated manager changes. Moreover, if intended relational disruptions are associated with performance improvements as hypothesized in H2b, I expect $\beta_2 > 0$ in the sample of company-initiated manager changes.

2.6.2.1. The Effect of Unintended Relational Disruptions

The results for the difference-in-differences analyses of the effect of unintended relational disruptions are tabulated in Table 2.4 and Figure 2.4. Table 2.4 examines the effect of unintended relational disruptions on service provider performance, and the relevant outcome variables are *Call Acceptance* in columns 1 and 2, and *ETA Accuracy* in columns 3 and 4. The coefficient of interest, β_2, in columns 2 and 4 suggests that unintended relational disruptions are associated with performance deteriorations in terms of lower *Call Acceptance* and lower *ETA accuracy*. More specifically, the reported coefficients on the interaction terms suggest that, on average, service providers in zones that are subject to an unintended network manager change exhibit a 0.52%

Table 2.4. Effects of Unintended Relational Disruptions

This table examines the empirical results from testing H2a, and reports the results from OLS regressions of equation (6) in Section 2.6.2 examining the effect of unintended relational disruptions. Unintended relational disruptions are defined as employee-initiated network manager changes. The sample is based on the 12-months surrounding an employee-initiated network manager change. For each unique employee-initiated network manager change event, *TREAT* is defined as 1 for the service providers in zones that experienced an employee-initiated network manager change, and 0 for the service providers in zones that did not experience any network manager change for the same corresponding 12-month time period. *POST* is defined as 1 for the 6 months subsequent to the employee-initiated network manager change, and 0 otherwise. The dependent variables are *Call Acceptance* (in columns 1 and 2) and *ETA Accuracy* (in columns 3 and 4). Control variables include *# of SPs in Zip5*, *# of Potential SPs in Zip5*, and all cost contract terms. Year-month- , network manager-, and service provider- fixed effects are included throughout. Robust standard errors are reported in brackets and are clustered at the service provider- and year-month level. ***, **, and * indicate significance at the 1, 5, and 10% level, respectively.

	Unintended NM Changes			
	(1)	(2)	(3)	(4)
VARIABLES	Call Acceptance		ETA Accuracy	
POST	0.17**	0.23***	0.02	0.07
	(0.083)	(0.086)	(0.129)	(0.131)
TREAT*POST		-0.52***		-0.35**
		(0.152)		(0.169)
Constant	49.24***	49.22***	76.80***	76.79***
	(2.385)	(2.383)	(2.297)	(2.296)
Observations	642,005	642,005	610,632	610,632
Controls	YES	YES	YES	YES
NM FE	YES	YES	YES	YES
NM Change-SP FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Adjusted R-squared	0.0413	0.0414	0.0215	0.0216

Figure 2.4. Effect of Unintended Relational Disruptions

Figure 2.4a graphically examines the empirical results from testing H2a, and reports the results from OLS regressions of equation (6) in Section 2.6.2 by plotting the coefficients with the corresponding 95% confidence intervals of the interactions between TREAT and each month in the 12-month period (instead of interacting *TREAT* only with *POST*). The sample is based on the 12-months surrounding an employee-initiated network manager change. For each unique employee-initiated network manager change event, *TREAT* is defined as 1 for the service providers in zones that experienced an employee-initiated network manager change, and 0 for the service providers in zones that did not experience any network manager change for the same corresponding 12-month time period. The dependent variables is *Call Acceptance*. Control variables are the same as in Table 5, and are defined in Panel H of Table 2.1. Figure 2.4b plots the mean Call Acceptance of service providers that underwent an employee-initiated network manager change (i.e. TREAT), and service providers that did not undergo any network manager change (i.e. Control). 0 on the x-axis represents the month immediately prior to the employee-initiated network manager change.

Figure 2.4a: Plot of coefficient for monthly interaction with TREAT

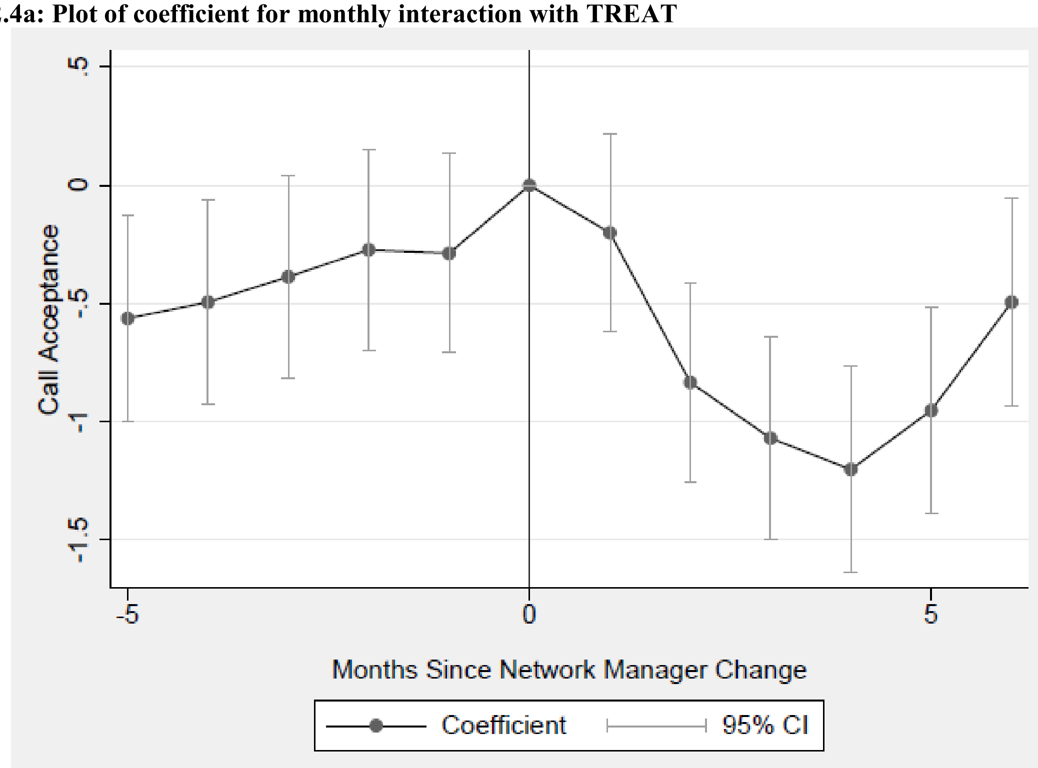
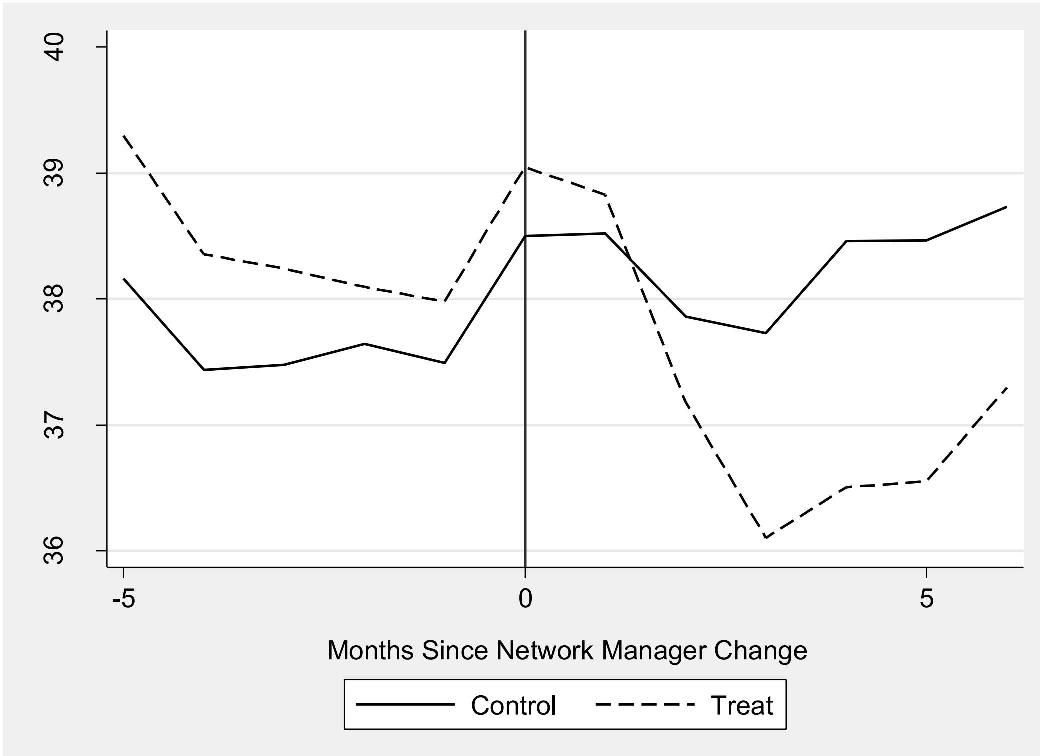


Figure 2.4. Effect of Unintended Relational Disruptions (Continued)

Figure 2.4b: Plot of Monthly Mean for Treatment and Control SPs



lower call acceptance rate, and 0.35% lower rate of services that were rendered within the estimated time frame relative to service providers for which the network manager did not change. Gauging from the descriptive statistics in Table 2.1, this translates into approximately 142 less overall services accepted for a zone in the months subsequent to a network manager change. To put the performance improvement of intended relational disruptions into perspective for ROAD, I highlight the fact that the number of service assignments to out-of-network service providers only constitutes of about 570 services per month for a zone (Panel B of Table 2.1). ROAD's utmost priority is to avoid out-of-network service assignments as it increases costs dramatically because there is no pre-established rate for the service. Put differently, this suggests that an unintended relational disruption may increase the number by about 25%, and the associated costs of having to resort to out-of-network service providers.

To examine whether the performance effects identified in Table 2.4 are indeed driven by performance deteriorations from unintended relational disruptions, Figure 2.4a graphically illustrates the results from the same regression as in equation (6) allowing for interactions with *TREAT* for all 12 pre- and post- manager change months. The omitted interaction (i.e. the base line) is the month immediately prior to the network manager change. Specifically, Figure 2.4a plots the coefficients for all interactions with the corresponding 95% confidence intervals. The solid line at 0 on the x-axis represents the month immediately prior to the network manager change. Note that the y-axis is zero when $t = 0$. This corresponds to setting the base-line at that month. Whereas there is no statistical difference in the pre-manager change period, zones that experienced network manager changes are associated with a lower call acceptance rate starting the first month of the network manager change. In addition, Figure 2.4b plots the mean *Call Acceptance* separately for service providers that underwent an unintended network manager change (i.e. Treat) and service providers that did not undergo any change (i.e. Control). It shows that the performance between the treated and control service providers largely follows parallel trends, whereas the performance of treated service providers starts to drop significantly subsequent to the unintended network manager change event date. Most importantly, the performance of treated service providers subsequent to the unintended network manager change declines below the performance level of control service providers. Taken together, these findings are consistent with H2a, and document the costs of relational incentives when organizations undergo unintended relational disruptions.

2.6.2.2. The Effect of Intended Relational Disruptions

The results for the difference-in-differences analyses of the effect of intended relational disruptions are tabulated in Table 2.5 and Figure 2.5. Table 2.5 examines the effect of intended relational disruptions on service provider performance. The relevant outcome variables are *Call Acceptance* in columns 1 and 2, and *ETA Accuracy* in columns 3 and 4. The coefficient of interest, β_2, in column 2 suggests that intended relational disruptions are associated with performance improvements in terms of a higher call acceptance rate. More specifically, on average, service providers in zones that are subject to an intended network manager change exhibit a 0.6% higher increase in the call acceptance rate relative to service providers for which the network manager did not change. Gauging from the descriptive statistics in Table 2.1, this translates into approximately 171 more services accepted for a zone in the months subsequent to a network manager change, and suggests that an intended relational disruption is able to decrease the number of services to out-of-network service providers by up to 30%.[17]

---

[17] In order to examine the potential mechanisms by which such performance improvements are achieved, I examine whether network manager behaviors are associated with changes. Network managers can influence service providers

Table 2.5. Effects of Intended Relational Disruptions

This table examines the empirical results from testing H2b, and reports the results from OLS regressions of equation (6) in Section 2.6.2 examining the effect of intended relational disruptions. Intended relational disruptions are defined as company-initiated network manager changes. The sample is based on the 12-months surrounding a company-initiated network manager change. For each unique company-initiated network manager change event, *TREAT* is defined as 1 for the service providers in zones that experienced a company-initiated network manager change, and 0 for the service providers in zones that did not experience any network manager change for the same corresponding 12-month time period. *POST* is defined as 1 for the 6 months subsequent to the company-initiated network manager change, and 0 otherwise. The dependent variables are *Call Acceptance* (in columns 1 and 2) and *ETA Accuracy* (in columns 3 and 4). Control variables include *# of SPs in Zip5*, *# of Potential SPs in Zip5*, and all cost contract terms. Year-month- , network manager-, and service provider- fixed effects are included throughout. Robust standard errors are reported in brackets and are clustered at the service provider- and year-month level. ***, **, and * indicate significance at the 1, 5, and 10% level, respectively.

	Intended NM Changes			
	(1)	(2)	(3)	(4)
VARIABLES	Call Acceptance		ETA Accuracy	
POST	0.24*	0.15	-0.01	0.02
	(0.123)	(0.127)	(0.207)	(0.210)
TREAT*POST		0.60***		-0.23
		(0.225)		(0.233)
Constant	10.53***	10.59***	80.33***	80.31***
	(3.092)	(3.093)	(2.441)	(2.439)
Observations	330,399	330,399	314,543	314,543
Controls	YES	YES	YES	YES
NM FE	YES	YES	YES	YES
NM Change-SP FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Adjusted R-squared	0.0346	0.0347	0.0231	0.0231

either through (1) their ability to negotiate explicit contract terms, and/or (2) their ability to assign services, and/or (3) the selection of new service providers. Accordingly, I first estimate the empirical model by replacing the outcome variable with cost reimbursement terms. Untabulated results suggest that cost reimbursement terms are not affected by intended manager changes. Second, I estimate the empirical model on data at the market area-level by replacing the outcome variable with (1) a dummy equal to 1 if the "Preferred" status service provider changed in the market area, and (2) the number of newly selected service providers in the market area. In untabulated results, the coefficient β_2 is positively significant when the relevant dependent variable is the number of newly selected service providers in the market area. This suggests that network managers primarily resort to selection of new service providers to induce changes in service provider performance.

Figure 2.5. Effect of Intended Relational Disruptions

Figure 2.5a graphically examines the empirical results from testing H2b, and reports the results from OLS regressions of equation (6) in Section 2.6.2 by plotting the coefficients with the corresponding 95% confidence intervals of the interactions between TREAT and each month in the 12-month period (instead of interacting *TREAT* only with *POST*). The sample is based on the 12-months surrounding a company-initiated network manager change. For each unique company-initiated network manager change event, *TREAT* is defined as 1 for the service providers in zones that experienced a company-initiated network manager change, and 0 for the service providers in zones that did not experience any network manager change for the same corresponding 12-month time period. The dependent variable is *Call Acceptance*. Control variables are the same as in Table 6, and are defined in Panel H of Table 2.1. The solid line at 0 on the x-axis represents the month immediately prior to the company-initiated network manager change. The dotted line represents the approximate initial announcement date of the company-initiated network manager change. Figure 2.5b plots the mean Call Acceptance of service providers that underwent a company-initiated network manager change (i.e. TREAT), and service providers that did not undergo any network manager change (i.e. Control). The solid line at 0 on the x-axis represents the month immediately prior to the company-initiated network manager change. The dotted line represents the approximate initial announcement date of the company-initiated network manager change.

Figure 2.5a: Plot of coefficient for monthly interaction with TREAT

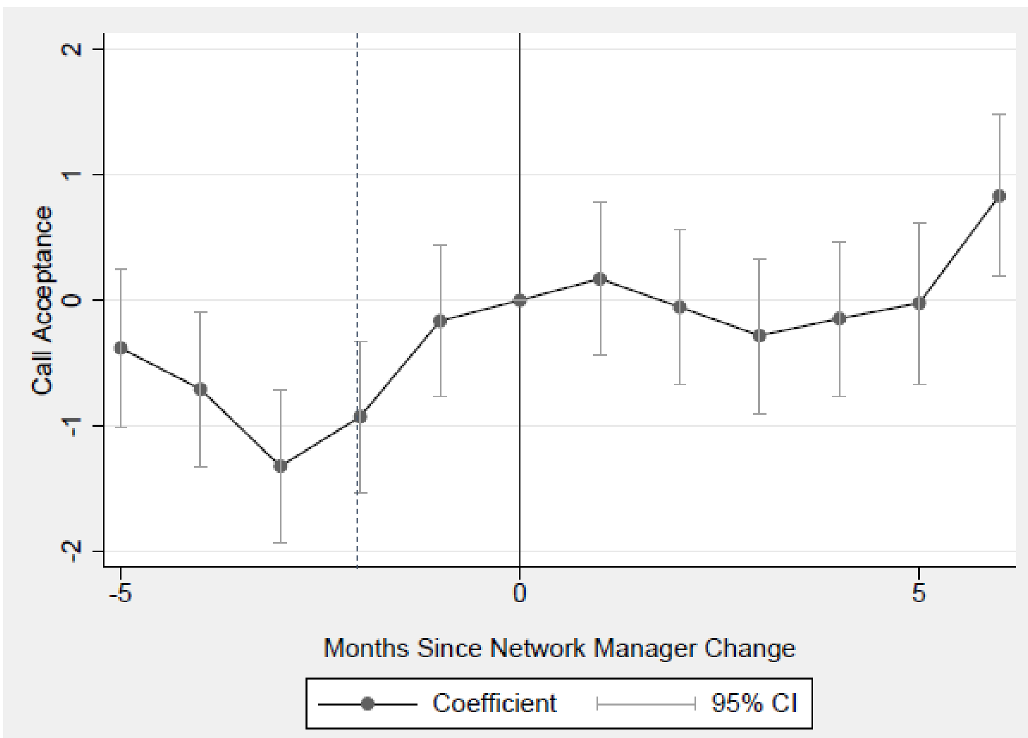
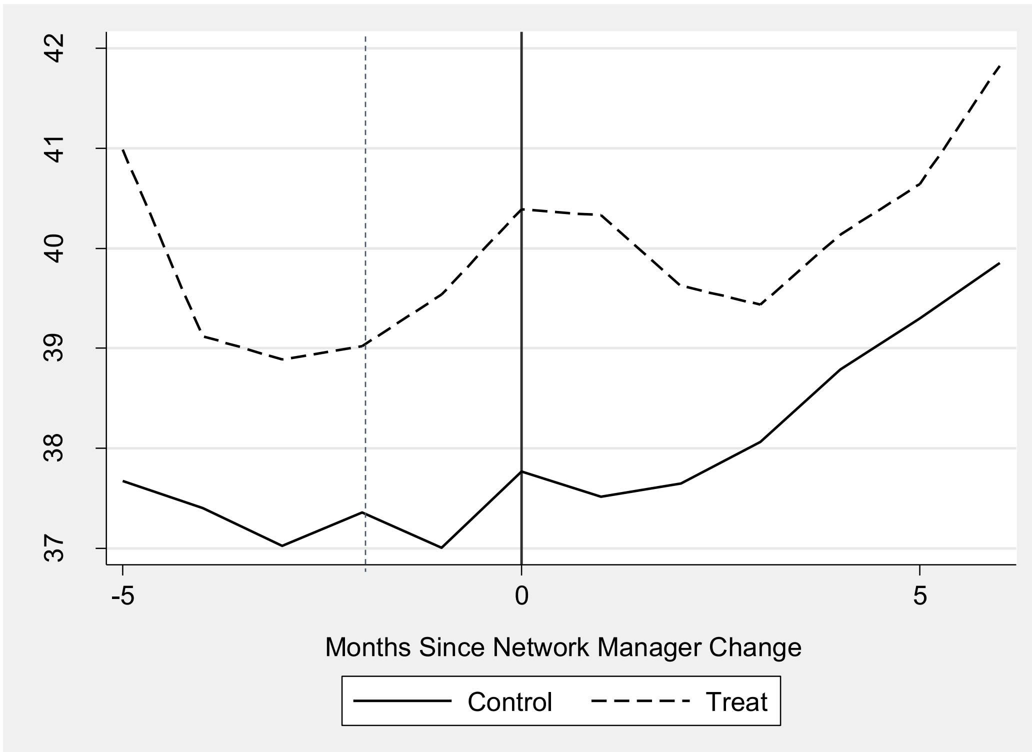


Figure 2.5. Effect of Intended Relational Disruptions (Continued)

Figure 2.5b: Plot of Monthly Mean for Treatment and Control SPs



Similar to Figure 2.4a, Figure 2.5a graphically illustrates the results from the same regression as in equation (6) examining the effects of intended relational disruptions allowing for interactions with *TREAT* for all 12 pre- and post- manager change months. The omitted interaction (i.e. the base line) is the month immediately prior to the network manager change. Specifically, Figure 2.5a plots the coefficients for all interactions with the corresponding 95% confidence intervals. Consistent with H2b, these findings suggest that the organization initiates manager changes when zones suffer from underperformance. Intended relational disruptions are preceded by service provider underperformance which reverses to a similar level as the control service providers subsequent to the manager change.[18]

If intended network manager changes constitute a control mechanism by which incumbent service providers expect a reevaluation of their relationship by the newly-assigned manager, the relevant event date for the change in service provider performance should be the *announcement* date and not the actual date of the manager change. Service providers are likely to exert greater effort to maintain their position in the network upon knowing that there will be a manager change. Consistent with this, the figure suggests that the largest decline in service provider performance is observed 3 months prior to the intended network manager change, and service providers start to show performance improvement even before the actual manager change date. The date subsequent to the largest performance decline coincides with the approximate initial announcement date for

[18] The service provider performance improvements subsequent to intended relational disruption should be primarily driven by network manager changes that ROAD initiated due to "strategic initiatives" related to performance concerns, and implies that intended relational disruptions due to the network managers' promotion reasons should not be preceded by performance deteriorations. To test this prediction, in untabulated tests, I examine differences in performance consequences subsequent to intended relational disruptions distinguishing between the two different reasons, separately. Consistent with expectations, the performance pattern as documented in Figure 2.5 is evident for intended relational disruptions due to ROAD's strategic initiatives, but not due to network managers' promotions. More importantly, network managers that are subsequently promoted to another position do not exhibit significant performance differences with the service providers in zones that did not experience any network manager changes before and after the change. Collectively, this suggests that network manager promotions are planned by ROAD such that no performance disruptions are evident in the vendor network.

an intended network manager change which happens about one or two months before the actual network manager change event.[19] The dotted line in Figure 2.5 represents the approximate initial announcement date of the network manager change. In addition, Figure 2.5b plots the mean *Call Acceptance* separately for service providers that underwent an intended network manager change (i.e. Treat) and service providers that did not undergo any change (i.e. Control). The performance of treated service providers exhibits a significant decline prior to the intended network manager announcement date, and starts to improve thereafter.

There are two additional notable observations from Figure 2.5b. First, the figure suggests that subsequent to the actual manager change event date control service providers start to improve their performance. This may be either due to (1) a seasonal effect – if the intended manager change event dates are clustered in a particular business season where the upward performance trend is correlated with favorable market conditions, or (2) a spillover effect – if the intended network manager change event date constitutes a signal to the control service providers of a potential change in their network manager as well. The intended manager change event dates are spread out evenly, and are not clustered around a particular business season such that the performance improvement is not likely driven by seasonal effects. Instead, the evidence suggests that intended manager changes have positive spillover effects on service providers that are not directly subject to an intended manager change. Second, the figure also suggests that intended network manager changes are associated with some adjustment costs in that performance immediately declines for treated service providers subsequent to the actual manager change event date. However, these

[19] The data do not allow me to retrieve the exact *announcement* date of the intended network manager change. The approximate initial announcement for an intended network manager change is obtained from interviews with network managers. The announcement of an intended network manager change is made via an internal e-mail circulation to the service providers. In additional untabulated tests, I re-estimate equation (6) by approximating the date by which service providers learned about the intended network manager change – i.e. the announcement date – as two months prior to the actual change. The coefficient on the interaction *TREAT*POST* increases to 0.9 which suggests that service providers react to the intended network manager change announcement.

performance declines are only minor and temporary such that performance between treated and control service providers are not statistically different from each other following the actual manager change event date as illustrated in Figure 2.5a. Taken together, the results suggest that service providers respond to the announcement of the intended network manage change and show performance improvement until the actual manager change happens.

2.7. Additional Tests

2.7.1. Cross-sectional Tests for the Effect of Relational Disruptions

In additional cross-sectional tests, I examine which service providers are more susceptible to the performance changes due to intended and unintended network manager changes. In particular, I am interested in whether service providers not facing competition from other service providers in ROAD's network exhibit different performance changes. Table 2.6 reports the results of these cross-sectional tests. In Panel A (Panel B), I examine whether the performance deteriorations (improvements) resulting from unintended (intended) manager changes are borne differently depending on whether the service provider faces competition from other potential service providers in its market area. Service providers constituting the single service provider for a market area may have relatively more bargaining power in the relationship with ROAD due to lacking outside alternative potential service providers in the area. Accordingly, such service providers may have less incentives to improve performance.

Regarding the unintended relational disruptions, I examine if service providers with relatively more bargaining power have less incentives to quickly adjust to a new relationship. If so, the performance deteriorations subsequent to unintended relational disruptions should be primarily driven by service providers enjoying more bargaining power. To examine this prediction,

Table 2.6. Cross-Sectional Tests for the Effect of Relational Disruptions

This table examines the empirical results from the cross-sectional tests as described in Section 2.7.1, and reports the results from OLS regressions examining the effect of relational disruptions on different subsamples partitioned based on the extent of the service provider's bargaining power in the contractual relationship with ROAD. Panel A (Panel B) re-estimates Table 2.4 (Table 2.5) based on the different subsamples. The partitioning variable is whether the service provider operates in a market area with more than one service provider or not (i.e. *Single SP*). Columns 1 through 4 (columns 5 through 8) are based on the subsample when *Single SP* is equal to one (zero). Service providers operating in market areas without any other service providers are regarded as enjoying greater bargaining power in relation with ROAD. The dependent variables are *Call Acceptance* (in columns 1, 2, 5, and 6) and *ETA Accuracy* (in columns 3, 4, 7, and 8). Control variables are the same as in Tables 2.4 and 2.5. Year-month-, network manager-, and service provider- fixed effects are included throughout. Robust standard errors are reported in brackets and are clustered at the service provider- and year-month level. ***, **, and * indicate significance at the 1, 5, and 10% level, respectively.

Panel A: Unintended Network Manager Changes

	Unintended NM Changes & Single SP				Unintended NM Changes & Multiple SP			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Call Acceptance		ETA Accuracy		Call Acceptance		ETA Accuracy	
POST	0.15	0.21**	0.11	0.17	0.21	0.28*	-0.18	-0.20
	(0.100)	(0.103)	(0.155)	(0.157)	(0.155)	(0.159)	(0.240)	(0.244)
TREAT*POST		-0.51***		-0.53***		-0.59**		0.17
		(0.181)		(0.201)		(0.292)		(0.333)
Constant	50.93***	50.95***	77.41***	77.43***	42.20***	42.08***	73.02***	73.06***
	(2.449)	(2.448)	(2.714)	(2.712)	(5.825)	(5.819)	(5.304)	(5.304)
Observations	457,038	457,038	434,495	434,495	184,967	184,967	176,137	176,137
Controls	YES	YES	YES	YES	YES	YES	YES	YES
NM FE	YES	YES	YES	YES	YES	YES	YES	YES
NM Change-SP FE	YES	YES	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES	YES	YES
Adjusted R-squared	0.0427	0.0427	0.0208	0.0210	0.0375	0.0376	0.0305	0.0305

Table 2.6. Cross-Sectional Tests for the Effect of Relational Disruptions (Continued)

Panel B: Intended Network Manager Changes

	Intended NM Changes & Single SP				Intended NM Changes & Multiple SP			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Call Acceptance		ETA Accuracy		Call Acceptance		ETA Accuracy	
POST	0.21	0.12	0.08	0.18	0.29	0.27	-0.20	-0.29
	(0.149)	(0.153)	(0.253)	(0.258)	(0.222)	(0.228)	(0.367)	(0.373)
TREAT*POST		0.70**		-0.71**		0.17		0.64
		(0.274)		(0.279)		(0.409)		(0.462)
Constant	16.30***	16.36***	80.27***	80.22***	-10.20	-10.19	79.51***	79.54***
	(3.071)	(3.068)	(2.594)	(2.592)	(7.785)	(7.783)	(6.934)	(6.932)
Observations	235,238	235,238	223,783	223,783	95,161	95,161	90,760	90,760
Controls	YES	YES	YES	YES	YES	YES	YES	YES
NM FE	YES	YES	YES	YES	YES	YES	YES	YES
NM Change-SP FE	YES	YES	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES	YES	YES
Adjusted R-squared	0.0357	0.0358	0.0224	0.0227	0.0321	0.0321	0.0293	0.0295

I re-estimate Table 2.4 on the *Single SP* (columns 1 through 4) and *Multiple SP* (columns 5 through 8) subsamples. The results are tabulated in Panel A of Table 2.6, and show that the performance deteriorations in terms of *Call Acceptance* are evident in both subsamples (i.e. the coefficient on the interaction is significantly negative in columns 2 and 6, and is not statistically different from each other).[20] When *ETA Accuracy* is considered, the difference in the coefficients on the interaction based on the *Single SP* subsample (in column 4) and *Multiple SP* subsample (in column 8) is statistically significant at the 5% level. Specifically, the results indicate that service providers not facing any competition from other potential service providers for service assignments in their market area exhibit worse *ETA Accuracy* subsequent to an unintended network manager change.

Regarding intended relational disruptions, I examine whether the performance improvements subsequent to intended relational disruptions are primarily driven by service providers with less bargaining power – i.e. service providers serving market areas with other service providers (*Multiple SP* subsample). To examine this prediction, I re-estimate Table 2.5 on the *Single SP* (columns 1 through 4) and *Multiple SP* (columns 5 through 8) subsamples. The results are tabulated in Panel B of Table 2.6, and show that the performance improvement in terms of *Call Acceptance* is statistically significant only in the *Single SP* subsample. However, there is no statistical difference between the coefficients on the interaction based on the *Single SP* subsample (in column 2) and *Multiple SP* subsample (in column 6) which suggests that regardless of the degree of bargaining power, service providers have incentives to perform better on *Call Acceptance*. More interestingly, however, the results when the performance measure is *ETA*

[20] To test whether the coefficients in columns 2 and 6 are significantly different from each other, I estimate a regression model on the entire sample whereby I include the variable *Single SP*, and the interactions of *Single SP* with POST, TREAT, and TREAT*POST as additional independent variables. I interpret the coefficient on TREAT*POST in columns 2 and 6 to be statistically different from each other if the coefficient on the three-way interaction between TREAT, POST, and *Single SP* is statistically significant.

Accuracy exhibit statistically significant differences. The difference in the coefficients on the interaction based on the *Single SP* subsample (in column 4) and *Multiple SP* subsample (in column 8) is statistically significant at the 5% level. Taken together, both results suggest that service providers with relatively high bargaining power in relation with their network managers exhibit worse performance quality (measured as ETA Accuracy) subsequent to any type of relational disruption, and suggests that such service providers have relatively less incentives to adjust to relationship renewals.

2.7.2. Long-term Effects of Relational Disruptions

My main analyses in Tables 2.4 and 2.5 examining the effects of unintended and intended relational disruptions are designed as a difference-in-differences whereby I compare zones that experienced either an unintended or intended relational disruption (i.e. the treatment in Table 2.4, and 2.5, respectively), to zones that did not experience any relational disruption at all (i.e. the control) covering a 12-months period surrounding the event month of the relational disruption. In order to examine the long-term performance effects of unintended relational disruptions, I estimate the following panel regression model where I am interested in the performance differential in the periods following an unintended relational disruption relative to the periods following an intended relational disruption:

$$\begin{aligned} \textit{SP Performance}_{i,t} = {} & \alpha + \beta_1 * \textit{Single SP}_{i,t} + \beta_2 * \textit{POST\_Unintended}_{z,t} \\ & + \beta_3 * \textit{Single SP}_{i,t} * \textit{POST\_Unintended}_{z,t} \\ & + \textit{Controls}_{i,t} + u_t + v_i + w_m + \varepsilon_{i,t} \qquad (7) \end{aligned}$$

POST_Unintended is defined as a dummy variable equal to 1 that flags the periods subsequent to an unintended relational disruption. The definition of S*ingle SP* is the same as before and the inclusion of control variables is the same as in equation (6). The dependent variable is service provider performance measured as either *Call Acceptance* or *ETA Accuracy*. A significant negative coefficient of β_2 would suggest that periods subsequent to unintended relational disruptions are associated with relatively lower performance. β_3 examines whether the performance differentials from unintended relational disruptions are borne out differently depending on whether the service provider faces competition from other potential service providers in its market area or not. The results from estimating equation (7) are tabulated in Table 2.7, and confirm the under-performance following unintended relational disruptions. In contrast with the short-term performance effect analyses from the difference-in-differences research design, the under-performance over the long-term is not primarily driven by service providers that enjoy greater bargaining power as evidenced by the insignificant coefficient on the interaction term.

2.8. Conclusion

In this paper, I document the benefits and costs of using relational incentives in motivating performance. In doing so, I use data from a company that outsources its core business operations to a nationwide vendor network which, instead of relying on explicit performance-based contracts, encourages relationship-building between network managers and well-performing vendors. The empirical findings demonstrate the prevalence of relational incentives in the absence of explicit performance-based contracts, and that relational disruptions via changing managers can be used as a viable control mechanism to sustain high levels of vendor performance. However, the use of relational incentives are also associated with costs when faced with unintended disruptions in relationships. My empirical findings show that vendor performance deteriorates subsequent to unintended manager changes suggesting that relationship-building with new managers is subject

Table 2.7. Long-term Effects of Relational Disruptions

This table examines the long-term effects of relational disruptions, and reports the results from OLS regressions of equation (7) in Section 2.7.2. The sample constitutes monthly panel data for all service providers in the time period between January 2011 and May 2016. The dependent variable is either *Call Acceptance* (in columns 1 and 2) or *ETA Accuracy* (in columns 3 and 4). *Single SP* is a dummy equal to 1 if the service provider is the only service provider in its market area. *POST_Unintended* is a dummy equal to 1 if the period is subsequent to a company-initiated network manager change event, and 0 if the period is subsequent to an employee-initiated network manager change. Control variables include *# of SPs in Zip5*, *# of Potential SPs in Zip5*, and all cost contract terms. Service provider-, network manager-, and year-month fixed effects are included throughout. Robust standard errors are reported in brackets and are clustered at the service provider- and year-month level. ***, **, and * indicate significance at the 1, 5, and 10% level, respectively.

	(1)	(2)	(3)	(4)
	Call Acceptance		ETA Accuracy	
Single SP		-0.57		-0.22
		(0.690)		(0.483)
POST_Unintended	-4.32***	-4.68***	-1.43*	-1.49*
	(1.043)	(1.104)	(0.858)	(0.899)
Single SP *POST_Unintended		0.55		0.09
		(0.534)		(0.397)
Constant	40.64***	41.07***	77.69***	77.97***
	(4.509)	(4.587)	(2.356)	(2.430)
Observations	199,035	199,035	190,127	190,127
Controls	YES	YES	YES	YES
NM FE	YES	YES	YES	YES
SP FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Adjusted R-squared	0.667	0.667	0.301	0.301

to adjustment costs. Moreover, I show that such costs are more pronounced for vendors that have relatively more bargaining power in the contractual relationship with the organization providing insights into what types of relationships are more susceptible to sudden, unintended disruptions in incumbent relationships.

This study is subject to several limitations. Despite the advantages of the field setting to examine the performance consequences of changing relationships, as my findings are based on archival data from a single company, they may not generalize to other settings. My findings are most likely to generalize to settings in which managerial relationships are only weakly governed by explicit performance-based conditions due to high contracting costs. Therefore, my study cannot address the performance consequences from changing relationships under alternative contracting relationships. Despite the limitations of this study, my hope is that it will be a first step in documenting the importance of relationship dynamics in incentive contracting, and how disruptions in incumbent relationships affect subsequent performance when relational incentives constitute an integral part of performance incentive systems in organizations. Disruptions to incumbent relationships are unavoidable as organizations operate in a dynamic environment with changing strategic objectives and market conditions. Future research could examine the mechanisms by which relational disruptions have an effect on subsequent performance, and could attempt to disentangle different consequences depending on alternative contracting environments. Moreover, having documented the costs to the use of relational incentives due to unintended relational disruptions, future research could address potential solutions to mitigate such costs in the design of management control systems.

CHAPTER 3

MANAGING THROUGH ORGANIZATIONAL CHANGE: EMPLOYEE ALIGNMENT IN THE PRESENCE OF UNEXEPCTED CAREER CONCERNS

3.1. Introduction

This study examines performance consequences in the presence of unexpected career concerns. A large body of analytical research examines career concerns where career-related incentives arise due to the existence of a labor market that allows for the valuation of the agents' ability. In these models, agents derive incentives to exert better performance in order to mitigate adverse consequences over time as the market learns about their true ability. In line with this stream of research, a number of empirical studies have examined career concern-related incentive effects and its relationship with incentive contracting, especially in the context of executives (e.g. Fama 1980; Holmstrom 1999; Gibbons and Murphy 1992) or professionals (e.g. Hong, Kubik, and Solomon 2000; Hong and Kubik 2003).

Yet, career-related incentives that hinge on the market's valuation of the agent's ability are less prevalent for lower-level employees as their ability is relatively easily replaceable. Instead, such employees are subject to rather *unexpected* career concerns that arise due to institutional reasons. For example, unexpected career concerns arise when employees experience feelings of job insecurity due to sudden organizational changes that can accompany a discontinuation of existing business units or product lines. A focus on the effects from unexpected career concerns is particularly important as recent rapid developments in the fields of technology, transportation, and communication force organizations to undergo frequent organizational changes that can include restructuring efforts. However, we do not yet fully understand whether and how such unexpected career concerns affect employee behaviors.

This paper aims to fill this void by examining how employees change their behaviors when subject to unexpected career concerns. We study a rental car company with store locations across US airports at which its management made an explicit announcement to all its employees of its intent for a merger. Consistent with the literature on horizontal mergers (Fee and Thomas 2004), industry experts predict a consolidation of duplicate resources which unexpectedly create heightened career concerns for current employees.[1] Exploiting rich micro-level data on employee performance surrounding the merger announcement, we are able to examine how unexpected career concerns affect employee behaviors and performance. Thereby, our study provides a more comprehensive view of career-concern-related incentive effects within organizations.

In particular, our study examines the incentive effects and effort allocation effects associated with unexpected career concerns. First, we examine how overall employee performance is affected when subject to unexpected career concerns. Unexpected career concerns may be associated with two countervailing incentive effects. On the one hand, employees may have incentives to exert better performance in order to minimize potential layoff risks. On the other hand, the associated termination threat may discourage employees to perform at the current organization and result in adverse performance consequences. Second, if performance is measured based on multiple measures, unexpected career concerns may also create incentives to allocate effort across these measures differently. For example, unexpected career concerns may create pressures to mitigate layoff risk that can incentivize myopic behaviors to fixate on short-term at the expense of long-term performance. Thereby, we shed light on whether employees consider the

[1] Mergers and acquisitions comprise a significant business strategy for an increasing number of organizations. For example, according to the IMAA, since 1985, more than 300,000 mergers and acquisitions transactions have been announced with a known value of almost 33,200 billion US dollars (https://imaa-institute.org/m-and-a-us-united-states/). Moreover, a larger number of anecdotal evidence suggests that mergers and acquisitions are associated with subsequent restructuring efforts, and accompany layoff risks for employees. (For example, http://www.businessinsider.com/signs-your-company-is-conducting-mass-layoffs-2015-10)

sensitivity-congruity trade-off of performance measures, and allocate their effort considering the different measure properties when subject to unexpected career concerns.

Several studies in economics, management, and accounting highlight the advantages of employee alignment – the extent by which employees are aligned with the overall organizational objectives and strategy – as an important organizational control channel (as opposed to the alignment of incentives via explicit contracting mechanisms).[2] Moreover, anecdotal evidence suggests that greater employee alignment is correlated with heightened retention rates and better execution.[3] Building on this argument, we further hypothesize and investigate empirically whether employee alignment constitutes a significant moderating force in explaining the effects arising from unexpected career concerns. To do so, we first operationalize employee alignment based on the notion of the 'clan mechanism' articulated by Ouchi (1979) which describes an informal control system sustained by shared values of the organization's constituents that is consistent with the overall organizational objectives and strategy. We then show whether and how employee alignment is associated with different performance consequences in the presence of unexpected career concerns.

Our research site provides us with an ideal setting to address our research question for several reasons. First, the management of the company made an internal announcement to all company-affiliated employees regarding its plans to merge with another major rental car company. Especially customer-facing employees were subject to termination threats due to the possibility that the merger would involve restructuring efforts to combine existing store locations at each

---

[2] See literatures in economics (e.g. Prendergast 2008), accounting (Campbell 2012; Abernethy et al. 2015), and management (Kaplan and Norton, 1996; Joshi, Kathuria, and Porth, 2003) that point to the importance of goal alignment in organizational management systems for increasing organizational performance.

[3] A survey of nearly 100 respondents from large companies revealed that firms with higher-performing employees report "a formal linkage between corporate and individual goals", and that such firms were "2.2 times more likely to be top performers than their peers" (Available at https://hbr.org/sponsored/2016/06/how-employee-alignment-boosts-the-boosts-the-bottom-line).

airport. Accordingly, we exploit the internal merger announcement date as a trigger date to proxy for heightened unexpected employee career concerns. Second, we examine performance effects of the customer-facing employees at each store location whose performance is measured at the individual-level based on two different performance measures: a relatively sensitive short-term sales-based measure and a relatively congruent long-term customer satisfaction measure. Examining performance measures that exhibit different properties allows us to study not only overall performance incentive effects, but also effort allocation effects. [4] Third, our research site implements periodic employee engagement surveys at each store location. Using the survey results, we can distinguish between store locations with higher and lower levels of employee alignment. Finally, alternative formal management control channels including the management team and the design of employee incentive contracts remain constant over the course of the merger event. Therefore, the empirically documented performance effects subsequent to the merger announcement are not prone to changes in explicit management control channels.

Our findings document significant incentive and effort allocation effects subsequent to the merger announcement. First, we find that unexpected career concerns are associated with positive incentive effects. In particular, we show that performance on both performance measures increases significantly subsequent to the merger announcement. This result is consistent with predictions from prior literatures that career concerns may substitute explicit incentives and provide additional performance incentive effects due to the threat of replacement. Moreover, we find corroborating empirical evidence that the source for such incentive effects are unexpected career concerns. Specifically, we distinguish between airport store locations where employees are more or less

[4] Empirical studies that examine career concern-related incentive effects for executives, or CEOs are subject to the limitation of having to rely on aggregate firm performance measures. In contrast, our empirical results are based on individual-level performance measures, and are, thus, not confounded by noise in the performance measure that captures group- or firm-level outcomes.

likely to be subject to the merger-related career concerns. Employees at airport store locations where the target rental car company also maintains store operations face relatively higher merger-related career concerns as the perceived likelihood of potential restructuring efforts is relatively higher. The results show that the positive incentive effects are driven by employees at such airport store locations where the merger-related career concerns are more predominant.

Second, we find evidence that unexpected career concerns are associated with effort allocation effects. In particular, we examine the extent of improvement on the sales-based measure relative to the customer satisfaction measure, and whether it varies between store locations with higher and lower levels of employee alignment. We find that the improvement of the customer satisfaction measure is significantly greater than the improvement of the sales-based measure for employees at locations that exhibit higher levels of employee alignment. Conversely, employees at locations that exhibit lower levels of employee alignment improve significantly greater on the sales-based measure than the customer satisfaction measure. Collectively, these findings suggest that, when subject to unexpected career concerns, employees may have incentives to fixate effort levels on relatively short-term performance measures at the expense of decreasing effort towards more long-term performance measures. More importantly, our results suggest that greater employee alignment can mitigate such myopic employee behaviors due to unexpected career concerns. Further additional tests corroborate this finding. In particular, we exploit the different subcomponents for overall customer satisfaction to identify more direct proxies for employee effort. Whereas subcomponents such as “speed of service”, and “staff courtesy” are directly reflective of the employee’s effort, other subcomponents such as “billing as expected”, and “vehicle condition” are managed centrally by the organization. We show that the results are driven by the former which provides further evidence that the performance effects are a result of corresponding employee behaviors.

This study makes contributions to largely three streams of literatures. First, this study contributes to the large body of literatures that examine the optimal design of employee incentive systems. A large stream of literature in economics and accounting is devoted to studying the design of optimal contracts to mitigate incentive problems in economic relationships. In the standard principle-agent framework, the classic theoretical insight suggests that employee performance should be evaluated using performance measures that are informative about managerial effort or talent (Ross 1973; Holmstrom 1979). Many studies have examined the provision of incentives within the bounds of the explicit contract. For example, research investigates performance measure properties, and the optimal balance of the combination of different performance measures.[5] A significant number of research is devoted to setting appropriate targets based on the performance measures used.[6] Yet, another stream of related research examines how to mitigate contracting limitations due to the incompleteness of performance measures through the use of relative performance evaluation and/or subjectivity.[7] This study sheds light on incentive effects that are outside of the bounds of explicit contracting mechanisms by empirically documenting performance consequences that arise due to unexpected career concerns. Thereby, we add to the literature by providing a more comprehensive perspective of employee performance incentives. Moreover, this

---

[5] See, for example, prior work by Lambert and Larcker (1987); Banker and Datar (1989); Feltham and Xie (1994); Ittner, Larcker, and Rajan (1997); Baker (2000); Banker, Potter, and Srinivasan (2000); Datar, Kulp, and Lambert (2001); Dikolli (2001); Core, Guay, and Verrecchia (2003); Dutta and Reichelstein (2005); Moers (2006); Bouwens and Van Lent (2007); Raith (2008); Indjejikian and Matejka (2009); etc.

[6] See, for example, prior work by Aranda, Arellano, and Davila (2014); Indjejikian, Matejka, and Merchant (2014); Indjejikian, Matejka, and Schloetzer (2014); etc.

[7] For literatures on the use of relative performance evaluation, see Antle and Smith (1986); Janakiraman, Lambert, and Larcker (1992); Albuquerque (2009) etc. For literatures on the use of subjectivity, see Baker, Gibbons, and Murphy (1994); Baiman and Rajan (1995); Ittner, Larcker, and Meyer (2003); Gibbs, Merchant, Van der Stede, and Vargus (2004); Bol (2008); Ederhof (2010); Bol (2011); Bol and Smith (2011); Hoeppe and Moers (2011) etc.

study contributes to the growing stream of literature that examines alternative means as a viable control mechanism to maximize desirable organizational outcomes other than formal contracting channels. For example, research shows that the delegation of authority via organizational design choices,[8] and strengthening of relationships and/or social norms[9] may address such limitations in the design of explicit contracts. Our findings contribute to these studies by shedding light on how firms can benefit from greater employee alignment with overall organizational objectives and strategy in the presence of unexpected career concerns.

Second, this study contributes to the literature on management control systems. The existing literature primarily emphasizes the role of management control systems as a significant determinant for successful firm strategy execution. Successful management control systems shape the culture of the firm, and guide employees to execute effort on desirable behaviors that is consistent with the organizational objectives (Simons 1987; Sandino 2007). Accordingly, prior management accounting research has primarily studied the deliberate choice of management control system of firms, and factors that are associated with different types of management control systems adopted by firms. By examining employee behaviors subject to a sudden organizational change event (i.e. a merger announcement), this study highlights the possibility that existing management control systems may result in distorted employee incentives with constantly changing business strategies to adapt to the dynamic business environment.

Finally, our findings contribute to the literature on mergers and acquisitions. A large number of works investigate factors that are associated with post-acquisition firm performance.[10]

---

[8] See, for example, prior work by Jensen and Meckling (1992); Baiman, Larcker, and Rajan (1995); Nagar (2002); Campbell, Datar and Sandino (2009); Indjejikian and Matejka (2012); etc.
[9] See, for example, prior work by Baker, Gibbons, and Murphy (2002); Cardinaels and Yin (2015); Abernethy, Bouwens, Hofmann, and van Lent (2015); etc.
[10] See, for example, prior work by Larsson and Finkelstein (1999); Bowman and Singh (1993); Anand and Singh (1997); Kim and Finkelstein (2009); etc.

Whereas prior literature has primarily focused on assessing overall aggregated firm performance of the newly created firm, there is a lack of understanding regarding whether and how the performance of individual employees is affected by a merger decision. By documenting how individual employee performance is affected by career concerns due to a merger announcement, our study enhances the understanding of post-acquisition firm performance.

The remainder of the paper is organized as follows. In Section 3.2, we review the prior literature and develop our hypotheses. Section 3.3 describes our research setting and Section 3.4 describes our empirical research design. We explain our empirical results in Section 3.5, and conclude with Section 3.6.

3.2. PRIOR LITERATURE AND HYPOTHESES DEVELOPMENT

3.2.1. Unexpected Career Concerns and Incentive Effects

A number of studies investigate performance incentives due to career concerns. However, the prior analytical literature primarily focuses on one particular type of career concern where the assumption of a well-functioning labor market for managerial talent is crucial. In other words, the fundamental incentive problem embedded in such career concerns arises due to the expected valuation of the manager's true ability in the labor market over time. For example, Fama (1980) argued that explicit incentive contracts are not necessary because managers are disciplined through the managerial labor market such that superior performances will generate high wage offers; and poor performances will result in low offers. Since the market infers the ability of managers by gauging the overall level of compensation, the manager is also incentivized to exert greater effort through the signaling aspects attached to higher compensation levels. Holmstrom (1999) demonstrates the dynamic incentive problem analytically. His model assumes that output is a function of the manager's true ability and effort, and that the market only observes the output level, but does not observe the manager's true ability and effort. Over time, when more output data points become available, the market learns about the manager's true ability. The model demonstrates that the optimal level of managerial effort declines as the market's learning progresses (i.e. the market can approximate the manager's true ability more accurately as time progresses). Gibbons and Murphy (1992) examine how such career concerns interact with the design of optimal incentive contracts. Specifically, they define career concerns as the expected effects of current performance on future compensation, and show that career concerns can still create important incentives, even in the presence of incentive contracts.

Despite the wealth of theoretical justification for such career-concern-induced incentive effects, due to empirical research design limitations, there is only limited empirical evidence, primarily on CEOs or professionals, that examines the incentive effects due to career concerns, and their association with incentive contract design choices. Using the Compustat population of firms, Matejka, Merchant, and Van der Stede (2009) show that loss-making firms put more emphasis on nonfinancial performance measures in their annual bonus plans. Arguing that managers at loss-making firms are likely to leave the firm in the near future (i.e. have a short employment horizon), they suggest that employment horizon concerns affect the relative emphasis on financial versus nonfinancial performance in annual bonus plans. Hallman, Hartzell, and Parsons (2011) exploit industry-specific organizational features that CEOs at certain companies are much harder to terminate than at other firms. They show that firms take into account the incentive effects of such inherent termination threats in the design of their financial incentives by showing that the financial incentives at Real Estate Limited Partnerships (RELPs) where termination threats are less credible, exhibit higher pay-for-performance sensitivity. Using data on security analysts, Hong, Kubik, and Solomon (2000) show that inexperienced analysts are more prone to herding behaviors in terms of issuing forecasts that are more timely and closer to the consensus. These findings suggest that security analysts are subject to implicit career concern incentives by trying to manage their reputation in the labor market.

Contrary to the focus of the bulk of career concern-related literatures, anecdotal evidence suggests that the majority of employees are subject to rather *unexpected* career concerns that do not critically hinge on the existence of a well-functioning labor market. For example, unexpected career concerns arise when employees experience feelings of job insecurity due to sudden organizational changes that can accompany a discontinuation of existing business units or product lines. Under such circumstances, the career concerns (primarily for lower-level employees) arise due to rather "exogenous" reasons, and is independent of the employees' concern of how his/her ability will be valued in the labor market. In fact, many modern firms operate in a fast-paced dynamic business environment with rapid developments in the fields of technology, transportation, and communication. In order to maintain competitive advantage, adapting to such environmental changes becomes critical, and such efforts frequently involve mergers and acquisitions to increase market share and/or acquire key capabilities in-house. For example, according to the IMAA, since 1985, more than 300,000 mergers and acquisitions transactions have been announced with a known value of almost 33,200 billion US dollars.[11] This study aims to examine the performance effects arising from such unexpected career concerns.[12]

Theoretically, unexpected career concerns are associated with two countervailing incentive effects. On the one hand, the theorized positive incentive effects from prior literatures may also

---

[11] See https://imaa-institute.org/m-and-a-us-united-states/

[12] Unlike the bulk of literatures on expected career concerns that focuses on managers or executives, the subject of interest to examine the incentive effects for unexpected career concerns are likely lower-level employees. In particular, in this study, we consider one type of unexpected career concern that arises from potential layoff decisions due to mergers and acquisitions.

apply to unexpected career concerns. Anecdotal evidence suggests that many of the changes associated with mergers and acquisitions are evolutionary, and that final outcomes are often not known during the negotiation process. This allows for merger-related rumors to spread amongst employees. In order to avoid a potential layoff, employees may have stronger incentives to signal their ability by exerting better performance (even above and beyond what is expected based on their explicit incentive contracts). Accordingly, we state our first hypothesis as follows:

H1: Employees exhibit better performance when subject to unexpected career concerns.

On the other hand, the associated termination threat may discourage employees to perform at the current organization. In fact, the merging process involves the integration of two different entities with distinct organizational cultures. Anecdotal evidence suggests that after a merger, employees often feel that the organization has changed so much that "it is no longer their company" (Ashkanasy and Holmes 1995; Hogg and Terry 2014), and literatures in strategy and management highlight the importance of organizational culture fit as a significant driver of post-merger performance (Weber 1996; Van den Steen 2010; etc.). If employees are dissatisfied with the merger prospects, and their desire to stay with the organization are not sufficient, the merger-related unexpected career concerns may result in rather adverse performance incentives.

3.2.2. Unexpected Career Concerns and Effort Allocation Effects

If performance is measured based on multiple measures, unexpected career concerns may also create incentives to allocate effort across these measures differently (i.e. effort-allocation effects). For example, if unexpected career concerns generate increased performance incentives to avoid potential layoff risk, employees may focus on improving short-term performance. Prior literatures provide supporting evidence that managers have tendencies to engage in rather myopic behaviors when career concerns are present. For example, Chen et al. (2015) look at whether executives exhibit differences in performance when faced with contracts with varying degrees of protection against the downside of potential dismissals (in the form of employment agreements and severance pay agreements). They find that CEOs with less contractual protection (i.e. greater career concerns) are under more pressure to maintain high short-term performance and, thus, are more likely to engage in myopic behavior compared to those with contractual protection. Moreover, González-Uribe, and Groen-Xu (2017) also find that a longer executive contract duration can motivate executives to invest more in innovation because they provide protection against dismissals.

The incentive conflict in allocating effort across different performance activities arises due to the incompleteness of performance measures. As employee effort is inherently unobservable (Holmstrom 1979), incentive contracts rely on a diverse set of performance measures in practice. The optimal performance measure exhibits two desirable attributes. First, it should insure the agent against risk by mitigating the impact of uncontrollable events (i.e. sensitivity). Second, it should achieve interest alignment between the principal and the agent (i.e. congruity). However, any observable performance measure faces different degrees of the sensitivity-congruity trade-off (Banker and Datar 1989; Feltham and Xie 1994; etc.). Whereas sales-based financial measures exhibit relatively higher sensitivity, non-financial performance measures such as customer satisfaction constitute leading indicators for long-term firm performance (Ittner and Larcker 1997), and, thus, exhibit relatively higher congruence. As congruent performance measures are frequently intangible and many of the desired organizational outcomes have a long-term horizon, employees may have incentives to overweight relatively more sensitive performance measures when subject to unexpected career concerns. Such employee behaviors, however, may not be desirable from the overall organization's perspective in that greater effort is diverted from tasks that are detrimental in sustaining the organization's long-term performance.[13]

Prior literatures in economics and management suggest that the extent by which employees are aligned with the overall organizational objectives and strategy (i.e. "employee alignment") may be a significant moderating force in mitigating the incentive conflicts in effort allocation effects arising from unexpected career concerns. We define employee alignment consistent with Ouchi's (1979) notion of "clan control" which hinges on shared values and beliefs among the constituents of an organization as a form of management controls. Embedded in this form of management control is the idea that employees exhibit alignment on preferences with organizational values as opposed to achieving alignment of incentives via explicit contracting mechanisms. For example, Akerlof and Kranton (2005) show that agents who identify with the firm gain utility in taking actions that benefit the firm. Similarly, Van den Steen (2010a, 2010b) demonstrates the benefits of attracting employees with values and beliefs aligned to the firm. Employees that are more aligned with organizational values also exert greater effort, and are associated with greater utility, and coordination as they are more motivated and satisfied in the work environment (Van den Steen 2005). In addition, research shows that firms rely on management controls to improve employee alignment, especially in uncertain and complex decision contexts subject to high levels of contracting difficulty (Snell 1992, Abernethy and Brownell 1997, Prendergast 2011, Campbell 2012, Abernethy, Dekker, and Schulz 2015). Consistent with the stipulated advantages from greater employee alignment, Campbell (2012) shows that improving employee selection mechanisms can be a successful means to do so. Using

[13] This incentive conflict has also been referred to as the "intertemporal choice" problem (Abernethy, Bouwens, and van Lent 2013). The source of the problem lies in that "the course of action that is best in the short-term is not the same course of action that is best over the long-run".

referral source as a proxy for the extent of employee alignment, he shows that referred employees are more likely to make decisions that are organizationally desirable.

Taken together, we hypothesize that employee alignment can mitigate adverse effort allocation incentives to fixate on short-term at the expense of long-term performance measures when subject to unexpected career concerns. Despite abundant research that stipulates the benefits of greater employee alignment, there is only limited research that examines conditions under which employee alignment is associated with positive organizational outcomes. This study fills this void in the literature by examining whether employee alignment can be beneficial in the presence of unexpected career concerns. We formulate our second hypothesis as follows:

H2: When employees are more aligned with the overall organizational objectives, unexpected career concerns result in relatively better performance on congruent measures than on sensitive measures.

3.3. RESEARCH SETTING AND DATA

3.3.1. Company Description

The data obtained for this study are from a rental car company (hereafter, RENT) with store operations across US airports.[14] RENT is one of the largest players in the rental car industry, holding more than one fifth of the entire market share in the US. RENT customers usually make a reservation for a rental car specifying their pick up and return location in advance primarily via online booking channels. During the reservation making process, customers also select their preferred options for their upcoming trip, including vehicle type and additional services such as GPS device, radio, and pre-paid gas. Accordingly, customer-facing employees have only limited ability to improve sales through customer interactions.

[14] Due to data confidentiality reasons, the name of the company, and the exact dates for events remain unidentified. The research site will be referred to as RENT hereafter.

Considering this nature of the industry, RENT mainly relies on two primary performance measures to monitor and incentivize its employees. First, when customers pick up their reserved car at the predetermined location, employees have the opportunity to solicit customers into an upgrade of their initial reservation. These can include an upgrade in the vehicle type or purchases of additional services. Such upgrades are referred to as "upsell" transactions, and result in additional revenue stream from customers. Employees receive commissions based on the number of upsell transactions. Moreover, they may be liable for termination, or demotion to less desirable positions, if they fail to meet pre-determined upsell quotas in a given period. An employee who used to work at RENT described that "if you miss a quota or two, [managers at RENT] stick you at the exit gate or something like that for a week or two," where the employee would be deprived of the opportunity to make additional earnings from commissions. Even experienced employees demonstrating a good service attitude may be fired if they are not able to generate sufficient upsell transactions.

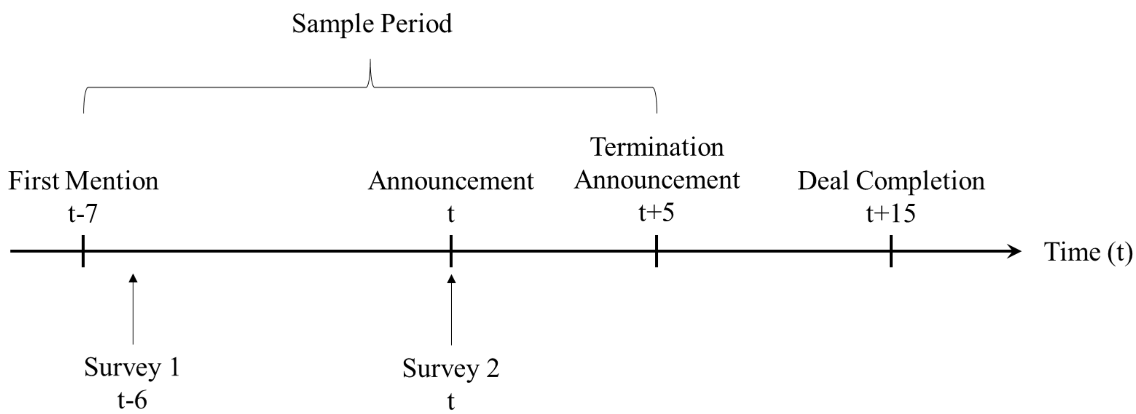
Second, employees can also improve their performance by exerting more effort in improving the overall customer experience. Maintaining higher customer satisfaction levels has been a core of RENT's business model, because it allows the company to attract more customers and charge a higher premium in an industry in which the focal good itself (i.e., renting a car) is a commodity. For example, one employee expresses the importance of customer satisfaction by saying: "If they leave here unhappy, we know they won't come back, and we just cannot afford to let that happen." To monitor customer satisfaction, RENT systematically collects customer responses after each rental transaction. Specifically, after returning the rental vehicle to the return location, customers are contacted and asked to fill out detailed customer satisfaction surveys. Several incentives, such as bonuses, are also provided if the overall customer satisfaction level at a location is particularly higher than at other locations. It is important to note, however, that the bonuses are often provided at the team level, rather than at the employee level (unlike the case for upsell transactions where commissions are based on employee-level performance). The reason is that it is difficult to attribute higher customer satisfaction to a specific employee – i.e. customer satisfaction is a performance measure that is congruent (with the long-term success of RENT), but less sensitive to the employee's actions.

3.3.2. Proxy for Unexpected Career Concerns: Internal Merger Announcement

RENT acquired another rental car company as part of their firm strategy to expand increasing their US market share to almost 30 percent. The press highlighted this merger as potentially "the last combination of major U.S. car-rental companies that regulators will tolerate". Figure 3.1 provides a timeline of the major merger-related events and our corresponding sample time period. The CEO made an explicit company-wide merger announcement to all its employees which we treat as the trigger date *(t)* for employees to perceive the merger as definite. The possibility of a potential merger was first mentioned seven months prior to the merger announcement date by management, but only constituted of outside media sources. We treat *(t)* as the event date to be associated with the highest credibility for the likelihood of the merger as the announcement was made directly by management. In order to not confound our analyses with the effects from the first media mention, we start our sample period in *(t-6)*. Under the assumption that a media mention by external sources is also associated with heightened unexpected career concerns for employees, our results show that the internal announcement has an incremental effect as we

Figure 3.1. Sample Period and Timeline of Merger-related Events

This figure provides a graphical illustration of the timeline for the occurrence of all merger-related announcements. The table below provides the list of all SEC filings that were filed on date *(t)* (i.e. the merger announcement date).



SEC Filings on *(t)*

SEC Filings	Event
Form S-4	The amount of securities registered represents the maximum number of shares of the Company's common stock that can be issued in the exchange offer and second-step merger.
Form 8-K	Referencing a press release that announces the Company's proposal to acquire Target
Form 425	Referencing a press release that announces the Company's proposal to acquire Target
Form 425	Referencing the CEO's email sent to the Company's employees informing them about the merger
Form 425	Referencing the presentation slides used to inform investors about the merger

Form S-4 "Registration of securities, business combinations"
Form 8-K "Current report"
Form 425 "Prospectuses and communications, business combinations"

compare the performance of employees in the period prior to the internal announcement, but after the initial media mention to the period following the internal announcement.[15]

As illustrated in Figure 3.1, the actual merger was only completed 15 months after the RENT-initiated merger announcement date and involved a long-running bidding process including a second company-wide announcement by RENT to pull out of the merger at *(t+5)*. Considering the first management-initiated merger announcement and the long-running bidding process until the finalization of the merger, the actual deal completion date of the merger in *(t+15)* only constitutes outdated news for firm insiders. We compare the annual reports that were filed in the time period covering the entire bidding process, and confirm that the annual report issued subsequent to the merger announcement date *(t)* exhibits the highest frequency of mentions regarding the merger relative to the annual report issued immediately before the deal completion date *(t+15)*.[16] Moreover, Figure 3.1 also lists relevant merger-related SEC filings that were filed with RENT's merger announcement. These firm-related disclosures provide corroborating evidence that the announcement date *(t)* is associated with the most heightened expectations by firm insiders for the completion of the merger. In order to avoid confounding our analyses with the effects from the announcement to pull out of the merger, we end our sample period at *(t+5)*.

The merger event considered in this study is a friendly merger where no drastic changes in RENT's management was expected. The only effective change resulting from the merger was a change in ownership of the target firm. All operations of the target firm were maintained under its own brand name such that customers did not experience a *de facto* change in their service experience for RENT rental cars. Employees of the target firm experienced some changes to

[15] The availability of our data do not allow for a sample period to estimate the results using (t-7) as the event date.
[16] Frequent references to the merger include mentions in the introductory note, under potential risk factors and legal proceedings, and the management, discussion and analysis (MD&A) section.

assimilate operational procedures with those at RENT which included for example, the merging process of customer membership data. Most importantly for our study, employees employed at RENT did not experience any changes in their daily operations, and in their contractual employment relationship with the firm.

3.3.3. Proxy for Employee Alignment: Employment Engagement Survey

RENT implements employee surveys to gauge employee-level engagement with the strategic directions at the management-level. Specifically, it includes survey items that directly ask for employees about their satisfaction with regards to the implemented organizational changes.[17] All employees are asked to provide a score between one and five (where five constitutes the highest level of agreement). Using the results from the survey items that specifically ask employees about their agreement related to the implemented organizational change (hereafter, "Alignment Score"), we distinguish between store locations with employees that exhibit relatively higher or lower levels of alignment with the strategic directions of management. In particular, we partition the store locations based on the median Alignment Score.[18] The employee engagement surveys were conducted every 6 months, and Figure 3.1 maps the timeline of the surveys into our sample time period. We use the employee engagement survey results in the same month, but prior to the merger announcement date, and define *Alignment* as stores with above-median Alignment Scores.

Our research site provides us with an ideal setting to measure employee alignment using a survey instrument at different stores within the same organization. The reason is that the airport

[17] For the purpose of this study, we are interested in gauging employee alignment with the implemented changes at the organization. Therefore, we focus only on the survey items asking employees specifically about the change. An example of such a survey item would be "I understand the reasons for change." Due to confidentiality reasons associated with the identity of our research site, we are not able to disclose the full survey.
[18] The results remain unchanged regardless of whether we partition the store locations based on the mean of all relevant survey items, or each individual survey item.

store locations operate in isolated markets such that intra-firm spillover effects are non-existent. Employee survey data to proxy for implicit cultural aspects at individual store locations are problematic in organizations with frequent intra-store interactions among employees. The reason is that in such highly interactive settings, it is difficult to attribute the survey results to a particular individual store. The survey results in our setting are not subject to such caveats as employee interactions across stores are only minimal due to the geographical dispersion of the individual stores.

3.3.4. Employee Performance Measures

At RENT, employee performance is evaluated on two performance measures: sales-based upsell transactions, and customer satisfaction. The former (latter) is representative of a performance measure that is relatively more (less) sensitive, but less (more) congruent.

3.3.4.1. Upsell Transactions

An upsell transaction is defined as a transaction that generates more profit by soliciting the customer into an upgrade of his/her initial reservation. It is a transaction where the customer either upgrades his/her reserved car class or where the customer includes an add-on device such as a radio, GPS, and/or fuel. Therefore, an upsell transaction directly translates into higher revenues. Compared to customer satisfaction, upsell transactions are an indicator of employee performance over which employees have relatively more control (i.e. more sensitive). Yet, from management's perspective, upsell transactions are relatively more myopic in that they purely incent improving sales-based measures at the expense of customer satisfaction by potentially sacrificing customers' service experience. For instance, one frequent customer of RENT describes the tradeoffs by saying:

"the top performer (in upselling) unfortunately is usually not the friendly one but usually the jerk who tries to scare people into buying insurance or gas or stretching to truth to convince the customer to get an upgraded car." Therefore, a discontinuous increase in upselling performance may be good for the company's short-term financial performance, but potentially detrimental to its long-term performance if it damages the brand image around superior customer service. We define the variable *Upsell* as a dummy variable equal to one for such upsell transactions, and zero otherwise.

3.3.4.2. Customer Satisfaction

Customer satisfaction is measured using a survey instrument whereby customers are asked to fill out a customer satisfaction survey with each rental experience. To complete the survey, customers are provided a hyperlink in one of two ways after they return the car: through email or on their printed receipt. Whereas it constitutes a leading indicator for future financial performance, it is a measure that is relatively less sensitive to the employee's effort levels as they have relatively less control over customer perceptions than the actual sales numbers generated via upsell transactions. We define a variable *Overall Experience* which is the raw survey-based score for the survey item that asks customers about their overall rental experience. This raw survey-based score is used to create net promoter scores as a measure of customer satisfaction quality. It transforms the survey results into either one of three values, -100, 0, or 100 to represent "detractors", "neutrals" and "promoters", respectively. This measure translates RENT's customer satisfaction surveys into comparable results with that of other competing rental car companies, and is, thus, primarily used by RENT in evaluating customer satisfaction quality for management control purposes. Accordingly, we base our main measure for customer satisfaction on the net promoter score, and construct a dummy variable *Will Recommend* that is equal to one if the net promoter score is in the most satisfactory "promoter" category, and zero otherwise.

3.4. EMPIRICAL RESEARCH DESIGN

First, to examine whether unexpected career concerns are associated with significant incentive effects (H1), we test whether performance significantly improves subsequent to the merger announcement. Exploiting the monthly panel data structure of 111,078 transaction records with survey responses from 81 locations over a 12-month time period, we estimate the following regression model:

$$Y_{iejt} = Location_i + Employee_e + \beta_1 \cdot Merger\ Announcement_t + X_j'\gamma + \epsilon_{ijt} \quad (1)$$

The dependent variable, Y_{iejt}, is either one of the three employee performance measures: *Upsell*, *Overall Experience*, or *Will Recommend*. *Merger Announcement*$_t$ is a dummy variable that equals to one for the months following the merger announcement, and zero otherwise. *Location*$_i$ are location-fixed effects, and *Employee*$_e$ are employee-fixed effects. X_j constitute various transaction-level and survey-level covariates. These variables include *Duration* which measures the length of the car rental in days; *Weekend* which is a dummy variable that equals one if the car rental was made on a weekend, and zero otherwise; *Membership* which is a dummy variable that equals one if the car rental was made by a customer who has a membership with RENT, and zero if not; and *Business* which is a dummy variable that equals one if the car rental was made for business purposes, and zero otherwise. We also include several categorical variables that control for car types and booking channels. We do not include time-period fixed effects as they would subsume our coefficient of interest β_1 on *Merger Announcement*$_t$ which captures the incremental performance following the merger announcement. In other words, we conduct pre-announcement and post-announcement comparisons around the merger announcement to evaluate how unexpected career concerns affect employee behaviors.

Such pre-post comparisons may be subject to several empirical threats. While imperfect, in this study we conduct several robustness tests to provide more corroborating evidence on the effects of unexpected career concerns. First, to minimize any temporal trends, we use relatively short time windows surrounding the event date. Second, we distinguish between store locations where employees are more or less likely to be subject to the merger-related career concerns. We define a variable *Less Risk* that indicates airport store locations at which RENT employees presumably face less layoff risks due to the merger announcement which are airports at which the target rental car company has no operating store location. *Less Risk* is defined as one for such locations, and zero otherwise. If career concerns from a potential consolidation is the main driver to the changes in employee behaviors, we expect that the effect should be much smaller in locations without target rental car company stores. To examine whether the merger announcement-related performance effects are driven by store locations where the unexpected career concerns are more predominant, we estimate the following equation:

$$Y_{iejt} = Merger\ Announcement_t + \ Location_i + \ Employee_e$$
$$+\beta_2 \cdot Merger\ Announcement_t \times Less\ Risk_i + X_j'\gamma + \ \epsilon_{ijt} \qquad (2)$$

After establishing that unexpected career concerns affect employee performance, we then move to H2 and examine whether effort allocation tendencies differ depending on the extent of employee alignment. We employ a difference-in-differences research design, and compare locations with different levels of employee alignment *prior to* the merger announcement. We use the degree of employee alignment *prior* to the announcement to avoid reverse causality, because the actual merger announcement may affect both the level of management-employee alignment and employee behavior simultaneously. Another advantage of employing a difference-in-differences research design is that we can avoid measuring a potential spurious relationship that may arise due to the organizational change and the associated management practice over the relevant time period. For instance, the expectation of having to consolidate the customer memberships at RENT and the target firm prior to the merger may affect employee behaviors during our sample period. A simple temporal comparison cannot distinguish the effect of the merger announcement on employee behaviors from the effect due to such changes in associated management practices. However, by including location- and time-fixed effects together, we can rule out such possibility. We estimate the difference-in-differences specifications using the following model:

$$Y_{iejt} = Location_i + Employee_e + Year\text{-}Month_t$$
$$+\beta_3 \cdot Merger\ Announcement_t \times Alignment_i + X_j'\gamma + \epsilon_{ijt} \quad (3)$$

The dependent variable, Y_{iejt}, is the same as in the above regressions. The location fixed effects, *Location$_i$*, control for time-invariant, location specific characteristics, the time-period fixed effects, *Year-Month$_t$*, control for any time-specific effects affecting all locations equally during the sample period, and the employee-fixed effects, *Employee$_e$*, control for any employee-specific characteristics. *Merger Announcement$_t$* is a dummy variable that equals to one for the months following the merger announcement, and zero otherwise. *Alignment$_i$* is a dummy variable that equals to one for airport store locations with above-median Alignment Scores, and zero for airport store locations with below-median Alignment Scores. The transaction-level and survey-level covariates X_j are the same as above. The coefficient of interest is β_3 which is the coefficient on the interaction term *Merger Announcement*$_t$ × *Alignment*$_i$. It captures the differential response of the employees in high- and low-*Alignment* locations to the merger announcement.

3.5. DISCUSSION OF EMPIRICAL RESULTS

3.5.1. Descriptive Statistics

Table 3.1 provides overall summary statistics of our data across all airport store locations. As shown in Panel A, there are a total of 81 major US airport store locations in our sample. The mean Alignment Score across all stores is 3.89. A histogram that graphically illustrates the distribution of stores on the Alignment Score is provided in Figure 3.2.

Our performance data are at the transaction-level for car rentals at all 81 major US airport store locations over the relevant sample time period, and constitute a total of 111,078 rental car transactions. Panel B of Table 1 summarizes the rental car transaction-related characteristics. From the variable *Duration*, we observe that the average duration between pick-up and return of the rented vehicle constitutes about 4 days. The variable *Weekend* indicates that about 22% of all transactions are made on a weekend, the variable *Membership* shows that about 74% of all transactions were made by RENT membership holders, and the variable *Business* shows that about 35% of all transactions were indicated to have been for business purposes. About 40% of all transactions constitute upsell transactions.

In Panel C of Table 3.1, we provide summary statistics on the customer satisfaction-related variables. The mean score on *Overall Experience* is about 7, and about 62% of all transactions are categorized into the most satisfactory "promoter" category. Panel C also provides the summary statistics for the raw scores for each of the subcomponents in the customer satisfaction survey.

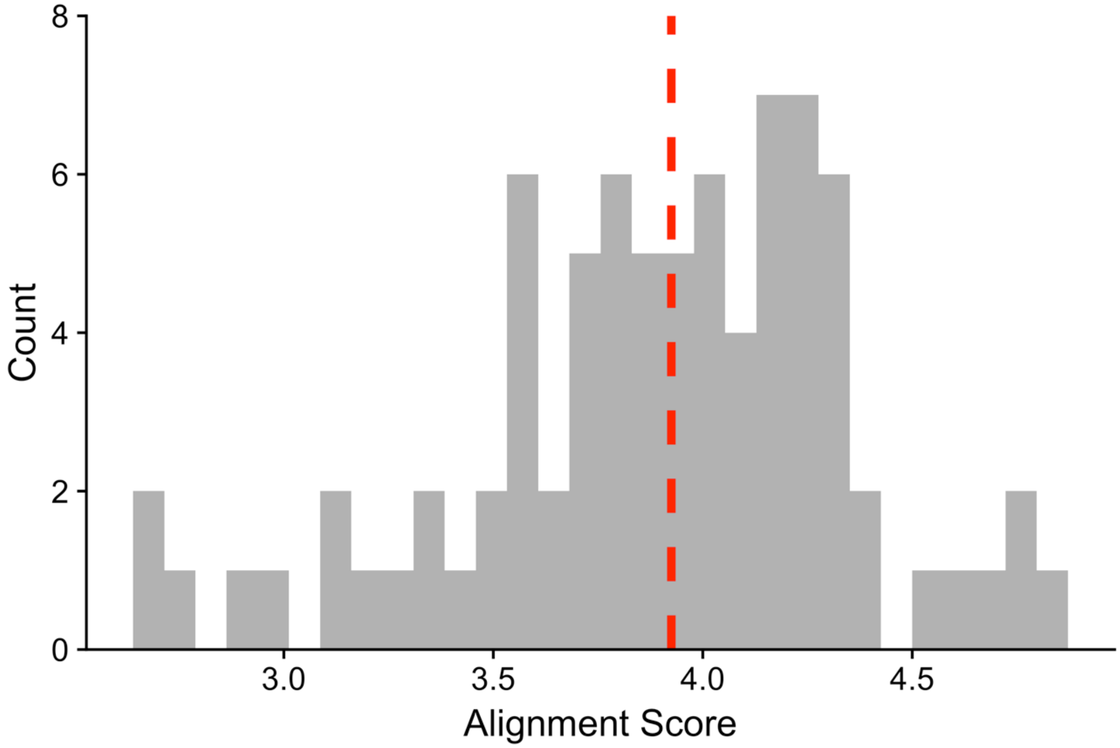
Table 3.1. Descriptive Statistics

This table provides the summary statistics of variables used in this study across all locations, or rental car transactions. Panel A reports the statistics on the location-specific variable *Alignment Score*. It refers to the average score on all change-related survey items in the employee engagement survey. Panel B and C report summary statistics on the transaction-level characteristics. *Duration* measures the length of the car rental in days; *Weekend* is a dummy variable that equals one if the car rental was made on a weekend, and zero otherwise; *Membership* is a dummy variable that equals one if the car rental was made by a customer who has a membership with RENT, and zero if not; and *Business* is a dummy variable that equals one if the car rental was made for business purposes, and zero otherwise. *Upsell* is a dummy variable that equals one if the rental car transaction qualifies as an Upsell transaction. Upsell transactions generate more profit by soliciting the customer into an upgrade of his/her initial rental car reservation. *Overall Experience* is the raw score from the customer satisfaction survey on the item "Overall Experience" and can range from zero to nine. *Net Promoter Score* is the raw score that the company uses to evaluate customer satisfaction, and is either -100, 0, or 100. *Will Recommend* is a dummy variable that equals to one if the raw *Net Promoter Score* is 100, zero otherwise. *Staff Courtesy, Speed of Service, Vehicle Condition, Billing as Expected,* and *Value for Money* are raw scores on each of the corresponding customer satisfaction survey items.

Statistic	N	Mean	Std. Dev.	Min	Max
Panel A: Location					
Alignment Score	81	3.890	0.467	2.667	4.824
Panel B: Transaction					
Duration	111,078	4.019	4.376	0	91
Weekend	111,078	0.220	0.414	0	1
Membership	111,078	0.743	0.437	0	1
Business	111,078	0.345	0.475	0	1
Upsell	111,078	0.409	0.492	0	1
Panel C: Customer Satisfaction Survey					
Overall Experience	111,078	6.976	2.548	0	9
Net Promoter Score	111,078	43.344	79.371	-100	100
Will Recommend	111,078	0.626	0.484	0	1
Staff Courtesy	99,689	7.826	1.906	0	9
Speed of Service	99,688	7.305	2.466	0	9
Vehicle Condition	99,687	7.262	2.420	0	9
Billing as Expected	99,677	7.609	2.376	0	9
Value for Money	99,675	6.911	2.334	0	9

Figure 3.2. Histogram of Alignment Score

This figure provides the histogram for all change-related survey items (i.e. Alignment Score) in the employee engagement survey across all stores in our sample. The X-axis represents the Alignment Score which can range from one to five. The Y-axis represents the number of stores at each score bracket on the X-axis.



These include “staff courtesy”, “speed of service”, “vehicle condition”, “billing as expected”, and “value for money”.

Table 3.2 provides the summary statistics for our main variables of interest separately for high-*Alignment* and low-*Alignment* store locations to ensure that there are no significant fundamental differences between these two location types that may impact our empirical results for our second hypothesis. The summary statistics provide confidence that both store location types are comparable in terms of their underlying operation characteristics.

3.5.2. Incentive Effects

Table 3.3 presents our results on how employee performance is affected subsequent to the merger announcement from estimating equation (1). The dependent variables are *Upsell*, *Overall Experience*, and *Will Recommend* in columns 1 through 3, columns 4 through 6, and columns 7 through 9, respectively. The first two columns pertaining to each dependent variable vary in terms of the inclusion of fixed-effects. All specifications are estimated using OLS. Therefore, we interpret the estimated coefficients as marginal effects (Angrist and Pischke 2009; Ai and Norton 2003). Interpreting the second column with the inclusion of location- and employee- fixed effects, we observe that the likelihood of an upsell increases by 8.9 percent in the post-merger announcement period. Moreover, as shown in column 8, the likelihood for a rental transaction to be classified into the highest customer satisfaction category increases by 2.7 percent subsequent to the merger announcement. In columns 3, 6, and 9, we estimate the same regression model on a smaller time period window of six months which compares the three months prior to the three months after the merger announcement. Finding the effect over a smaller time period window allows us to attribute the effect more confidently to the merger announcement event as a large time

Table 3.2. Comparisons between High and Low Alignment Locations

This table compares High- and Low- Alignment locations, and provides key summary statistics separately across these two types of locations. High (Low)-Alignment locations are locations that scored above (below)-median on the change-related survey items. *Alignment Score* refers to the average score on all change-related survey items in the employee engagement survey. *# Transactions* reports the average number of total monthly rental car transactions. *Upsell* is a dummy variable that equals one if the rental car transaction qualifies as an Upsell transaction. Upsell transactions generate more profit by soliciting the customer into an upgrade of his/her initial rental car reservation. *Overall Experience* is the raw score from the customer satisfaction survey on the item "Overall Experience" and can range from 0 to 9. *Will Recommend* is a dummy variable that equals to 1 if the raw *Net Promoter Score* is 100, 0 otherwise.

Statistic	N	Mean	St. Dev.	Min	Max
Panel A: High-Alignment Locations					
Alignment Score	40	4.255	0.223	3.943	4.824
#Transactions	40	100.820	84.620	22.427	354.006
Upsell	40	0.356	0.083	0.056	0.521
Overall Experience	40	6.984	0.265	6.427	7.605
Will Recommend	40	0.625	0.041	0.530	0.728
Panel B: Low-Alignment Locations					
Alignment Score	41	3.534	0.353	2.667	3.925
#Transactions	41	131.736	99.590	15.384	421.575
Upsell	41	0.351	0.083	0.117	0.467
Overall Experience	41	6.938	0.284	6.153	7.565
Will Recommend	41	0.620	0.055	0.442	0.737

Table 3.3. Effect of Merger Announcement on Employee Performance

This table provides the results from estimating the following model:

$$Y_{iejt} = \beta 1 \cdot Merger\ Announcement_t + \ Location_i + Employee_e + X_j^{'}\gamma + \ \epsilon_{iejt}$$

In columns 1 through 3, the dependent variable is the sales-based performance measure *Upsell*. *Upsell* is a dummy variable that equals one if the rental car transaction qualifies as an Upsell transaction. In columns 4 through 6, the dependent variable is the customer satisfaction measure *Overall Experience*. In columns 7 through 9, the dependent variable is the customer satisfaction measure *Will Recommend*. *Merger Announcement* is a dummy variable that is equal to 1 for the months following the merger announcement, and 0 otherwise. *Location*$_i$ are location-fixed effects, *Employee*$_e$ are employee-fixed effects. *X* is a vector of transaction-related characteristics, and includes *Duration* which measures the length of the car rental in days; *Weekend* which is a dummy variable that equals one if the car rental was made on a weekend, and zero otherwise; *Membership* which is a dummy variable that equals one if the car rental was made by a customer who has a membership with RENT, and zero if not; and *Business* which is a dummy variable that equals one if the car rental was made for business purposes, and zero otherwise. Additional transaction controls include indicator variables for car types and booking channels. Columns 1 and 2, 4 and 5, 7 and 8 differ in the inclusion of fixed effects. Columns 3, 6, and 9 estimate the model specification over a 6-month time period window. Standard errors are clustered by locations, and presented in parentheses. All specifications are estimated by OLS. *, **, and *** denote statistical significance at 10%, 5%, and 1% level, respectively.

	Upsell (=1)			Overall Experience			Will Recommend (=1)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Merger Announcement	0.088***	0.089***	0.066***	0.139***	0.157***	0.055*	0.024***	0.027***	0.009*
	(0.006)	(0.006)	(0.008)	(0.033)	(0.030)	(0.028)	(0.006)	(0.005)	(0.005)
Duration	-0.002***	-0.002***	-0.002***	0.017***	0.019***	0.017***	0.002***	0.002***	0.002***
	(0.0004)	(0.0005)	(0.001)	(0.002)	(0.002)	(0.003)	(0.0004)	(0.0004)	(0.0005)
Weekend	-0.009**	-0.009**	-0.005	-0.004	0.040*	0.055*	0.0001	0.007*	0.015***
	(0.004)	(0.004)	(0.006)	(0.022)	(0.024)	(0.029)	(0.003)	(0.004)	(0.005)
Business	-0.017***	-0.018***	-0.016***	-0.029	-0.058***	-0.040	-0.006	-0.011***	-0.008
	(0.004)	(0.004)	(0.005)	(0.021)	(0.022)	(0.030)	(0.004)	(0.004)	(0.006)
Membership	-0.045***	-0.041***	-0.043***	0.448***	0.269***	0.226***	0.091***	0.057***	0.050***
	(0.007)	(0.006)	(0.008)	(0.040)	(0.037)	(0.050)	(0.007)	(0.006)	(0.009)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Employee FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Sample Period (in Months)	12	12	6	12	12	6	12	12	6
Observations	111,078	108,558	56,233	101,973	99,713	51,624	101,973	99,713	51,624
Adjusted R^2	0.146	0.154	0.163	0.018	0.032	0.031	0.017	0.027	0.025

period window is more likely to be convoluted by other events. The results provide corroborating evidence that the merger announcement is associated with the resulting performance improvements.

Table 3.4 provides the results from estimating equation (2), and provides another corroborating evidence that unexpected career concerns drive the performance improvements subsequent to the merger announcement. We estimate the model specification including all fixed effects, over a 12-months and six-months window in the first and second columns pertaining to each outcome variable, respectively. Airport store locations that are relatively less subject to the merger-related layoff risks exhibit upsell performance declines as evidenced by the negative coefficient on the interaction between *Merger Announcement* and *Less Risk*. In other words, the documented performance improvements are primarily driven by store locations at which employees are more subject to unexpected career concerns. Such performance declines at *Less Risk* store locations are less evident when considering the customer satisfaction measures. The coefficient on the interaction term is only negative over the six-months window, and only significantly negative when considering the *Overall Experience* measure. The finding that upsell performance is more responsive to the merger announcement is consistent with it being a relatively more sensitive measure that the employee is more likely able to influence.

3.5.3. Effort Allocation Effects and Employee Alignment

So far our results show that performance increases on both measures – the sales-based measure *Upsell* and customer satisfaction – subsequent to the merger announcement. In this section, we examine whether there is variation in how employees allocate their effort in improving

Table 3.4. Unexpected Career Concerns as Driver

This table provides the results from estimating the following model:

$$Y_{iejt} = Merger\ Announcement_t + Location_i + Employee_e + \beta 2 \cdot Merger\ Announcement_t * Less\ Risk + X_j^{'}\gamma + \epsilon_{iejt}$$

In columns 1 and 2, the dependent variable is the sales-based performance measure *Upsell*. *Upsell* is a dummy variable that equals one if the rental car transaction qualifies as an Upsell transaction. In columns 3 and 4, the dependent variable is the customer satisfaction measure *Overall Experience*. In columns 5 and 6, the dependent variable is the customer satisfaction measure *Will Recommend*. *Merger Announcement* is a dummy variable that is equal to one for the months following the merger announcement, and zero otherwise. $Location_i$ are location-fixed effects, $Employee_e$ are employee-fixed effects. *Less Risk* indicates airport store locations at which the target rental car company has no operating store location. *X* is a vector of transaction-related characteristics, and includes *Duration* which measures the length of the car rental in days; *Weekend* which is a dummy variable that equals one if the car rental was made on a weekend, and zero otherwise; *Membership* which is a dummy variable that equals one if the car rental was made by a customer who has a membership with RENT, and zero if not; and *Business* which is a dummy variable that equals one if the car rental was made for business purposes, and zero otherwise. Additional transaction controls include indicator variables for car types and booking channels. Columns 1, 3, and 5 (columns 2, 4, and 6) estimate the model specification over a 12-month time period window (6-month time period window). Standard errors are clustered by locations, and presented in parentheses. All specifications are estimated by OLS. *, **, and *** denote statistical significance at 10%, 5%, and 1% level, respectively.

	Upsell (=1)		Overall Experience		Will Recommend (=1)	
	(1)	(2)	(3)	(4)	(5)	(6)
Merger Announcement	0.091***	0.067***	0.159***	0.067***	0.026***	0.009**
	(0.004)	(0.005)	(0.019)	(0.023)	(0.004)	(0.005)
Merger Announcement x	-0.099***	-0.113***	0.022	-0.363**	0.029	-0.009
Less Risk	(0.021)	(0.030)	(0.142)	(0.156)	(0.023)	(0.028)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes	Yes	Yes
Employee FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Period (in Month)	12	6	12	6	12	6
Observations	108,558	56,233	99,713	51,624	99,713	51,624
Adjusted R^2	0.154	0.163	0.048	0.041	0.028	0.025

the sales-based measure relative to customer satisfaction. Specifically, we test if alignment induces employees to exert greater effort on the more congruent performance measure (i.e. customer satisfaction) relative to the more sensitive performance measure (i.e. *Upsell*). In other words, we examine whether employee alignment is a significant moderating factor that influences upsell performance relative to customer satisfaction subsequent to the merger announcement. The results are tabulated in Table 3.5. Panel A examines *Upsell* as the dependent variable, and Panel B examines the customer satisfaction-related performance measures as the dependent variable. Column 1 presents the baseline descriptive results without any fixed effects. Omitting the fixed effects allows for the estimation of the coefficients on *Merger Announcement* and *Alignment*. The estimated constant term, 0.367, indicates that before the merger announcement, the likelihood of an *Upsell* transaction is about 36.7 percent. The likelihood of an upsell increases about 10.7 percent after the merger announcement ($p < 0.01$). This finding is consistent with unexpected career concerns being associated positive incentive effects.

More importantly for the purpose of examining the moderating effect of employee alignment, the likelihood of an upsell starts to diverge significantly between high- and low-*Alignment* locations after the merger announcement. There are no significant differences between high- and low-*Alignment* locations before the merger announcement as shown by the insignificant coefficient on *Alignment*. However, following the merger announcement, high-*Alignment* locations are significantly less likely to engage in an upsell transaction than low-*Alignment* locations ($p < 0.01$) as evidenced by the negative coefficient on the interaction term, *Merger Announcement × Alignment*. In other words, employees at high-*Alignment* locations are associated with an increase in the likelihood of an *Upsell* by 20 percent (= (0.107 - 0.035) / 0.367) whereas employees at low-*Alignment* locations are associated with an increase in the likelihood of an Upsell

Table 3.5. Moderating Effects of Employee Alignment

This table provides the results from estimating following model over the entire sample period:

$$Y_{iejt} = Location_i + Employee_e + Year\text{-}Month_t$$
$$+\beta 3 \cdot Merger\ Announcement_t \times Alignment_i + X_j^{'}\gamma + \epsilon_{iejt}$$

In Panel A, the dependent variable is the sales-based performance measure *Upsell*. *Upsell* is a dummy variable that equals 1 if the rental car transaction qualifies as an Upsell transaction. Upsell transactions generate more profit by soliciting the customer into an upgrade of his/her initial rental car reservation. In Panel B, the dependent variable are the customer satisfaction-based performance measures, *Overall Experience* in columns 1 through 4 and *Will Recommend* in columns 5 through 8. *Location*$_i$ are location-fixed effects, Employee$_e$ are employee-fixed effects, and *Year-Month*$_t$ are time-period-fixed effects. *Merger Announcement* is a dummy variable that is equal to one for the months following the merger announcement, and zero otherwise. *Alignment* is a dummy equal to 1 for locations with above-median Alignment scores, and 0 for below-median Alignment scores. *X* is a vector of transaction-related characteristics, and includes *Duration* which measures the length of the car rental in days; *Weekend* which is a dummy variable that equals one if the car rental was made on a weekend, and zero otherwise; *Membership* which is a dummy variable that equals one if the car rental was made by a customer who has a membership with RENT, and zero if not; and *Business* which is a dummy variable that equals one if the car rental was made for business purposes, and zero otherwise. Additional transaction controls include car types and booking channels. Standard errors are clustered by locations, and presented in parentheses. All specifications are estimated by OLS. *, **, and *** denote statistical significance at 10%, 5%, and 1% level, respectively.

Panel A: Upsell

	Upsell (=1)			
	(1)	(2)	(3)	(4)
Constant	0.367***			
	(0.003)			
Merger Announcement	0.107***			
	(0.004)			
Alignment	0.005			
	(0.004)			
Merger Announcement x Alignment	-0.035***	-0.030***	-0.027**	-0.029**
	(0.006)	(0.011)	(0.011)	(0.012)
Rent Duration			-0.002***	-0.002***
			(0.0004)	(0.0005)
Weekend			-0.009**	-0.008**
			(0.004)	(0.004)
Business			-0.019***	-0.020***
			(0.004)	(0.004)
Membership			-0.045***	-0.041***
			(0.007)	(0.006)
Other Controls	No	No	Yes	Yes
Location FE	No	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes
Employee FE	No	No	No	Yes
Observations	111,078	111,078	111,078	108,558
Adjusted R^2	0.009	0.034	0.149	0.157

Panel B: Customer Satisfaction

	Overall Experience				Will Recommend (=1)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	6.965***				0.620***			
	(0.014)				(0.003)			
Merger Announcement	0.083***				0.013***			
	(0.021)				(0.004)			
Alignment	0.007				0.005			
	(0.021)				(0.004)			
Merger Announcement x Alignment	0.107***	0.099*	0.109*	0.095*	0.022***	0.021**	0.021**	0.018*
	(0.032)	(0.059)	(0.061)	(0.057)	(0.006)	(0.010)	(0.010)	(0.010)
Other Controls	No	No	Yes	Yes	No	No	Yes	Yes
Location FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Employee FE	No	No	No	Yes	No	No	No	Yes
Observations	101,935	101,935	101,935	99,675	101,935	101,935	101,935	99,675
Adjusted R^2	0.001	0.007	0.036	0.050	0.001	0.005	0.019	0.029

by 29 percent (= 0.107 / 0.367) after the merger announcement. The results from estimating the difference-in-differences specification including location-, and year-month fixed effects, are reported in column 2. The location fixed effects allows to control for unobservable factors that are correlated with some seasonality that may affect the likelihood of an upsell, and thus investigate the likelihood of an upsell within each location over time. The year-month fixed effects account for common time trends that affect all locations. Our new estimate on the interaction term confirms that high-*Alignment* locations are less likely to engage in an upsell transaction than low-*Alignment* locations in the post-period. We also report the results from the difference-in-differences specification that includes the additional transaction-level control variables (in column 3), and additional employee-fixed effects (in column 4). The direction and economic magnitude of the coefficient on *Merger Announcement* $\times$ *Alignment* are comparable across all columns.

Panel B of Table 3.5 presents our results on how the merger announcement affects employee performance on the customer satisfaction-based measures. The dependent variable is *Overall Experience* in columns 1 through 4, and *Will Recommend* in columns 5 through 8, respectively. Columns 1 and 5 present the baseline descriptive results without any fixed effects. Following the merger announcement, we observe a significant 0.08 increase in the rating for *Overall Experience* (column 1), and a 1.3 percent higher likelihood for a rental transaction to be classified into the highest customer satisfaction category (column 5). Similar as in Panel A, we estimate a difference-in-differences specification to examine how customer satisfaction in high- and low-*Alignment* locations is differently affected subsequent to the merger announcement. We include location-, and year-month fixed effects (in columns 2 and 6), additional transaction-level control variables (in columns 3 and 7), and additional employee-fixed effects (in columns 4 and 8). Again, there are no significant differences between high- and low-*Alignment* locations before the merger announcement as shown by the insignificant coefficient on *Alignment*. However, contrary to the results for *Upsell*, we observe that the coefficient on *Merger Announcement × Alignment* is significantly positive across all columns which suggests that high-*Alignment* locations are significantly more likely to improve customer satisfaction than low-*Alignment* locations. The economic magnitude of the coefficients are comparable across all model specifications. For example, interpreting the results in column 4 we observe that high-*Alignment* locations experience a 2.7 percent (= (0.083 + 0.107) / 6.965) increase after the merger announcement, whereas low-*Alignment* locations experience only a 1.2 percent increase (= 0.083 / 6.965).

In additional tests tabulated in Table 3.6, we estimate column 4 of Table 3.5 Panel B using the raw scores of the different subcomponents in the customer satisfaction survey. Columns 1 through 5 use scores on the subcomponents "value for money", "speed of service", "billing as expected", "staff courtesy", and "vehicle condition" as the dependent variable, respectively. Whereas employees have control over dimensions such as "speed of service" or "staff courtesy" by taking less breaks, contemplating innovative ways to make the rental car process more time-efficient, and/or being more polite to customers; subcomponents such as "vehicle condition" and "billing as expected" are categories over which the employees at each airport store location do not have control over as they are managed centrally. The results show that the effect is primarily driven by the subcomponent "speed of service" which provides corroborating evidence that the observed improvements in customer satisfaction at high-*Alignment* locations are due to greater effort exertion by employees to improve customer satisfaction subsequent to the merger announcement.

In Table 3.7 we test for the lack of pre-trends (often described as the parallel trends assumption) by estimating the dynamic version of the difference-in-differences specification. To

Table 3.6. Customer Satisfaction Subcomponents

This table estimates Table 3.5 column 4 replacing the dependent variable with the raw scores of the customer satisfaction subcomponents. The subcomponents in columns 1 through 5 are “value for money”, “speed of service”, “billing as expected”, “staff courtesy”, and “vehicle condition”, respectively. Standard errors are clustered by locations, and presented in parentheses. All specifications are estimated by OLS. *, **, and *** denote statistical significance at 10%, 5%, and 1% level, respectively.

	Value for Money (1)	Speed of Service (2)	Billing as Expected (3)	Staff Courtesy (4)	Vehicle Condition (5)
Merger Announcement x Alignment	0.074	0.198**	0.061	0.047	0.021
	(0.046)	(0.086)	(0.046)	(0.035)	(0.050)
Other Controls	Yes	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Employee FE	Yes	Yes	Yes	Yes	Yes
Observations	89,626	89,637	89,628	89,638	89,636
Adjusted R^2	0.013	0.026	0.028	0.017	0.019

Table 3.7. Dynamic Moderating Effects of Employee Alignment

This table provides the estimates of the dynamic version of the following model over the entire sample period:

$$Y_{iejt} = Location_i + Employee_e + Year\text{-}Month_t + \Sigma\beta \cdot Year\ Month_t \times Alignment_i + X_j'\gamma + \epsilon_{iejt}$$

The month immediate prior to the merger announcement is treated as the baseline year-month. The dependent variables are *Upsell* which is a dummy variable that equals 1 if the rental car transaction qualifies as an Upsell transaction (column 1), *Overall Experience* which is the raw score on the corresponding item in the customer satisfaction survey (in column 2), and *Will Recommend* which is a dummy variable that equals to 1 if the raw *Net Promoter Score* is 100, 0 otherwise (in column 3). $Location_i$ are location-fixed effects, $Employee_e$ are employee-fixed effects, and $Year\text{-}Month_t$ are time-period-fixed effects, *Alignment* is a dummy equal to 1 for locations with above-median Alignment scores, and 0 for below-median Alignment scores. *X* is a vector of transaction-related characteristics, and is the same as before. Standard errors are clustered by locations, and presented in parentheses. All specifications are estimated using OLS. *, **, and *** denote statistical significance at 10%, 5%, and 1% level, respectively.

	Upsell (=1) (1)	Overall Experience (2)	Will Recommend (=1) (3)
Merger Announcement (t-7) x Alignment	0.005	-0.035	-0.002
	(0.018)	(0.115)	(0.019)
Merger Announcement (t-6) x Alignment	0.003	-0.091	-0.038**
	(0.018)	(0.087)	(0.017)
Merger Announcement (t-5) x Alignment	-0.001	0.052	-0.003
	(0.015)	(0.094)	(0.015)
Merger Announcement (t-4) x Alignment	-0.004	-0.015	-0.00001
	(0.017)	(0.084)	(0.016)
Merger Announcement (t-3) x Alignment	0.003	0.082	-0.007
	(0.019)	(0.126)	(0.021)
Merger Announcement (t-2) x Alignment	-0.009	-0.044	-0.006
	(0.015)	(0.112)	(0.023)
Merger Announcement (t) x Alignment	-0.031	-0.077	-0.014
	(0.019)	(0.119)	(0.020)
Merger Announcement (t+1) x Alignment	-0.046**	0.078	0.001
	(0.020)	(0.114)	(0.020)
Merger Announcement (t+2) x Alignment	-0.047**	0.017	0.009
	(0.020)	(0.105)	(0.019)
Merger Announcement (t+3) x Alignment	0.002	0.233**	0.036**
	(0.020)	(0.113)	(0.017)
Merger Announcement (t+4) x Alignment	-0.019	0.246**	0.030*
	(0.019)	(0.119)	(0.018)
Other Controls	Yes	Yes	Yes
Location FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Employee FE	Yes	Yes	Yes
Observations	108,518	99,675	99,675
Adjusted R^2	0.149	0.050	0.029

do so we interact *Alignment* with each of the individual time period dummies, and omit the month immediately prior to the merger announcement to test whether the high-*Alignment* locations and low-*Alignment* locations exhibit similar trends until the merger announcement. The dependent variable is *Upsell*, *Overall Experience*, and *Will Recommend* in column 1, 2, and 3, respectively. Figure 3.3 reproduces the same results as in Table 3.7 by plotting the estimated coefficients for each interaction over the entire sample period. Panel A plots the results for *Upsell*, and Panel B plots the results for *Will Recommend*. Consistent with the parallel trend assumption, we observe that the stores exhibit no significant performance differences in the months prior to the merger announcement, and that the effect is apparent only after the merger announcement.

To summarize, the empirical results provided in this section show that the merger announcement is associated with different effort allocation effects depending on the extent by which employees are aligned with the overall organizational objectives. Following the merger announcement, high-*Alignment* locations exhibit more improvements related to the customer satisfaction-based performance measures, but less improvement on the *Upsell* measure relative to low-*Alignment* locations. Collectively, these findings provide support for H2 that employee alignment can mitigate myopic employee incentives to fixate on short-term at the expense on long-term performance arising from incentives to minimize layoff risk when subject to unexpected career concerns.

3.6. CONCLUSION

This study examines incentive and effort allocation effects due to unexpected career concerns. We use data from a rental car company, and exploit an internal announcement by management regarding its intent for a horizontal merger as a source of heightened unexpected

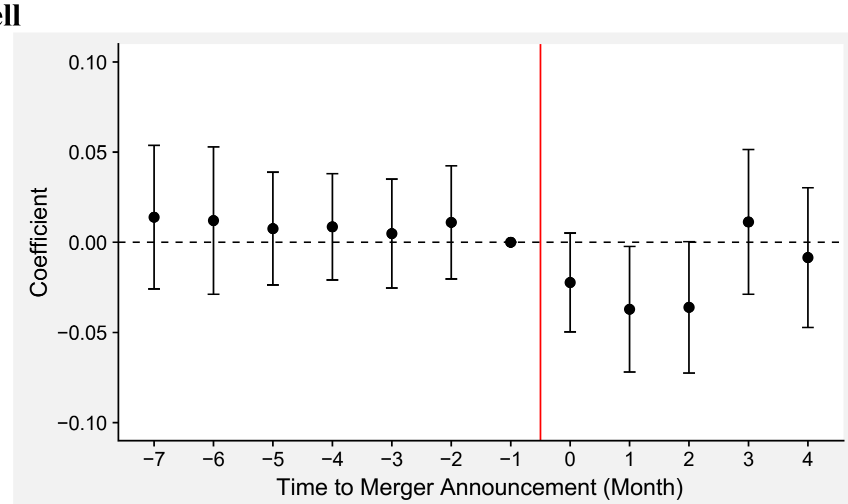
Figure 3.3. Dynamic Effects of Merger Announcement

This figure provides a graphical illustration of the results from estimating the dynamic version of the following model over the entire sample period:

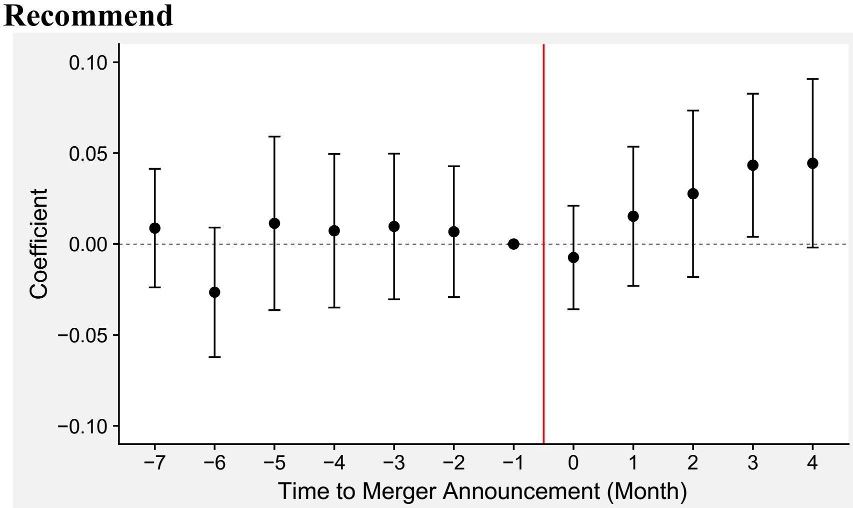
$$Y_{iejt} = Location_i + Employee_e + \ Year\text{-}Month_t + \Sigma\beta \cdot Year\ Month_t \times Alignment_i + X_j'\gamma + \ \epsilon_{iejt}$$

The dependent variable is either the sales-based performance measure *Upsell* (i.e. a dummy equal to one if a rental car transaction qualifies as an Upsell transaction, and zero otherwise) or the customer satisfaction-based performance measure *Will Recommend* (i.e. a dummy equal to one if the net promoter score is in the most satisfactory category, and zero otherwise). $Location_i$ are location-fixed effects, $Employee_e$ are employee-fixed effects, $Year\text{-}Month_t$ are time-period-fixed effects, *Alignment* is a dummy equal to one for locations with above-median Alignment Scores. *X* is a vector of transaction-related characteristics, and includes *Duration* which measures the length of the car rental in days; *Weekend* which is a dummy variable that equals one if the car rental was made on a weekend, and zero otherwise; *Membership* which is a dummy variable that equals one if the car rental was made by a customer who has a membership with RENT, and zero if not; and *Business* which is a dummy variable that equals one if the car rental was made for business purposes, and zero otherwise. Specifically, this figure plots the estimated coefficients on the interaction terms between Alignment and each time-period-fixed effect, $Year\text{-}Month_t$. The month immediate prior to the merger announcement is the baseline year-month. The relevant dependent variable is *Upsell* in Panel A, and *Will Recommend* in Panel B, respectively. All estimated coefficients are reported in Table 3.7. The black dots correspond to coefficient estimates on the interaction between *Alignment* and each of the time dummy variables as tabulated in Table 3.7. The corresponding 95% confidence intervals around the estimates are plotted as error bars.

Panel A: Upsell



Panel B: Will Recommend



career concerns. We examine employee performance subsequent to the merger decision, and examine whether such performance effects vary depending on the degree of employee alignment. Employee performance is evaluated based on two measures: (1) a sales-based measure that is relatively more reflective of immediate short-term financial performance, and (2) customer satisfaction which is relatively more reflective of intangible long-term performance. All existing formal management control channels remain constant over our sample period including the employee incentive contracts.

First, we document positive incentive effects associated with unexpected career concerns. Performance measures on upselling and customer satisfaction exhibit significant improvement subsequent to the merger announcement which suggests that unexpected career concerns trigger employee performance incentives to minimize potential layoff risks. Second, we document effort allocation effects associated with such unexpected career concerns. We find that greater employee alignment is associated with relatively greater levels of improvement in customer satisfaction, but not so in the sales-based performance measure. In other words, more aligned employees exhibit more willingness to trade-off improvements in customer satisfaction at the expense of immediate result in their sales-based performance measures. Taken together, these findings suggest that employee alignment can mitigate career concern-related myopic employee behaviors to fixate effort levels on relatively short term-oriented performance measures at the expense of decreasing effort levels towards more goal-congruent performance measures.

Our study contributes to a comprehensive understanding of performance incentives embedded in unexpected career concerns within organizations, and sheds new light on the benefits of employee alignment under periods of organizational change. Despite rapidly changing business environments, the management accounting literature has devoted limited attention on how to adjust existing management control systems in order to cope with organizational changes. Consistent with prior literatures stipulating the importance of employee alignment as an important complementary control mechanism to formal contracting channels, our study proposes that creating a more aligned organizational culture can mitigate disruptions from misalignment between firm strategy changes and the associated management control system adaptations. We hope that such findings contribute to the literature on the optimal design of employee incentive systems, and particularly, the interplay between formal and informal management control systems during rapid organizational changes.

Like any empirical study, our study has several limitations. First, relying on rich administrative data from a single field site exposes our study to external validity concerns. This study examines the effects subsequent to a specific kind of organizational change, namely a management-level horizontal merger decision that may lead to corporate restructuring and layoffs. We do not claim that our findings can be generalizable to different kinds of organizational changes. Future research should examine other kinds of organizational changes to clarify important boundary conditions on whether and how the performances of incumbent employees are affected. Second, studies in this area should explore other means by which potential unintended consequences resulting from unexpected career concerns may be mitigated. This study proposes strengthening employee alignment as one such means. However, alternative means may also include *ex-post* adjustments to existing formal management control systems, and *ex-ante* inclusion of preemptive measures in the formal management control systems by which such potential unintended consequences can be mitigated.

CHAPTER 4

RELATIVE PERFORMANCE BENCHMARKS: DO BOARDS FOLLOW THE INFORMATIVENESS PRINCIPLE?

4.1. Introduction

Measurement of performance is a critical element in managerial evaluation and in the design of incentives. A classic insight of principal-agent theory specifies that the principal – in practice, the board – should evaluate managers' performance using performance metrics that are informative about effort or talent ("the informativeness principle") (Holmstrom 1979; Shavell 1979). In particular, when a performance metric contains common noise that is beyond the CEO's control, a metric of relative performance that filters out such noise can be desirable in two ways. First, designing explicit incentive contracts around such relative performance metrics should increase the efficacy of risk sharing and help elicit costly unobservable effort from risk-averse managers. Second, greater precision in measurement of non-systematic performance can increase the agent's *implicit* incentive (e.g., due to career concerns) to exert additional effort (Holmstrom 1999).

In keeping with this insight, a large empirical literature has devoted itself to understanding whether and to what degree corporate managers are evaluated and rewarded on the basis of the systematic and non-systematic components of firm performance (e.g. Antle and Smith 1986; Janakiraman, Lambert, and Larcker 1992). However, a significant challenge to this literature has been the limited visibility to researchers of the actual relative performance metrics chosen by firms – including the metric itself and the peers used to isolate the non-systematic component of it (as discussed in e.g. Albuquerque (2009)) – a due to limited disclosure of such information in public proxy statements.[1]

The SEC's 2006 disclosure reforms provided new opportunities for researchers to revisit these classic questions by requiring firms to disclose the details of their performance-based incentives, including relative performance metrics and peer benchmarks. By 2006, furthermore, compensation practices had evolved significantly since publication of the classic empirical work in this area. This reform therefore provides an exciting opportunity to distill new insights into the evaluation of managerial performance, by allowing for *direct* assessment of the properties of firms' relative performance metrics.

In this paper, we exploit the disclosure reform to examine the relative performance evaluation metrics – in particular, relative shareholder returns (rTSR) – that firms choose to evaluate managers in their performance-based contracts. We assess the extent to which firms' chosen peer-return benchmarks yield rTSR measures that evaluate managers on the basis of the systematic and non-systematic components of firm performance. These questions are increasingly important in light of the shifting landscape in corporate governance and executive compensation. Over the last ten years, rTSR – that is, a firm's own TSR relative to that of an index or a group of peer firms – appears to have become the single most widely used performance metric by which market participants judge companies and their executives. For example, since 2006 the SEC has

[1] As a result, much of the empirical literature has been forced to rely on indirect tests – tests of whether variation in executive compensation is explained by variation in firm performance relative to some measure of "systematic" performance selected by the researcher (e.g., industry or competitive peers) – that have produced mixed findings. For example, Albuquerque (2009) summarizes the evidence for and against relative performance evaluation in CEO compensation and turnover, and argues that the mixed findings in the literature are likely the result of the joint hypothesis problem: that is, researchers are testing both the theory and the choice of an empirical metric to measure the systematic component of firm performance.

required firms to disclose rTSR in their annual reports to shareholders.[2] Also, rTSR has become the predominant metric used in executives' relative-performance-based compensation.[3]

To guide our analysis, in the standard framework for relative performance evaluation, which assumes a linear factor structure (e.g., $p = c + e$ where c is the systematic or common component of performance p and e is the non-systematic or idiosyncratic component), the common component of firm performance should exhibit two properties. First, the firm's performance should have a slope ("benchmark-return-beta") of 1 with respect to the common component. Second, the common component should capture all of the contemporaneous variation in firm performance that can be explained (i.e., non-idiosyncratic to the firm). We examine to what extent firms' self-chosen rTSR peer benchmarks satisfy these theoretical properties.

Our analyses produce four main findings. First, with regards to the first property, we find that all firms on average choose rTSR benchmarks that exhibit benchmark-return-betas of 1, consistent with the systematic component of firm returns. However, with regards to the second property, these benchmarks on average produce rTSR metrics that continue to retain some degree of systematic noise in that alternative benchmarks generate larger time-series R^2 s from contemporaneous return regression than the chosen RP benchmarks. Specifically, we compare firms' chosen rTSR benchmarks to three normative peer benchmarks: the S&P500, firms' self-

[2] The New York Stock Exchange's Listing Company Manual (Section 303A.05) recommends that compensation committees consider a firm's rTSR in determining long-run executive incentives. The influential proxy advisory firm Institutional Shareholder Services (ISS) relies on an analysis of the relationship between a firm's rTSR and its executive's pay relative to peers to judge whether an executive's compensation is justified by performance and to formulate its say-on-pay recommendations. Activist investors often focus on poor rTSR as evidence of poor management quality or poor performance (Brav, Jiang, Partnoy and Thomas, 2008). Finally, as part of its implementation of the 2010 Dodd-Frank Act, the SEC recently proposed Rule No. 34-74835 requiring firms to disclose comparisons of executive compensation to that of peers in terms of rTSR in annual proxy statements.

[3] Figure 4.1 indicates that the proportion of firms with explicit relative performance incentives has increased from 20% in 2006 to 48% in 2014. Back-of-the-envelope estimates suggest that rTSR-based incentives have an economically important and increasing impact on CEO compensation. Between 2006 and 2014, on average, 73% of relative performance target payouts are accounted for by rTSR-based payouts; moreover, meeting rTSR targets increases the CEO's incentive-plan-based compensation (assuming that all other performance-based targets are met) by an average of 40%.

chosen compensation-benchmarking peers, and search-based peers (SBPs).[4] Based on these comparisons, we document that firms' disclosed benchmarks (1) outperform the S&P500 index by 47% in terms of time-series R^2s from contemporaneous-return regressions; (2) perform equally well relative to compensation-benchmarking peers; and (3) underperform SBPs by about 7%.

We also find significant difference between the two predominant approaches to selecting rTSR benchmarks: (a) based on a customized set of peer firms ("specific peers") and (b) based on an industry or market index. Specifically, we find that the underperformance of firms' chosen benchmarks is concentrated in the 40% of firms that choose index-based benchmarks; compared to the index-based benchmarks, SBPs explain 16% more of the time-series variation in firms' monthly stock returns (i.e., time-series regression R^2); and, surprisingly, firms' chosen compensation-benchmarking peers perform 8% better than the index-based benchmarks. These findings raise questions about the appropriateness of choosing indexes – effectively, benchmarking against a large number of peers – in lieu of a narrower but possibly more relevant peer set (e.g., firms' self-chosen compensation-benchmarking peers that are already available). On the other hand, the approximately 60% of firms that use specific peers fare well relative to the normative benchmarks. For example, SBPs only outperform firms' chosen specific peers by 2%; moreover, firms' chosen compensation-benchmarking peers underperform the specific peers chosen for rTSR purposes by over 4%. Viewing SBPs as a normative upper bound in capturing systematic performance, these results indicate that the majority of boards are remarkably efficient at isolating the non-systematic component of performance through their chosen rTSR benchmarks.

Our second set of findings shows that the R^2 differences we document are economically meaningful. We provide an economic interpretation of the R^2 differences in the context of a

[4] Lee, Ma, and Wang (2015, 2016) show that SBPs are superior to other state-of-the-art peer identification schemes at explaining variation in firms' stock returns, valuation multiples, and fundamental performance characteristics.

standard principal-agent framework absent any frictions in peer selection. For any set of performance benchmarks, this framework enables us to translate our R^2 results to the variance in measurement errors (for the systematic component of performance) resulting from the chosen peers (up to a scalar constant). More importantly, this framework allows us to calibrate the performance consequences of poor rTSR benchmarks. Our results suggest that, in the absence of frictions against selecting a precise set of peers, measurement errors due to firms' choices of performance benchmarks reduce managerial effort and result in an on-average performance penalty of 60 to 153 basis points in annual stock returns, under plausible values of risk aversion. Again, these effects are driven by the subset of firms that use index-based benchmarks, for which we find an on-average penalty of 106-277 bps in annual stock returns. These magnitudes are economically significant, particularly in light of the relatively large size of the firms in our sample.

Our third set of findings shows that the observed selection of poorly performing rTSR benchmarks, due to the selection of indexes, is probably a result of governance-related frictions and compensation consultants' tendencies. We begin with a theoretical analysis of potential sources of economic friction that can rationalize the observed underperformance of firms' chosen benchmarks. We generalize the standard principal-agent framework by endogenizing the board's choice of relative performance-benchmarking efficacy, and show that less efficacious benchmarking can be expected when the firm has greater idiosyncratic volatility, and when the manager or the board is of lower quality or ability (due to the lower marginal incentive effects of improving benchmarks). Our empirical evidence does not lend support to firm-level volatility or managerial quality as explanations of benchmarking inefficacy.

Moreover, we argue that the empirical evidence does not support the interpretation that observed patterns of benchmark selection are expected outcomes of plausible alternative theories

– that is, explanations external to the model – about how boards select peers for relative performance metrics, such as (1) firms' own actions influencing peer performance (Janakiraman, Lambert and Larcker 1992; Aggarwal and Samwick 1999a), (2) firms trading off ex-ante vs. ex-post efficiency due to the perception that indexes are less gameable (Godfrey and Bourchier 2015; Walker 2016), (3) tournament theory (Lazear and Rosen 1981; Hvide 2002), (4) firms selecting benchmarks on the basis of aspiration (Scharfstein and Stein 1990; Hayes and Schaefer 2009; Hemmer 2015; Francis, Hasan, Mani and Ye 2016), (5) managers' ability to self-insure against the systematic factor (Garvey and Milbourn 2003), (6) firms having alternative production technology (Hoffmann and Pfeil 2010; DeMarzo, Fishman, He and Wang 2012), or (7) firms facing differential implementation costs between specific peers and index-based benchmarks. At a basic level, none of these alternative theories predicts that firms' chosen rTSR benchmarks would exhibit a benchmark-return-beta of 1.

Instead, the selection of poorer (i.e. index-based) benchmarks is systematically associated with proxies for governance weaknesses, such as excess compensation, large board size, and heavy director workload. Furthermore, our empirical evidence also reveals that compensation consultants' systematic tendencies to recommend indexes or specific peers, in combination with governance weaknesses, can explain why certain firms choose relatively poor rTSR benchmarks, even given the availability of such superior options as the specific peers that boards have chosen for compensation-benchmarking purposes. Overall, our evidence suggests that weaker boards' failure to carefully scrutinize the default recommendations of a compensation consultant will result in managerial evaluation based on poor performance measures.

Finally, our fourth set of findings shows that the choice of an index-based benchmark is cross-sectionally associated with lower *realized* annual ROA. Firms with index-based benchmarks perform 70 basis points lower in ROA than firms with specific-peer benchmarks. We find similar results using annual stock returns. This reduced-form analysis treats firms that use specific-peer benchmarks as counterfactuals to firms that use index-based benchmarks; to the extent that our control variables do not fully capture differences in the underlying characteristics of these two types of firms that could be associated with firm performance, it is possible that the estimated differences are overstated. Nevertheless, these reduced-form estimates provide empirical support for the calibration exercise's results that poorer benchmark selection is associated with economically significant consequences for firm performance.

Our results make several contributions to the literature. First, given the rising importance of rTSR for evaluating managerial performance, we provide novel evidence on the properties of the rTSR measures chosen by boards to evaluate executives in relative-performance contracts. In particular, our findings provide more direct and novel answers to the longstanding questions about whether and to what extent boards evaluate managers on the basis of systematic and non-systematic components of performance (Antle and Smith 1986; Lambert and Larcker 1987; Albuquerque 2009). Our results suggest that, at the majority of firms, boards do a remarkable job of capturing the systematic component of firm returns via the appropriate selection of peers. This finding is particularly surprising in light of the conventional view that the executive-pay-setting process is often compromised and captured by powerful CEOs (Bebchuk, Cremers, and Peyer 2011; Morse, Nanda, and Seru 2011). Second, we provide novel evidence on how compensation consultants impact the managerial evaluation and pay-setting process. Our results show how compensation consultants' preferences interact with board-level governance weaknesses to produce relatively poor performance-evaluation metrics. Third, our findings suggest that the properties of the performance metrics chosen by boards to evaluate executives can serve as novel indicators of board-level governance quality.

Finally, we add to the body of work that has emerged since the 2006 mandate on compensation benchmarking practices. Examining only firms with specific peers as relative performance benchmarks, Gong, Li, and Shin (2011) suggests that firms' chosen relative-performance benchmarks perform better than randomly chosen benchmarks in eliminating common shocks, but the analysis does not investigate the overall extent to which firms' benchmarking practices adhere to the informativeness principle and its implications. [5] Complementing our results, Bizjak, Kalpathy, Li and Young (2016) focuses on the effect of the choice of performance benchmarks on CEO compensation levels, finds that boards' selection of performance peers does not affect the level of pay in an economically significant manner.

The rest of the paper is organized as follows. Section 4.2 lays out data and descriptive statistics illustrating the rise of explicit grant-based relative-performance benchmarking; it also provides empirical evidence on the efficacy of the board's choice of benchmarks. Section 4.3 maps our empirical test to the principal-agent framework in order to recover primitives that describe the efficacy of relative performance benchmarks and its implications for firm performance. Section 4.4 investigates the possible sources of, and alternative theories about, the observed patterns of benchmarking. Section 4.5 investigates cross-sectional consequences in realized firm performance associated with benchmarking quality. Section 4.6 concludes.

[5] A related set of papers tests for whether a firm's choice in tying incentives to relative performance, on the extensive margin, is consistent with its costs and benefits as predicted in Gibbons and Murphy (1990). Carter, Ittner, and Zechman (2009) studies a sample of UK firms and finds that the propensity to tie incentives to relative performance – identified via explicit disclosures – is not associated with the degree of a firm's exposure to systematic shocks. Gong et al. (2011) finds the opposite result in the US.

4.2. Data and Descriptive Evidence of Benchmarking Behavior

This section provides empirical evidence on the quality of firms' chosen relative-performance benchmarks, in terms of the extent to which they follow the informativeness principle and eliminate common performance shocks. Our analyses focus on the sample of firms that explicitly tie executive compensation to rTSR; the quality or informativeness of rTSR is expected to be of greater importance to such firms.

4.2.1. Data Description

Our data come from ISS Incentive Lab, which collected details on compensation contracts and incentive-plan-based awards of named executive officers, at the individual-grant level, from firms' proxy statements. Incentive Lab covers every U.S. firm ever ranked in the top 750 in terms of market capitalization in any year since 2004. Due to backward- and forward-filling, the raw Incentive Lab data (2004-2014) encompasses the entire S&P500, most of the S&P Midcap 400, and a small proportion of the S&P Small-Cap 600. Thus, roughly speaking, each annual cross-section encompasses the largest 1,000 firms listed on the U.S. stock market in terms of market capitalization. Our analysis focuses on the sample from 2006 onward; mandatory disclosure of compensation details began in 2006, and coverage of firms is more comprehensive after that year.

For each grant, ISS Incentive Lab collected information on the form of the payout (cash, stock options, or stock units); conditions for payout (tenure [Time] and for fulfillment of absolute performance criteria [Abs] or relative performance criteria [Rel] or a combination of the two [Abs/Rel]); and specific accounting- or stock-based performance metrics associated with performance-based grants. Finally, ISS Incentive Lab collected information on the specific peer firms or indexes selected for purposes of awarding grants based on relative performance.

Table 4.1, Panel A, provides summary statistics on 34,321 CEO grants awarded by 1,547 unique firms in the 2006-2014 period. During this period, on average, companies awarded 3.2 CEO grants per year. The proportion of incentive awards paid out in cash is stable within the sample period at roughly 35% of all CEO grants; in the same time period, stock-based payouts increased from 36% to 49% while option-based payouts declined from 29% to 15%. Notably, the proportion of CEO grants that included a relative performance component (Abs/Rel or Rel) more than doubled, from 8% in 2006 to 17% in 2014.[6]

Table 4.1, Panel B, suggests that, at the firm level, usage of relative performance incentives has more than doubled since 2006. Relative to the total number of companies in our sample, the proportion of firms with explicit relative performance (RP) incentives increased from 20% in 2006 to 48% in 2014 (see the solid line in Figure 4.1). Moreover, Panel C suggests that the use of rTSR has been increasingly prevalent at such firms: whereas 70% of the companies that provide RP incentives used rTSR in 2006, 87% did so by 2014 (see the dashed line in Figure 4.1). Jointly, the summary statistics presented in Table 4.1 and Figure 4.1 illustrate the increasing pervasiveness of explicit RP-based incentives and the prominence of rTSR in such incentive plans.

To further assess the economic magnitude of CEOs' RP-based incentives, Table 4.2 provides back-of-the-envelope estimates of the relative importance of meeting RP targets. We estimate how much incremental incentive-plan-based compensation the CEO would earn by meeting RP-based targets, assuming that all other incentives are earned. Column 3 estimates

[6] This increase in the explicit use of relative performance grants is consistent with descriptive evidence from the prior literature. For example, our summary statistics are comparable to those of Bettis, Bizjak, Coles and Young (2014) which also uses data from ISS Incentive Lab spanning the time period 1998-2012 (e.g., see their Table 1).

Table 4.1. Summary Statistics on CEO Grants 2006-2014

Panel A reports summary statistics for compensation grants awarded to the CEO in fiscal years 2006-2014 using the ISS Incentive Labs data prior to any sample selection restrictions. We report the total number of unique firms, the average number of grants awarded to the CEO in each year, the average of the proportion of each award payout type (cash, option, or stock) to the total number of grants awarded to the CEO, and the average of the proportion of each performance evaluation type (absolute performance, relative performance, a mix of the two, and time-based) to the total number of grants awarded to the CEO. Panels B and C report the same summary statistics for sub-samples conditional on CEO grants with a relative performance component and a rTSR component respectively.

			Payout Type [Grant-level]			Evaluation Type [Grant-level]			
Fiscal Year	Unique # of Firms	Unique # of Grants	Cash	Option	Stock	Abs	Abs/Rel	Rel	Time
Panel A: All CEO Grants									
2006	1,278	2.86	0.35	0.29	0.36	0.42	0.04	0.04	0.49
2007	1,283	3.06	0.35	0.26	0.39	0.44	0.05	0.04	0.48
2008	1,249	3.06	0.35	0.25	0.40	0.44	0.05	0.04	0.47
2009	1,153	3.13	0.35	0.24	0.41	0.43	0.05	0.04	0.47
2010	1,165	3.30	0.34	0.21	0.45	0.43	0.06	0.05	0.46
2011	1,159	3.29	0.33	0.20	0.47	0.44	0.07	0.05	0.43
2012	1,173	3.31	0.35	0.18	0.47	0.46	0.09	0.06	0.40
2013	1,155	3.31	0.34	0.17	0.49	0.46	0.10	0.06	0.38
2014	1,108	3.56	0.35	0.15	0.49	0.47	0.11	0.06	0.36
Panel B: CEO Grants with RP Component									
2006	257	1.22	0.35	0.02	0.62	-	0.55	0.45	-
2007	279	1.27	0.36	0.02	0.62	-	0.54	0.46	-
2008	289	1.24	0.29	0.02	0.69	-	0.52	0.48	-
2009	289	1.29	0.32	0.01	0.67	-	0.53	0.47	-
2010	343	1.24	0.28	0.01	0.72	-	0.52	0.48	-
2011	384	1.23	0.23	0.01	0.76	-	0.52	0.48	-
2012	456	1.27	0.21	0.01	0.78	-	0.56	0.44	-
2013	489	1.22	0.19	0.00	0.81	-	0.59	0.41	-
2014	530	1.28	0.17	0.00	0.82	-	0.63	0.37	-
Panel C: CEO Grants with rTSR Component									
2006	180	1.18	0.24	0.02	0.73	-	0.49	0.51	-
2007	206	1.18	0.27	0.01	0.72	-	0.50	0.50	-
2008	217	1.18	0.20	0.01	0.79	-	0.49	0.51	-
2009	220	1.21	0.22	0.01	0.77	-	0.48	0.52	-
2010	264	1.18	0.19	0.00	0.81	-	0.47	0.53	-
2011	312	1.17	0.16	0.00	0.83	-	0.47	0.53	-
2012	380	1.17	0.15	0.01	0.84	-	0.53	0.47	-
2013	420	1.13	0.13	0.00	0.86	-	0.57	0.43	-
2014	459	1.18	0.12	0.00	0.88	-	0.62	0.38	-

Figure 4.1. Fraction of Firms Using Relative Performance Contracts 2006-2014

The solid line plots the fraction of firms in the ISS Incentive Labs sample prior to any sample selection restrictions that disclose awarding at least one performance grant based on relative performance (RP) in a given fiscal year; the dotted line plots the fraction of firms with at least one performance grant based relative performance that use rTSR as the metric of relative performance.

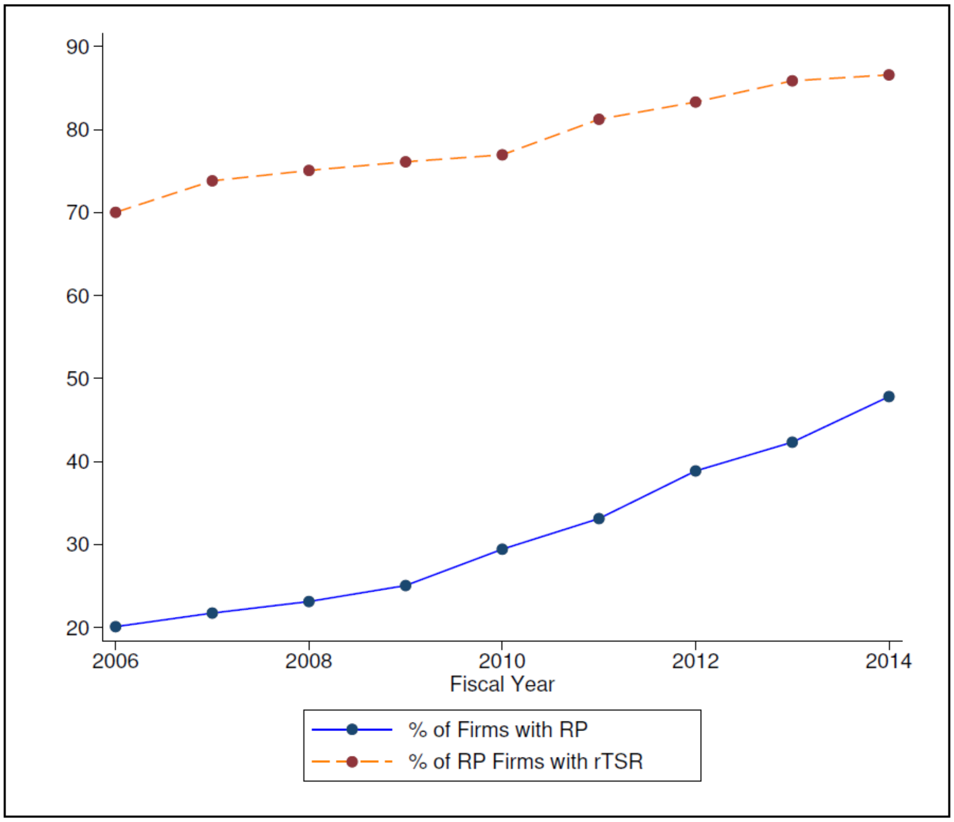


Table 4.2. Importance of CEO Relative Performance Incentives

This table reports summary statistics on the relative performance incentive ratio and the relative performance stock incentive ratio of compensation grants awarded to CEOs in fiscal years 2006-2014 using the ISS Incentive Labs data prior to any sample selection restrictions. The RP incentive ratio measures the incremental potential incentive when the CEO meets all RP-based targets; it is calculated as (expected incentive-plan-based compensation if all targets are met)/(expected incentive-plan-based compensation if all other targets excluding RP metric-based targets are met). The rTSR incentive ratio measures the incremental potential incentive when the CEO meets all RP-based stock price targets (i.e. rTSR); it is calculated as (expected incentive-plan-based compensation if all targets are met)/(expected incentive-plan-based compensation if all other targets excluding rTSR targets are met). The amount of expected incentive-plan-based compensation is calculated using the values reported in the Grants of Plan-Based Awards Table in the proxy statement which includes both annual and long-term incentive plans. Specifically, it is computed by adding the target dollar value of Estimated Future Payouts Under Non-Equity Incentive Plan Awards and Grant Date Fair Value of Stock and Option Awards (which are based on meeting the performance target). For grants that use multiple performance metrics, we calculate the weighted portion of expected compensation that corresponds to each performance metric. We assume that each performance metric is weighted equally in the calculation of the grant. Column 3 reports the average expected incentive-plan-based compensation. Columns 4 and 5 report the portion of column 3 attributable to RP-based metrics and rTSR metrics, respectively. Column 6 reports the average proportion of RP-based compensation attributable to stock price-based metrics. Columns 7 and 8 report the accompanying incentive ratios.

		Expected Incentive-Plan-Based Compensation				**Incentive Ratios**	
Fiscal Year	Unique # of Firms	Total	RP	rTSR	Fraction	RP	rTSR
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2006	235	6,640,461	1,768,730	1,233,457	0.64	1.63	1.38
2007	265	7,710,858	1,962,250	1,369,169	0.67	1.55	1.34
2008	277	7,266,247	2,050,361	1,563,382	0.70	1.61	1.42
2009	280	6,078,541	1,787,294	1,336,082	0.71	1.65	1.4
2010	332	6,707,443	1,898,931	1,404,596	0.71	1.59	1.39
2011	377	7,160,502	1,820,329	1,389,227	0.75	1.55	1.39
2012	446	7,445,771	2,077,538	1,596,345	0.78	1.58	1.43
2013	478	7,727,030	2,082,448	1,650,666	0.81	1.56	1.41
2014	524	7,950,395	2,005,443	1,505,951	0.81	1.50	1.35

expected total plan-based compensation when all incentives are earned, including meeting all RP-based targets.[7] Columns 4 and 5 estimate the allocated expected compensation stemming from meeting RP-based targets and from meeting rTSR-based targets respectively. Overall, RP-based incentives comprise a significant proportion of the total expected plan-based compensation, with rTSR accounting for the vast majority of RP-based incentives (73% on average as reported in column 6).

We also estimate the improvement in incentive-plan-based compensation expected from meeting RP-based and rTSR-based targets (the "Incentive Ratio" reported in columns 7 and 8). Column 7 suggests that, relative to not meeting RP-based targets, meeting them increases CEOs' plan-based compensation by an average of 58%, assuming all other incentives are earned. Column 8 suggests that, assuming that all non-rTSR-based RP targets are met and that all other incentives are earned, meeting rTSR-based targets increases CEOs' plan-based compensation by an average of 40%.[8]

Our back-of-the-envelope estimates are consistent with existing and growing evidence of the importance of performance-based – and in particular RP-based – incentives for CEOs. For example, Bettis et al. (2014) shows that the RP-related components of compensation at RP-grant-issuing firms between 1998 to 2012 consistently determined more than 30% of the realized total compensation amount. Similarly, De Angelis and Grinstein (2016) shows that, for a hand-collected

[7] Expected compensation is calculated using values reported in the Grants of Plan-Based Awards Table by adding the dollar values of Estimated Future Payouts Under Non-Equity Incentive Plan Awards based on target performance and the Grant Date Fair Value of Stock and Option Awards reported in the proxy statements.

[8] The incentive ratio in column 7 (8) is calculated as the expected plan based compensation from meeting RP-based targets, assuming that all other incentives are earned as reported in column 3, divided by the counterfactual expected compensation excluding RP-based allocations (rTSR-based allocations). For example, the RP-based incentive ratio of 1.50 in 2014 implies that on average, CEOs who achieve their RP-based targets can earn 50% more than the counterfactual in which they do not earn their RP-based threshold performance payouts. When an incentive grant involves multiple performance criteria, we attribute the total expected payout from meeting all targets equally to each metric.

sample of S&P500 firms in 2007, about one-third of firms explicitly mentioned that their performance-based awards were RP-based, and that firms with RP contracts attributed about half of the estimated total performance award value to RP. The paper also documents that about 75% of the performance metrics associated with RP are market measures; this finding is consistent with the notion that stock-price-based measures prevail for relative performance purposes.

Table 4.3 provides information on the different types of benchmarks used to measure relative performance. The sample of RP grants is identical to Table 4.1, Panel B. Specifically, we consider four benchmark categories: a specific peer set, the S&P500 index, the S&P 1500 index, and other indexes (typically industry-based). Columns 4-7 report the percentages of RP grants that use each type of benchmark in a given fiscal year. Column 8 reports the percentage of RP grants whose benchmark cannot be identified. Because each grant can be associated with multiple types of benchmarks, the sum of the values across columns 4-8 can exceed one. Finally, column 9 reports the average number of peer firms used by firms that opt for a specific peer set.

Overall, we observe that around half of all relative-performance grants use specific peers as a benchmark, and that the average number of peers is 15-18. At firms that choose an index benchmark, the most popular choice is the S&P500. In fiscal year 2014, for example, 48% of RP grants to CEOs identify specific peer firms as the relative benchmark; 21% use the S&P500 or 1500 indexes, 19% use another index (e.g., narrower or industry-specific indexes), and 15% do not specify a peer benchmark. The distribution of relative benchmark types remained stable over the eight-year period from 2006 to 2014. Among the firms that chose an index, the distribution of index choices also remained stable; in 2014, for example, 40% chose the S&P500, 12.5% chose the S&P1500, and the remaining 47.5% chose other indexes.

Table 4.3. Summary Statistics on Types of Relative Performance Benchmarks 2006-2014

This table summarizes the percentages of rTSR-based grants associated with different types of relative benchmarks for fiscal years 2006-2014 using the ISS Incentive Labs data prior to any sample selection restrictions. Columns 2 and 3 report the unique numbers of firms and grants respectively. Columns 4-8 report the percentages of RP that use each type of benchmark: specific peers, the S&P500 index, the S&P1500 index, other indexes (typically industry-based), and unspecified. Because each grant can be associated with multiple types of benchmarks, the values across columns 4-8 can exceed one. Column 9 reports the average number of peer firms chosen as benchmarks for RP grants associated with specific peers.

			Relative Performance Benchmark Type					
Fiscal Year	Unique # of Firms	Unique # of Grants	Specific Peer	S&P500	S&P1500	Other Index	Not Specified	# of Peers
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
2006	257	313	0.49	0.14	0.05	0.17	0.20	14.79
2007	279	355	0.55	0.16	0.05	0.14	0.14	15.80
2008	289	358	0.57	0.18	0.04	0.17	0.10	16.71
2009	289	373	0.55	0.17	0.05	0.13	0.15	16.43
2010	343	424	0.56	0.14	0.04	0.15	0.17	17.31
2011	384	471	0.55	0.15	0.05	0.15	0.16	18.09
2012	456	579	0.51	0.15	0.04	0.16	0.18	17.07
2013	489	596	0.49	0.18	0.05	0.17	0.15	17.49
2014	530	678	0.48	0.16	0.05	0.19	0.15	16.99

4.2.2. Assessing Properties of RP Benchmarks

Given the rising importance of rTSR as a metric for judging and incentivizing managerial performance, our paper seeks to assess the efficacy of boards' performance-measurement choices. We examine the extent to which boards' choices of relative-performance benchmarks follow the informativeness principle (Holmstrom 1979; Holmstrom and Milgrom 1987). In other words, how well do firms' choices of RP benchmarks capture the systematic component of their stock returns?

We examine whether chosen RP benchmarks satisfy the expected properties of the common component in firm performance. In the standard framework for relative performance evaluation, which assumes a linear factor structure (e.g., $p = c + e$ where c is the systematic or common component of performance p and e is the non-systematic or idiosyncratic component), the common component of firm performance should exhibit two properties.[9] First, the firm's performance should have a slope ("benchmark-return-beta") of 1 with respect to the common component.[10] Second, the common component should capture all of the contemporaneous variation in firm performance that can be explained (i.e., non-idiosyncratic to the firm).

To examine whether firms' chosen RP benchmarks exhibit these two properties, we estimate the following time-series returns regression for each firm:

$$R_{it} = \alpha_i + \beta_i R_{p_{it}} + \varepsilon_{it} \tag{1}$$

where R_{it} is firm i 's monthly cum-dividend returns in period t and $R_{p_{it}}$ is the returns of that firm's benchmark peers. Our empirical analysis focuses on those firms that tie their CEOs'

[9] That the factor structure is linear is without loss of generality. A unique linear structure with respect to a set of factors is guaranteed by the projection theorem. Note also that a linear factor structure is consistent with the relative performance metrics observed in practice, like rTSR, which are expressed as the difference between two performance measures. That the factor structure has a single common component is also without loss of generality: with multiple factors, the common component is simply $c = b'f$, where b' is a vector of factor sensitivities and f is a vector of factor returns.

[10] Together with the definition of rTSR, this prediction is also implied in Equation 3 of Janakiraman et al. (1992).

performance-based incentives to rTSR; the quality of the RP metric should be especially important to them. We therefore restrict our attention to the subsample of firms covered by ISS Incentive Lab that (1) issued rTSR-based grants to their CEOs (that is, the sample described in Table 4.1, Panel C), (2) disclose the peers or indexes used in determining performance payouts, and that (3) intersect with available alternative benchmark peers introduced by Lee et al. (2015). In total, our sample consists of 356 unique firm-benchmark-type (i.e., index vs. specific peers) observations between fiscal years 2006 and 2013; this sample represents 330 unique firms, due to the inclusion of 26 firms that switched benchmark types during the sample period. Returns data are obtained from CRSP monthly files, and firms with fewer than ten months of valid monthly returns in total are excluded from the sample. Detailed construction of our final sample is described in the Appendix Table I.

In estimating equation (1), we use the median of the peer set's returns for firms that select a set of specific RP peer firms. Although the choice of the order statistic from the peer-return distribution can be arbitrary, median is the most popular performance target in relative-performance contracts (Reda and Tonello 2015; Bennett, Bettis, Gopalan, and Milbourn 2017).[11] For firms that select an index as the relative benchmark, we use the corresponding index returns. For the RP benchmarks disclosed in the proxy statement for a given fiscal year, we use returns from the following fiscal year. For example, if firm i reports its fiscal-year-end date as December 2000, we obtain monthly stock-return data for the calendar window January 2001 to December 2001 for it and for its performance peers, disclosed in that proxy statement, to calculate $R_{p_{it}}$. Our choice reflects how the selected peers are *used* in RP contracts and thus how they relate to realized

[11] The optimal aggregation rule from an informativeness perspective is discussed in Holmstrom (1982) and Dikolli, Hofmann, and Pfeiffer (2012).

firm performance *ex post*. However, the results of our paper do not change if we use the prior-year stock returns, due to a low level of turnover in firms' chosen RP peers.

To examine the first property of the common component of firm performance, we analyze the slopes from estimating equation (1), summarized in Table 4.4. Interestingly, we find a cross-sectional average slope coefficient of 1.03 across all firms, which is statistically no different from the normative benchmark of 1. Moreover, we find that the average slope is close to, and statistically no different from, 1 for the subset of firms that choose specific-peer benchmarks and for the subset that choose index-based RP benchmarks. Thus, it appears that firms' chosen benchmarks, whether index or specific peers, exhibit a return-regression slope of 1, consistent with these benchmarks capturing the systematic component of firm performance.

To examine the second property of the common component of firm performance, we analyze the R^2s from estimating equation (1), summarized in Table 4.5. If firms' chosen RP benchmarks capture the common component of their stock returns, no other contemporaneous variables should exhibit incremental explanatory power. Equivalently, no other contemporaneous variables or alternative peer formulations should generate R^2(from estimating equation (1)) higher than can be obtained by using the chosen RP benchmark returns.

To test this implication, we choose three alternative RP benchmarks: the S&P500 index, firms' compensation-benchmarking peers and the search-based peer firms (SBPs) of Lee et al. (2015). We choose the S&P500 index as a normative benchmark since it is salient and a "free option" to firms. We examine compensation-benchmarking peers because they represent a set of peers already chosen by firms' boards of directors – in this case to set their CEOs' compensation levels – and thus a readily available alternative. Appendix Table II reports summary statistics

Table 4.4. Assessing Firms' Chosen RP Benchmarks: Slope Coefficient

This table estimates and compares the cross-sectional average slope coefficient values (β) from time-series regressions of the form

$$R_{it} = \alpha_i + \beta_i R_{p_{it}} + \epsilon_{it}$$

using CRSP monthly returns data. Column 1 reports the across-firm average slope coefficient from time-series regressions, regressing base firm *i*'s returns on the concurrent returns of a portfolio of peers. Column 2 reports the p-value of the null test of $\beta = 1$, and column 3 reports the average number of observations per firm.

Results are reported for the sample of base firms whose chosen benchmarks are identifiable in the data from ISS Incentive Lab. We use return data from 2006-2013 for firms for which there are at least 10 observations. The first row reports on all firms in our sample that satisfy these filters; the second row estimates the same regressions on the subset that select specific peers as benchmarks; the third row estimates the same regressions on the subset that select a stock index as a benchmark.

Standard errors are reported in brackets and significance levels are indicated by *, **, *** for 10%, 5%, and 1% respectively.

Sample	β	p-value $H_0: \beta = 1$	Mean Obs per Firm
	(1)	(2)	(3)
All	1.0255***	0.3272	36.820
(N=356)	[0.0258]		
Specific Peers	1.0052***	0.8765	40.552
(N=201)	[0.0329]		
Index	1.0520***	0.1864	31.981
(N=155)	[0.0387]		

Table 4.5. Assessing Firms' Chosen RP Benchmarks: R^2

This table estimates and compares average R^2 values from time-series regressions of the form

$$R_{it} = \alpha_i + \beta_i R_{p_{it}} + \epsilon_{it}$$

using CRSP monthly returns data. In Panel A, columns 1, 2, and 3 report across-firm average R^2 from time-series regressions, regressing base firm i's returns on the concurrent returns of a portfolio of peers. Column 1 uses the median returns of firms' chosen relative performance benchmarks; column 2 uses the returns of the S&P500 index; and column 3 uses the mean returns of search-based peers (Lee et al. 2015). Column 4 reports the differences between R^2 values reported in columns 2 and 1; column 5 reports the differences between columns 3 and 1. Column 6 reports the average number of observations per firm. Panel B re-estimates Panel A with compensation-benchmarking peers (CBP) as an additional alternative normative benchmark.

Results are reported for the sample of base firms whose chosen benchmarks are identifiable in the data from ISS Incentive Lab. We use return data from 2006-2013 for firms for which there are at least 10 observations. The first row reports on all firms in our sample that satisfy these filters; the second row estimates the same regressions on the subset that select specific peers as benchmarks; the third row estimates the same regressions on the subset that select a stock index as a benchmark.

To facilitate comparisons, all the regressions are conducted using the same underlying set of firms. The reported N in parentheses represents the number of firms-benchmark combinations contained in each sample. Standard errors are reported in brackets and significance levels are indicated by *, **, *** for 10%, 5%, and 1% respectively.

Panel A: Baseline Sample

Sample	RP Benchmarks (1)	S&P500 (2)	SBP (3)	(2)-(1) (4)	(3)-(1) (5)	Mean Obs per Firm (6)
All	0.483***	0.328***	0.518***	-0.155***	0.035***	36.820***
(N=356)	[0.012]	[0.010]	[0.012]	[0.011]	[0.008]	[1.375]
Specific Peers	0.548***	0.328***	0.560***	-0.220***	0.013	40.552***
(N=201)	[0.015]	[0.012]	[0.015]	[0.014]	[0.009]	[1.898]
Index	0.400***	0.329***	0.464***	-0.071***	0.064***	31.981***
(N=155)	[0.019]	[0.017]	[0.019]	[0.013]	[0.015]	[1.919]

Table 4.5. Assessing Firms' Chosen RP Benchmarks: R^2 (Continued)

Panel B: Compensation Benchmarking Peers Subsample

Sample	RP Benchmarks (1)	CBP (2)	SBP (3)	(2)-(1) (4)	(3)-(1) (5)	Mean Obs per Firm (6)
All	0.497***	0.496***	0.525***	-0.001	0.028***	37.547***
(N=322)	[0.013]	[0.012]	[0.013]	[0.007]	[0.008]	[1.433]
Specific Peers	0.559***	0.534***	0.568***	-0.025***	0.009	40.837***
(N=190)	[0.015]	[0.015]	[0.015]	[0.008]	[0.009]	[1.921]
Index	0.409***	0.442***	0.464***	0.033***	0.055***	32.811***
(N=132)	[0.020]	[0.020]	[0.021]	[0.011]	[0.015]	[2.079]

regarding the extent of overlap between firms' chosen RP and compensation benchmarking peers. Finally, we utilize SBPs, which represent firms' economic benchmarks as collectively perceived by investors and inferred from co-search patterns on the SEC's Electronic Data-Gathering, Analysis, and Retrieval (EDGAR) website, because the findings of Lee et al. (2015, 2016) suggest that SBPs prevail over other state-of-the-art methods for identifying economically related firms for purposes of explaining co-movement of stock returns, valuation multiples, growth rates, R&D expenditures, leverage, and profitability ratios.[12]

Table 4.5, Panel A, compares the R^2s generated by firms' selected benchmarks to those obtained from both the S&P500 index and the SBPs.[13] To facilitate comparisons, all regressions are conducted using the same underlying set of base firms. In columns 1, 2, and 4, the first row shows that, across all 356 unique firm-benchmark observations, using the S&P500 as the benchmark yields an average R^2 of 32.8%, which is significantly (at the 1% level) lower than the 48.3% generated by firms' chosen RP benchmarks, consistent with the latter's capture of the common component of performance. However, in columns 1, 3, and 5, the first row shows that SBPs produce R^2s (of 51.8%) that are significantly higher (at the 1% level) than those produced by firms' chosen RP benchmarks, inconsistent with the latter's capture of the common component of performance.

[12] Among S&P500 firms, for example, an equal-weighted portfolio of top-10 SBPs explains 63% more of the variation in base-firm monthly stock returns than a randomly selected set of 10 peers from the same 6-digit Global Industry Classification System industry. A search-traffic-weighted portfolio of top-10 SBPs, weighted by the relative intensity of co-searches between two firms (a measure of perceived similarity), explains 85% more of the variation in base-firm monthly returns.

[13] We weight peer returns by the relative magnitude of EDGAR co-search fractions, interpreted as a measure of similarity or relevance between firms. Lee et al. (2015, 2016) show that this weighting scheme performs best at explaining contemporaneous variations in base-firm returns. To avoid look-ahead bias, we follow Lee et al. (2015) in always identifying SBPs using search traffic from the prior calendar year.

We also examine how these results vary across the two prevailing approaches to selecting RP benchmarks: (a) based on a customized set of peer firms ("specific peers") and (b) based on an industry or market index. Prior literature has suggested that a narrower set of peer firms is generally more capable of measuring the common factor in performance than are broad indexes (Lewellen and Metrick 2010). Table 4.5, Panel A, rows 2 and 3, compares the R^2s generated by firms' selected benchmarks for the subsets of firms that use specific peers (N=201) and an index (N=155) respectively. In both cases, we find that firms' selected RP benchmarks outperform the S&P500, although the index-based benchmarks outperform by a smaller amount – a 21.6% proportional improvement – than do the specific peers, whose average R^2 outperforms those produced by the S&P500 proportionally by 67.1%.[14] However, we find that the overall underperformance of firms' benchmarks relative to SBPs is concentrated among the set of firms that use index-based benchmarks, among which the average time-series R^2 is 40.0%. By comparison, the average R^2s produced by SBPs (46.4%) represent not only a statistically significant improvement but also an economically significant one (a 16% proportional improvement). On the other hand, specific RP peers produce an average R^2 of 54.8%, which is statistically no different from the average R^2 of 56.0% generated from firms' SBPs.

In untabulated tests, we assess the robustness of the above findings in a few ways. First, we re-frame the analysis by comparing the variance in firms' chosen rTSR metrics and the variance in their realized TSR. Consistent with the R^2 results and with firms' intention to remove systematic noise, we find that rTSR exhibits significantly lower variance compared to TSR. However, we see relatively less improvement for the rTSR using index-based benchmarks (which exhibit 30% lower variance on average, significant at the 1% level) than for rTSR metrics using specific peers as

[14] Though index-based benchmarking firms include those that use the S&P500, the outperformance stems from narrower industry indexes that are also part of the group.

benchmarks (which exhibit 48% lower variance on average, again significant at the 1% level). Second, we examine alternative rules for aggregating peer performance, namely the mean and the 75^{th} percentile of peer portfolio returns. Since these variations do not affect index returns, these robustness tests focus only on firms that use specific RP peers. As in Panel A, the mean (75^{th} percentile) of chosen peers' returns yields an average time-series R^2 of 54.7% (51.7%) in return regressions across the set of specific-peer-benchmarking firms. The underperformance of 4.4% relative to SBPs is statistically significant at the 1% level for the 75^{th} percentile of specific peer performance, but not significant for the mean of the portfolio of specific peers. Third, we examine how results differ by using an alternative normative peer benchmark – the peers most commonly co-covered by sell-side analysts ("ACPs" of Lee et al (2016)) – and find results very similar to those using SBPs.

Table 4.5, Panel B, compares the R^2s generated by the selected benchmarks of the 322 firms whose compensation-benchmarking peers are reported in ISS Incentive Labs. Using this sample, and the subsample of firms that choose specific peers and index-based benchmarks, the relative comparisons to SBPs yield virtually identical results. However, the comparison to firms' compensation-benchmarking peers reveals surprising results. In particular, although the R^2s produced by firms' selected RP benchmarks are statistically no different from the R^2s produced by firms' compensation-benchmarking peers, the subsamples of specific peers and index-based benchmarks reveal very different patterns. Whereas specific peers produce R^2s that are on average 4.7% higher than those produced by compensation-benchmarking peers, a difference that is statistically significant at the 1% level, index-based benchmarks produce R^2s that are on average 7.5% lower than those produced by compensation-benchmarking peers, a difference that is statistically significant at the 1% level.

Overall, we draw two conclusions from the empirical findings in Tables 4.4 and 4.5. First, the majority of firms – those choosing specific peers – select RP benchmarks in accordance with the informativeness principle. Consistent with capturing the common component of firms' returns, these specific-peer benchmarks exhibit a slope of 1 and exhibit R^2s of magnitudes similar to those of SBPs. However, nearly half of the firms – those choosing indexes – do not select RP benchmarks in accordance with the informativeness principle. Although these index-based benchmarks also exhibit a slope of 1, they fail to capture a significant proportion of variation in performance that can be explained by the common component of returns. This finding is particularly puzzling in light of the availability of compensation-benchmarking peers that can serve as RP benchmarks: they not only capture well the common component of firms' returns, as shown in Panel B of Table 4.5, but untabulated tests also show that they exhibit a slope of 1 from estimating equation (1). One possibility is that these choices reflect alternative considerations of heterogeneous firms, a point we will return to in Section 4.5.

4.3. Interpreting R^2 Differences

The results on R^2s presented in the last section suggest that while the majority of firms adhere to the informativeness principle in selecting their RP benchmarks, over 40% of firms' chosen peers fail to capture the common component of firms' returns for relative performance assessments. Although these findings on R^2 are suggestive, they are difficult to interpret economically. In this section, we interpret these R^2 results in the context of a classic principle-agent model. Our intent is simply to provide a familiar and baseline framework for assessing the firm-performance ramifications of firms' RP-benchmark inefficacy. In Section 4.5, we generalize this model to examine potential channels that could rationalize the calibrated firm-performance implications in this baseline framework.

4.3.1. Basic Setup

Like Margiotta and Miller (2000), Milbourn (2003), Gayle and Miller (2009), and Coles, Lemmon, and Meschke (2012), the starting point of our model follows Holmstrom and Milgrom (1987)and Gibbons and Murphy (1990). We assume a risk-neutral principal (the board) and a risk-averse agent (the CEO), and assume further that firm performance follows a factor structure consisting of (i) unobserved managerial effort $[a]$, (ii) a common shock that is beyond the manager's control $[c \sim_{iid} N(0, \sigma_c^2)]$, and (iii) a firm-specific idiosyncratic shock $[c \sim_{iid} N(0, \sigma^2)]$

$$p = a + c + \epsilon \tag{2}$$

The empirical counterpart to p in the RPE literature often includes stock returns or accounting performance measures (Antle and Smith 1986; Barro and Barro 1990; Bertrand and Mullainathan 2001; Ittner and Larcker 2002; Aggarwal and Samwick 1999b; Mengistae and Colin Xu 2004; Albuquerque 2009, 2013; Lewellen 2015).

Under linear contracts of the form $w = \alpha + \beta[p - c]$, exponential utility, and a quadratic cost of managerial effort, the manager's problem is given by

$$\max_a e^{-\eta(w - \frac{\kappa}{2}a^2)}, \tag{3}$$

where η is the manager's CARA risk aversion. The manager's optimal effort choice (and expected firm performance) is given by

$$a^* = \frac{\beta}{\kappa}, \tag{4}$$

which is the performance sensitivity of the linear contract scaled by the cost of effort parameter κ.

In this framework, the risk-neutral board's problem is given by

$$\max_{a,\alpha,\beta} \mathbb{E}(p - w) \text{ , s.t.} \tag{5}$$

$$\mathbb{E}[-e^{-\eta\left[w-\frac{\kappa}{2}a^2\right]}] \geq u(\underline{w}) \text{ , and} \tag{5-PC}$$

$$a \in \text{argmax}\ \mathbb{E}[-e^{-\eta\left[w-\frac{\kappa}{2}a^2\right]} \text{ ,} \tag{5-IC}$$

and the optimal relative performance contract is given by

$$w^* = \alpha^* + \beta^*(p - c) \tag{6}$$

The first component of the optimal contract (α^*) is the manager's expected compensation when the firm meets its peers' performance, which depends on the manager's exogenously determined outside option. The second component, $\beta^* = \frac{1}{1+\eta\kappa\sigma^2}$, represents the pay-performance sensitivity portion of the contract. The contract form can also be interpreted within the stated objective of rTSR in Appendix Table III. Without loss of generality, under equation (2), relative returns ($p - c$) is equivalent to the excess returns of the firm ($a + e$), i.e., "shareholder alignment".

Finally, given the optimal contract chosen by the board, the manager's effort can be rewritten as

$$a^* = \mathbb{E}[p] = \frac{1}{\kappa+\eta\kappa^2\sigma^2} \tag{7}$$

Thus, to motivate optimal effort, the principal designs a contract that rewards managers for effort by perfectly eliminating the common shocks in firm performance since they are beyond the manager's control.

The key comparative static for our purposes is the negative effect of the variance in idiosyncratic shocks on performance through managerial effort: $\frac{\partial a^*}{\partial \sigma^2} < 0$ from equation (7). The intuition is that all else equal, higher σ^2 means that a greater proportion of the firm's performance is unpredictable – or explained by idiosyncratic shocks – even after filtering out the common component of performance. Thus, the increased risk in the compensation, coupled with the participation constraint, causes the optimal contract to be less sensitive to performance and thereby reducing the manager's incentives to exert effort. We also obtain the same comparative static under a standard career concerns framework whereby agents have greater incentive to exert effort when the variance of idiosyncratic shocks is lower (Holmstrom 1999). Our results can consistent with either interpretation.

4.3.2. Imperfect Common Shock Filtration

We now depart from the baseline case by introducing imperfect filtering and assuming that the principal (the board) observes the common shock with error,

$$\hat{c} = c + \omega_b \tag{8}$$

where $\omega_b \sim_{iid} N(0, \sigma_b^2)$.[15] Here, lower σ_b^2 represents the greater ability of performance peers or benchmarks to eliminate common shocks, and perfect common shock filtering reduces to the special case where $\sigma_b^2 = 0$.[16]

Under this framework, again assuming linear contracts, exponential utility, and quadratic cost of effort, the manager's optimal effort and expected equilibrium firm performance is given by

$$a^* = \mathbb{E}(p^*) = \frac{1}{\kappa + \eta\kappa^2(\sigma^2 + \sigma_b^2)} \tag{9}$$

Notably, poorer measurement of the common shock (higher σ_b^2) reduces the equilibrium effort level and expected firm performance.

[15] We assume that ω_b has mean zero and is iid which is without loss of generality in the manager's optimization choice of effort. A non-zero mean would enter into the level of the manager's expected compensation. Both Sinkular and Kennedy (2016) and Bennett et al. (2017) find that firms beat the target payout approximately half the time, which suggests that $\mathbb{E}[\omega_b]$ is close to zero.

[16] Note that this formulation produces the same analytical prediction as the original Holmstrom and Milgrom (1987) and Aggarwal and Samwick (1999b) framework where a second signal of firm performance (performance peers/benchmarks) exists with the two signals sharing a correlation ρ. One can think of choosing better peers/benchmarks in the two models as either a decrease in σ_b^2 or an increase in the absolute value of ρ.

The intuition is that measurement errors introduce additional noise into the manager's compensation, and in particular into the performance metric $(p - \hat{c})$ from which the manager's effort level is inferred $(a + \omega_b + \epsilon)$. Thus, as in the baseline case above, the incremental volatility stemming from poorer measurement of the common shock causes compensation for the manager to become riskier which then induces the optimal pay-for-performance sensitivity (β) to be lower and thereby forcing the manager to choose a lower level of effort.

The R^2 results in Table 4.5 can be interpreted in the framework of this model, since there is a one-to-one mapping with the variance in measurement errors. In particular, the time-series return regression $p_t = \delta_1 + \delta_2 \hat{c}_t + \epsilon_t$ yields an R^2 – the squared correlation coefficient ($\rho^2_{p,\hat{c}}$) that can be expressed as a function of the primitives of the model:

$$\hat{R}^2 = \rho^2_{p,\hat{c}} = \frac{\sigma_c^2 \sigma_c^2}{(\sigma_c^2 + \sigma^2)(\sigma_c^2 + \sigma_b^2)} \tag{10}$$

For a given firm – i.e., fixing σ^2 and σ_c^2 – lower σ_b^2 corresponds to higher R^2. Therefore, the results of Table 4.5 imply that SBPs produce lower measurement error variances in the common performance factor relative to firms' chosen performance benchmarks.

In fact, the assessment of peer benchmarking adequacy reported in Table 4.5 can be recast in terms of measurement error variances – ultimately the economic object of interest in our analysis.[17] Under the model assumptions, the following data moments – the variances in prediction errors from peer benchmarks – can identify the measurement error variances up to a scalar constant:

$$Var(p - \hat{c}_{peer}) = \sigma^2_{b,peer} + \sigma^2 \tag{11}$$

Although we cannot identify the magnitude of the measurement error variances, their differences between one set of peer benchmarks and another can be identified – a refinement over the R^2

[17] Dikolli et al. (2012) also model measurement error in the common performance shock in order to understand how various forms of the error structure can bias the standard tests for (implicit) RPE.

measures, which, as shown in equation (10), contain σ_c^2, but other factors as well. Moreover, these sample moments allow us to obtain a lower bound estimate on the proportional improvement between two benchmark candidates.[18]

4.3.3. Estimating Measurement Error Variances and Performance Implications

Table 4.6, panel A, rows 1-3 present simple method of moments parameter estimates of equation (11) for the S&P500 (row 1), firms' chosen performance benchmarks (row 2), and SBPs (row 3), where p represents the monthly stock returns of the base firm. $\hat{c}_{sp500}$, $\hat{c}_{sbp}$, and $\hat{c}_{firm}$ are monthly stock returns of the S&P500 index, the (traffic-weighted-average) returns of firms' SBPs, and the median returns of firms' chosen performance benchmarks, respectively. Note that in this exercise we are implicitly constraining the coefficient on c in equation (2) to equal 1, consistent with how the RP benchmarks are used in practice.

In column 1, the estimated $\sigma_{b,firm}^2 + \sigma^2$ across the whole sample equals 44.489, whereas $\sigma_{b,sbp}^2 + \sigma^2$ equals 40.742 for a statistically significant difference (at the 1% level) of 3.747. On a relative basis, firms' chosen performance benchmarks produce *at least* 9.2% greater variance in measurement errors. Columns 2 and 3 of the same panel recover estimates for the subset of firms that selected specific peers and indexes, respectively. Similar to our findings in Table 4.5, SBPs' out-performance of firms' chosen performance benchmarks is concentrated in the set of firms that choose indexes. Index-based benchmarks' measurement error variances are at least 16.4% greater than SBPs; for firms using specific peers, variances are at least 4.4% greater, both statistically significant at the 1% level. In summary, our findings on the ultimate construct of interest – the

[18] It is easily shown that $\frac{\sigma_{b,firm}^2 + c}{\sigma_{b,sbp}^2 + c} > 1 \Rightarrow \frac{\sigma_{b,firm}^2}{\sigma_{b,sbp}^2} > \frac{\sigma_{b,firm}^2 + c}{\sigma_{b,sbp}^2 + c}$.

Table 4.6. Estimating Measurement Error Variances and Calibrated Performance Implications: Firm Benchmarks vs. S&P500 and SBPs

This table reports method of moments estimates of equations 9 and 11 using pooled firm-month observations. $\sigma^2_{b,sp500}$ is the variance of the measurement error of the common factor using the S&P500 as the relative performance benchmark. $\sigma^2_{b,firm}$ is the variance of the measurement error of the common factor using the firm's chosen performance peers. $\sigma^2_{b,sbp}$ is the variance of the measurement error of the common factor using the firm's search based-peers (Lee et al. 2015). σ^2 is the variance of the firm's idiosyncratic performance where performance is measured via CRSP monthly stock returns, and peer performance is measured at the median of the peer set's returns. κ is the cost of effort parameter in the standard principal-agent model. The estimates are based on the assumption that the manager's CARA risk aversion η=0.625.

The four rows in Panel A report the individual parameter estimates of the measurement error variances. Panel B reports the differences in the measurement error variances of the firms' chosen benchmarks relative to two normative benchmarks: the S&P500 and SBPs. Panel C reports in annualized basis points the performance implications of using the firms' chosen benchmarks relative to the two normative benchmarks described in equation (12).

Column 1 reports estimates for the entire sample using the same sampling criterion as Table 4.5. Column 2 reports estimates for the sub-sample in which the base firm chooses specific firms as the performance benchmark. Column 3 reports estimates for the sub-sample in which the base firm chooses an index (e.g., S&P500, S&P1500) as its performance benchmark. Standard errors are reported in brackets and calculated via the delta method where appropriate. Significance levels are indicated by *, **, *** for 10%, 5%, and 1% respectively.

	All Firms	Specific Peers	Index Peers
	(1)	(2)	(3)
Panel A: Pooled Estimates			
$\sigma^2_{b,sp500} + \sigma^2$	60.197***	63.310***	55.080***
	[1.937]	[2.519]	[3.011]
$\sigma^2_{b,firm} + \sigma^2$	44.489***	41.161***	49.962***
	[1.593]	[1.877]	[2.867]
$\sigma^2_{b,sbp} + \sigma^2$	40.742***	39.418***	42.920***
	[1.367]	[1.712]	[2.269]
κ	0.170***	0.176***	0.162***
	[0.008]	[0.010]	[0.011]
Panel B: Δ Esimates			
$\sigma^2_{b,firm} - \sigma^2_{b,sp500}$	-15.709***	-22.149***	-5.118***
	[0.952]	[1.388]	[1.041]
$\sigma^2_{b,firm} - \sigma^2_{b,sbp}$	3.747***	1.743**	7.042***
	[0.907]	[0.853]	[1.946]

Table 4.6. Estimating Measurement Error Variances and Calibrated Performance Implications: Firm Benchmarks vs. S&P500 and SBPs (Continued)

Panel C: Performance Implications

$\mathbb{E}[p(\sigma^2_{firm}) - p(\sigma^2_{sp500})]$	281.83***	385.73***	97.22***
	[25.03]	[41.29]	[22.12]
$\mathbb{E}[p(\sigma^2_{firm}) - p(\sigma^2_{sbp})]$	-91.73***	-44.41***	-162.62***
	[22.08]	[21.53]	[44.91]
	13,108	8,151	4,957

performance of firms' chosen RP benchmarks in terms of measurement error variances – are empirically and theoretically consistent with the earlier R^2 results of Table 4.5.[19]

Given the greater measurement error variance implicit in the benchmarks chosen by firms, we proceed to quantify the economic implications of benchmark inefficacy in terms of managerial effort and expected firm performance, which can be estimated using the sample analogue of equation (9). In particular, given the manager's risk aversion (η) and effort-cost parameter (κ), the impact of poorer benchmarks on expected performance is given by

$$\mathbb{E}\left[\mathrm{p}\left(\sigma_{sbp}^{2}\right)-p\left(\sigma_{firm}^{2}\right)\right]=\frac{1}{\kappa+\eta\kappa^{2}(\sigma_{b,sbp}^{2}+\sigma^{2})}-\frac{1}{\kappa+\eta\kappa^{2}(\sigma_{b,firm}^{2}+\sigma^{2})} \quad (12)$$

This computation requires identification of the risk-aversion (η) and the effort-cost (κ) parameters; however, with three unknowns (κ, η, $\sigma_{b,firm}^{2}$) and two equations (9 and 11), the model is underidentified. The under-identification of the risk aversion parameter is common in estimating these models (see, e.g., Gayle and Miller 2015); our calibration, thus, borrows accepted ranges of the risk aversion parameter η from the prior literature. Following Haubrich (1994), we consider the range of η between 0.125 and 1.125. Consistent with the model, we also restrict κ >0, since equation (9) has two roots. Table 4.6, Panel A, row 4 reports method of moments estimates of κ, under the assumption that η=0.625, the midpoint of the range considered in Haubrich (1994).

The implications of SBPs' outperformance of firms' chosen RP benchmarks are obtained by applying the method of moments parameter estimates and the assumed risk-aversion parameter to equation (12). Across all firms in the sample, at the midpoint of risk aversion η =0.625, we estimate the counterfactual performance gained (lost) under SBPs (S&P500) to be 91.73 (281.83)

[19] An interesting question is whether idiosyncratic firm performance isolated through relative performance is driven by discount rate or cash flow news. Both cases can support the effort story, depending on one's assumptions about the nature of the manager's effort. In either case, Vuolteenaho (2002) shows that, at the firm level, the majority of the variation in returns is driven by cash flow news.

basis points in annual returns as reported in Table 4.6, Panel C. In other words, the on-average under-performance of firms' selected benchmarks – in terms of explaining the common component of firm performance – implies an economically significant performance effect. These performance implications are again driven by the set of firms that select index-based benchmarks. Interestingly, we find that firms that selected specific peers averted a loss of 385 basis points by not selecting the S&P500.

In un-tabulated robustness results, we also estimate the performance consequences (in annual returns) of counterfactually switching to SBPs using the lower (η=0.125) and upper bound (η=1.125) of the manager's risk-aversion parameter. We find an effect size of 60 basis points corresponding to the lower bound and 153 basis points corresponding to the upper bound for all firms and a range of 106 to 277 for index-based benchmarking firms. Relative to our sample median annual returns, this represents an economically significant 4.4 to 11.3% proportional decline overall, and 7.8 to 20.4% for the sub-sample of index-based benchmarking firms. Moreover, we examine in un-tabulated tests the robustness of the results in Table 4.6 (which is based on a pooled sample estimation to provide sufficient power to estimate κ) based on a firm-by-firm basis. We find similar but slightly larger magnitudes in our estimate of error variances than in the pooled estimates, suggesting that pooling and averaging across firm-specific estimates yield similar conclusions.

4.4. Understanding the Choice of Benchmark Precision

The calibration estimates in the prior section suggest that, in the absence of frictions generated by selecting precise peer benchmarks, the on-average underperformance of firms' selected RP benchmarks – particularly index-based benchmarks – implies performance penalties that are economically large. Our hypothesis is that these economic magnitudes could be rationalized, at least in part, by certain economic frictions, to which we now turn our attention. We begin by extending the baseline model to endogenize the board's problem of benchmark selection. We then use the model's predictions to guide our empirical investigation of plausible economic explanations for the observed underperformance in RP benchmarks.

4.4.1. Comparative Statics of Benchmarking Efficacy σ_b^2

To analyze the possible sources of ineffective benchmarking more formally, we generalize the problems faced by the board and the manager introduced in equations (5) and (8), which assume that the quality of the benchmarking technology available to the board (σ_b^2) is exogenously determined. We now assume instead that improving the benchmark peers' quality (lowering σ_b^2) entails costly effort on the part of the board, and model the cost function as quadratic in peer quality.

The board's optimal selection of a benchmark, characterized by its measurement error variance (σ_b^2), is the solution to the utility maximization problem based on the board's indirect utility function from substituting equations (6) and (7) into equation (5):

$$\sigma_b^{2*} = \underset{\sigma_b^2}{\operatorname{argmax}} f(\sigma_b^2; \theta, \kappa, \sigma^2) = \underset{\sigma_b^2}{\operatorname{argmax}} \frac{1}{2\kappa + 2\kappa^2 \eta(\sigma_b^2 + \sigma^2)} - \underline{w} - \frac{1}{2}\theta\left(\frac{1}{\sigma_b^2}\right)^2 \quad (13)$$

Thus, obtaining a precise estimate for the common component of firm performance (low σ_b^2) is more costly with θ, a cost shifter to capture differential cost of effort or monitoring among boards.

Because the objective function exhibits increasing differences in σ_b^2 with respect to each of the state variables (i.e., $\frac{\partial^2 f}{\partial \sigma_b^2 \partial \theta} > 0$, $\frac{\partial^2 f}{\partial \sigma_b^2 \partial \kappa} > 0$, and $\frac{\partial^2 f}{\partial \sigma_b^2 \partial \sigma^2} > 0$), by Topkis' Theorem (Topkis 1978), the model yields the following three predictions. First, the level of peer precision is decreasing in the board's cost of effort or monitoring ($\frac{\partial \sigma_b^{2*}}{\partial \theta} > 0$). In other words, a board will be more likely to exert effort to search for better benchmarks when board members are more skilled or higher-quality monitors (e.g., less distracted or less captured). Second, the level of peer precision is increasing with the CEO's quality or ability ($\frac{\partial \sigma_b^{2*}}{\partial \kappa} > 0$). Third, the level of peer precision is decreasing with the level of volatility in firm performance ($\frac{\partial \sigma_b^{2*}}{\partial \sigma^2} > 0$). The intuition for the latter two predictions is that boards are more likely to exert effort to produce better benchmarks when the marginal benefits are higher: that is, when managers are more likely to exert effort as a result of better filtering, either because their cost of effort is lower or they are more talented (lower κ) or because their efforts contribute more to firm performance (lower σ^2).[20] In the next section, we test these hypotheses by examining how the characteristics of the CEO, the board, and the firm may explain the observed variation in the quality of RP benchmarks.

4.4.2. Empirical Drivers of Benchmarking Efficacy

To test these hypotheses empirically, we first construct measures of benchmarking adequacy that assess the performance of firms' selected RP benchmarks relative to a normative alternative. Following Section 4.3, we use three candidate benchmarks: (1) the S&P500, (2) firms' own compensation benchmarking peers, and (3) search-based peers. We measure benchmarking adequacy based on the difference in the time-series R^2 between using a firm's chosen RP benchmarks and using each of the three alternatives. The higher the value of these measures, the

[20] This set up also yields the technical result that $\frac{\partial \sigma_b^{2*}}{\partial \eta} > 0$ when $\kappa\eta(\sigma^2 + \sigma_b^2)$>1 and $\frac{\partial \sigma_b^{2*}}{\partial \eta} < 0$ when $\kappa\eta(\sigma^2 + \sigma_b^2)$<1. That is, there is a non-linear relation in the marginal effect of the manager's risk aversion on the quality of peers selected. All else equal, boards choose lower-quality peers when managers with a "high" degree of risk aversion become more risk-averse (i.e., $\eta > \frac{1}{\kappa(\sigma^2+\sigma_b^2)}$); conversely boards choose higher-quality peers when managers with a "low" degree of risk aversion become more risk-averse (i.e., $\eta > \frac{1}{\kappa(\sigma^2+\sigma_b^2)}$). This is the case because the marginal benefit of improving benchmark quality is non-linear in the manager's risk aversion. We do not emphasize this comparative static since we cannot empirically observe or proxy for variations in managers' risk aversion.

more efficacious are a firm's chosen benchmarks. To ease interpretation, all of the adequacy measures are standardized to have zero mean and unit variance.

Table 4.7 reports the results of OLS regressions of the benchmarking-adequacy measures on an indicator for choosing index-based benchmarks as well as a set of CEO, board, and firm characteristics intended to capture variation in the model's primitives. We include three proxies for CEO characteristics – *Log of CEO Pay*, *CEO Tenure*, and *CEO Age* – that could capture managerial talent or cost of effort; five proxies for board characteristics – *% Busy Directors, Board Size, Director Workload, % Age 65+ Directors*, and *Staggered Board* – that could capture board quality or cost of monitoring; and four measures of firm-level characteristics – *Log Market Cap, Return Volatility, Book-to-Market*, and *Dual-Class Shares* – that could capture the relative importance of idiosyncratic shocks in firm performance and other governance characteristics. The specifics of variable construction and descriptive statistics are explained in Appendix Table IV.

Under the model's predictions, we expect positive associations between benchmark adequacy and *Log of CEO Pay*, *CEO Tenure*, and *CEO Age*, which can be interpreted as measures of CEO skill or ability. We expect negative associations with *% Busy Directors*, *Director Workload*, *Board Size*, and *% Age 65+ Directors*, which can be interpreted as measures of board distractedness (in the case of busy directors and director workload) and of incentives/proclivities to free ride (in the case of board size and retirement-aged directors), which approximate the board's cost of effort.[21] Similarly, we expect to see negative associations with *Staggered Board*, and *Dual-Class Shares*, both director-entrenching governance mechanisms that can increase directors' cost of monitoring or the extent to which board members are captured. We also expect to see positive

[21] One consultant explained that at certain firms, older board members find it easier to understand stock indexes as relative performance benchmarks.

Table 4.7. Explaining the Adequacy of Benchmark Selection

This table reports results from OLS regressions of the adequacy of a firm's choice of relative performance benchmark on CEO, board of directors, and firm characteristics. We consider three measures of relative performance benchmark adequacy. They are defined as the performance of the relative performance benchmark in terms of times-series R^2 relative to that of the S&P500 index (R^2_{sp500}), firms' own compensation benchmarking peer firms (R^2_{cb}), and search-based peers (R^2_{sbp}), in columns 1-3 respectively. All three measures are normalized to have a zero mean and unit variance. Observations are at the annual firm-benchmark level and all variables are defined in Table A3. All specifications include time, industry, and compensation consulting firm fixed effects. Industry-fixed effects use 2-digit Global Industry Classification Standard definitions. The reported p-values refer to joint F-tests of the significance of the compensation consultant fixed effects. Standard errors are clustered at the firm level and reported below the point estimates in brackets. Significance levels are indicated by *, **, *** for 10%, 5%, and 1% respectively.

	$R^2_{rp} - R^2_{sp500}$	$R^2_{rp} - R^2_{cb}$	$R^2_{rp} - R^2_{sbp}$
	(1)	(2)	(3)
Index	-0.646***	-0.548***	-0.515***
	[0.090]	[0.150]	[0.123]
<u>CEO Characteristics</u>			
Log CEO Pay	0.027	0.083	-0.026
	[0.071]	[0.094]	[0.090]
CEO Tenure	0.029***	-0.003	-0.002
	[0.006]	[0.011]	[0.012]
CEO Age	-0.001	0.013	0.014
	[0.008]	[0.011]	[0.009]
<u>Board and Firm Characteristics</u>			
% Busy Directors	-0.031	-0.241	-0.129
	[0.581]	[0.842]	[0.897]
Board Size	-0.027*	-0.028	0.004
	[0.016]	[0.024]	[0.028]
Director Workload	0.037	0.037	0.040
	[0.100]	[0.129]	[0.129]
% Age 65+ Directors	0.122	-0.154	-0.310
	[0.218]	[0.343]	[0.235]
Log Market Cap	0.005	0.046	0.041
	[0.041]	[0.077]	[0.066]
Return Volatility	0.106	0.185	-2.454**
	[0.734]	[0.944]	[0.958]
Book-to-Market	0.127	-0.123	-0.051
	[0.122]	[0.156]	[0.142]
Staggered Board	-0.133*	-0.138	0.033
	[0.078]	[0.116]	[0.094]
Dual-Class Shares	-0.114	-0.053	-0.535
	[0.196]	[0.267]	[0.361]

Table 4.7. Explaining the Adequacy of Benchmark Selection (Continued)

Industry Characteristics			
Census-based HHI Index	0.379	1.074	-2.802***
	[1.208]	[1.744]	[1.065]
Time FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Comp Consultant FE	Yes	Yes	Yes
p-value of F-test	0.4879	0.1545	0.3603
Observations	1,175	1,074	1,175
Adjusted R-squared	0.5172	0.1077	0.1320

(negative) associations with *Return Volatility* and *Book-to-Market* (*Log Market Cap*), since higher (lower) values in these variables reflect greater fundamental volatility.

Table 4.7, columns 1-3, report the OLS results using the R^2-based measure of benchmarking adequacy with standard errors clustered at the firm level as well as time, industry, and compensation-consulting-firm fixed effects. Across all specifications of benchmarking adequacy, the choice of an index is associated with a significant decline (at the 1% level) in the relative performance of firms' chosen benchmarks, consistent with the findings in Tables 4.5 and 4.6. The effects we document are not only statistically significant but also economically large: the choice of an index is associated with a -0.52 to -0.65 standard deviation decline in efficacy depending on the specification.[22] Interestingly, we find no joint significance in the compensation consulting firm fixed effects, suggesting that, conditional on the selection of index-based or specific peers, compensation consultants do not systematically differ from each other in terms of the quality of peers chosen. Indeed, among the explanatory variables we consider, the choice of an index is the only one that systematically and significantly explains the quality of the chosen peers.

Given this finding, we further test the model's predictions by investigating the economic determinants for selecting index-based benchmarks, an indicator of low-quality benchmarking. Table 4.8 reports the marginal effects from a probit regression of an indicator for having chosen an index-based benchmark on the same set of CEO, board, and firm characteristics as in Table 4.7.[23]

---

[22] We obtain qualitatively similar results using the difference in measurement error variances as the dependent variable. The choice of an index is associated (significantly at the 1% level) with a -0.41 and -0.33 standard deviation decline in $(\sigma^2_{sp500} - \sigma^2_{rp})$ and $\left(\sigma^2_{sbp} - \sigma^2_{rp}\right)$ respectively. This is not surprising, because the differences in R^2s are equivalent to differences in measurement error variances up to a scalar constant, i.e., $\left(R^2_{rp} - R^2_{sbp}\right) = \frac{(\sigma^2_{rp} - \sigma^2_{sbp})\sigma^2_c\sigma^2_c}{(\sigma^2_c\sigma^2)}$.

[23] For the tests in Tables 4.7, 4.8, and 4.9, we obtain firm-level characteristics of the 330 unique firms for fiscal years 2006-2013. Unlike the analyses in Table 4.7, those in Tables 4.8 and 4.9 are not based on our benchmark efficacy measures derived from R^2 measures (in Table 4.5) and measurement error variance measures (in Table 4.6). Instead, our main variable of interest is *Index* – an indicator variable equal to 1 if the firm-year involves an index-based

The results of Table 4.8 suggest that the choice of index-based benchmarks is associated with board-level governance weaknesses. In the first three columns, which vary by which fixed effects are included (i.e., time, industry, or compensation consultant), the likelihood of choosing an index-based benchmark is associated positively and both economically and statistically significantly (at the 5% level) with the size of the board and with directors' workload. Interpreting the marginal effects of column 3, a one standard deviation increase in *Board Size, Director Workload*, and *% Age 65+ Directors* is associated with 6.2%, 7.4%, and 8.4% higher likelihoods of choosing index-based benchmarks. Relative to a baseline likelihood of 34% for choosing index-based benchmarks, these estimates represent proportional increases of 18.2%, 21.7%, and 24.7% respectively.

We do not find evidence that less precision in RP benchmarks is explained by lower CEO ability or effort. In particular, we find no evidence of a negative and significant coefficient on *Log CEO Pay, CEO Tenure*, or *CEO Age.* In all three specifications, in fact, we find a positive and significant coefficient (at the 10% level) on *Log CEO Pay*. For example, column 3's point estimates suggest that a one standard deviation increase in *Log CEO Pay* increases the likelihood of choosing index-based benchmarks by 8.1%, or a proportional increase of 23.8% relative to the baseline likelihood. One interpretation of this result is that, conditional on the controls, higher CEO pay reflects *excess* pay and is thus an outcome of board-level governance weaknesses. If so, the observed positive and statistically and economically significant coefficient on *Log CEO Pay* is consistent with the model's prediction that less precision in the choice of benchmarks could result from lower board monitoring quality.

benchmark, and 0 otherwise. Accordingly, these tests are not subject to the additional data filters that are required for our R^2 analyses in Table 4.5, and comprise a larger number of firm-year observations than used in Tables 4.5 and 4.6.

Table 4.8. Explaining Selection of Benchmark Types

This table reports the marginal effects, evaluated at the sample mean for continuous variables and at zero for indicator variables, from probit regressions of the firm's choice of an index as its relative performance benchmark on CEO, board of directors, and firm characteristics. Observations are at the annual firm-benchmark level and all variables are defined in Table 4.11. All specifications include time fixed effects. Column 2 includes industry-fixed effects using the 2-digit Global Industry Classification Standard definitions. Column 3 includes compensation consultant fixed effects. Columns 4 and 5 report sub-sample results conditioned on whether the compensation consulting firm has a preference for an index benchmark versus specific peers. Index preference is determined based on whether an individual compensation consulting firm's fixed effects is above the median of all compensation consultant fixed effects in column 3. Compensation consulting firms in column 4 (5) include: Exequity, Frederick Cook, Pay Governance and Towers Watson (Aon Hewitt, Compensation Advisory, Compensia, Deloitte, Hay Group, Mercer, Meridan, Pearl Meyer, and Semler Brossy). The reported *p*-value refer to joint F-tests of the significance of the compensation consultant fixed effects. Standard errors are clustered at the firm level and reported below the point estimates in brackets. Significance levels are indicated by *, **, *** for 10%, 5%, and 1% respectively.

	Pr(Index) = 1				
	(1)	(2)	(3)	(4)	(5)
CEO Characteristics					
Log CEO Pay	0.140***	0.091**	0.081*	0.061	0.022
	;0.052'	[0.046]	[0.049]	[0.065]	[0.028]
CEO Tenure	0.008	0.007	0.007	0.009	0.001
	[0.005]	[0.005]	[0.005]	[0.007]	[0.001]
CEO Age	-0.001	0.005	0.006	0.011	0.000
	[0.005]	[0.006]	[0.006]	[0.007]	[0.002]
Board and Firm Characteristics					
% Busy Directors	0.365	0.550	0.541	0.496	0.005
	[0.491]	[0.470]	[0.464]	[0.735]	[0.106]
Board Size	0.033**	0.027*	0.030**	0.013	0.008
	[0.015]	[0.014]	[0.013]	[0.022]	[0.008]
Director Workload	0.202***	0.194***	0.213**	0.225**	0.043
	[0.067]	[0.072]	[0.074]	[0.103]	[0.042]
% Age 65+ Directors	0.164	0.262*	0.263*	0.236	0.078
	[0.151]	[0.149]	[0.153]	[0.204]	[0.086]
Log Market Cap	-0.063*	-0.014	-0.018	-0.007	-0.002
	[0.033]	[0.033]	[0.032]	[0.046]	[0.008]
Return Volatility	-0.254	0.203	0.399	1.087	0.069
	[0.553]	[0.559]	[0.533]	[0.872]	[0.149]
Book-to-Market	-0.135	-0.036	-0.005	-0.002	-0.005
	[0.088]	[0.094]	[0.089]	[0.130]	[0.020]
Staggered Board	0.041	0.054	0.047	-0.035	0.068
	[0.061]	[0.063]	[0.060]	[0.085]	[0.065]
Dual-Class Shares	-0.032	-0.012	-0.033	-0.013	0.017
	[0.118]	[0.119]	[0.102]	[0.150]	[0.053]

Table 4.8. Explaining Selection of Benchmark Types (Continued)

Industry Characteristics					
Census-based HHI Index	0.287	-0.100	-0.160	-0.736	0.066
	[0.637]	[0.625]	[0.576]	[1.068]	[0.116]
Time FE	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Yes	Yes	Yes
Comp Consultant FE	No	No	Yes	No	No
p-value of F-test			0.0000		
Observations	1,175	1,175	1,162	696	476
Pseudo R-squared	0.0580	0.1835	0.2583	0.2079	0.3888
Index Preference	All	All	All	Yes	No

Further, we do not find consistent evidence that lower precision in RP benchmarks is explained by higher volatility in firm performance. The coefficients on *Return Volatility* and *Book-to-Market* are not significantly positive in all three specifications. We do find, in column 1 (with no fixed effects), negative coefficients on *Log Market Cap* that are significant at the 10% level. However, the significance does not survive the inclusion of additional fixed effects in columns 2 and 3.

Finally, we find that compensation-consultant fixed effects are important in explaining the choice of index-based benchmarks. In column 3, the inclusion of these fixed effects increases the R^2 of the regression specification by over 40%; moreover, the results of an F-test show that these fixed effects are jointly significant (at the 1% level). That is, after for controlling for firm-level covariates (i.e., characteristics that can explain firms' selection of particular compensation consultants), compensation consultants exhibit systematic tendencies to recommend indexes or specific peers.

These results are consistent with our interviews with practitioners. We conducted interviews with four compensation consultants and two compensation experts involved in determining CEO compensation packages at their respective corporations. While these interviewees all acknowledged that a primary reason for using rTSR in performance contracts is to remove market or industry-level noise from performance, they differed in their preferences for index versus specific relative performance peer benchmarks. Certain consultants have built capabilities to better identify ideal specific-peer benchmarks; others choose indexes by default. Thus, a board that does not carefully scrutinize the recommendations of a compensation consultant will ultimately evaluate executives based on the systematic component of performance to some degree.

If our conjecture is correct, we would expect the effects of monitoring quality to be especially acute at firms whose compensation consulting firms have a penchant for index benchmarks. The argument proceeds as follows. Low-quality boards follow the default recommendation of the hired consultant; the maintained assumption is that on average the right method is use of specific peers. Therefore, firms with low-quality boards that hire index-preferring consultants will on average choose indexes and firms with high-quality boards that hire index-preferring consultants will, after careful scrutiny, on average choose specific peers; this pattern would result in a negative association between monitoring quality and the choice of an index. However, such a pattern would not characterize firms that hire consultants with a preference for specific peers, because both high-quality- and low-quality-monitoring firms would choose specific peers, albeit for different reasons, leading to pooling instead of separation.

When we partition the sample by the index preference of the compensation consulting firms, we find that the effects of board distractedness are concentrated in the set of firms whose compensation consultants prefer index-based benchmarks. The coefficients for *Busy Directors, Board Size, Director Workload*, and *% Age 65+ Directors* are all larger in column 4 (estimated on the subsample of firms that hired index-preferring consultants) than in column 5 (estimated on the subsample of firms that hired specific-peer-preferring consultants). Moreover, *Director Workload* is statistically significant at the 1% level in column 4 but none of the firm-level covariates is statistically significant in column 5, consistent with pooling. These results are consistent with Cai, Kini, and Williams (2016), which finds significant consultant-style effects in the determination of executive compensation levels and mix only at the set of firms with weak governance mechanisms. Our work highlights how compensation consultants' styles interact with firms' governance quality in the selection of relative-performance benchmarks.

Overall, we do not find consistent evidence that volatility in firm performance or CEO talent and skill explains the observed variation in benchmark adequacy. Instead, the evidence of Tables 4.7 and 4.8, along with our interviews, suggests that the friction that may explain the observed poor performance of benchmarks is associated with the quality of board monitoring.

4.4.3. Alternative Theories of Variation in Benchmarking Quality

We discuss several plausible alternative theories, external to our model, that could explain the observed variation in benchmarking quality (i.e., the selection of broad indexes).

4.4.3.1. Firms' Market Power

Choosing closely related industry peers may not always appropriately capture the exogenous component of performance. Incentivizing on relative-performance measures is predicated on the assumption that the firm's actions do not affect the performance of its benchmarks; but the actions of firms operating in oligopolistic industries may impact the performance of their industry competitors, invalidating their use for purposes of common shock filtration. Under such conditions, RPE contracts may incentivize managers to sabotage their competitors rather than improve their own firm's performance; as a result, it may be optimal to reward CEOs for the common shock (Janakiraman et al. 1992; Aggarwal and Samwick 1999a) and to encourage firms to soften product-market competition. Thus, one prediction of such a theory is that firms with greater market power are less likely to select direct competitors and more likely to adopt broad indexes in order to eliminate market-level volatility from their performance.

Our empirical results suggest that such a theory does not explain our findings. In Tables 4.7 and 4.8, we do not find evidence that larger firms are more likely to select poorer-performing peers or to select broad indexes. However, firm size may not be a good empirical measure of market power. Thus, in both tables, we also include a measure of SIC-based industry concentration (*Census-based HHI Index*).[24] In Table 4.7, we do find some limited evidence that firms in more concentrated industries on average select poorer-tracking rTSR benchmarks, conditional on the choice of index or specific peers. In economic magnitudes, a one-standard-deviation change in the HHI index (column 3) leads to a -0.11 standard deviation reduction in the adequacy of benchmarks relative to search-based peers ($R^2_{rp} - R^2_{sbp}$). The effect size is approximately one-quarter of the magnitude of the effect of choosing an index. However, in Table 4.8 we do not find evidence that the choice of an index – the primary driver of the variation in benchmark adequacy – is related to industry concentration. Together, these results do not support the hypothesis that variation in firms' market power drives the observed variation in benchmarking quality between indexes and specific peers.

4.4.3.2. Gameability

Index-based benchmarks may reflect ex-ante rather than ex-post efficiency concerns. For example, in our interviews, one compensation consulting firm offered the possible rationalization that "[w]ith the influence that proxy advisor like the ISS carry, many companies are concerned that a custom group of peers will be misconstrued by proxy advisors as cherrypicking." The presumption is that compensation packages tied to index benchmarks are more difficult to manipulate ex-post. We find this argument uncompelling for several reasons. First, most performance contracts in our sample provide payouts that are linearly interpolated from the threshold targets. Thus, there is less incentive to manipulate the performance metric due to the

[24] Following Ali, Klasa and Yeung (2008) and Keil (2017), we use a census-based Herfindahl-Hirschman Index of SIC industries obtained from Jan Keil's website: https://sites.google.com/site/drjankeil/data.

absence of discontinuous payoffs.[25] Supporting this view, Bennett et al. (2017) finds no evidence of asymmetry in firms' propensity to beat or miss a relative performance target. Second, cherrypicking requires that either the CEO or the board can forecast the returns of peer firms or indexes, since benchmarks are formulated ex-ante prior to the realization of performance.[26] Consistent with this argument, Bizjak et al. (2016) finds that boards' selection of performance peers does not affect the level of pay in an economically significant manner.

It is possible that firms are concerned that the selection of specific peers in relative-performance contracts may provide the *appearance* of gaming. For example, if firms with characteristics associated with poor governance are also more sensitive to the external perception of poor governance, they may prefer index benchmarks. Our empirical evidence is not consistent with this alternative, however. Under such a hypothesis, one would expect governance characteristics to be associated with the choice of index-based benchmarks regardless of the compensation consultant's systemic tendencies. Instead, the results in columns 4 and 5 of Table 4.8 find the association between governance weaknesses and the selection of index-based benchmark to be significant, both statistically and economically, only for the subset of firms hiring index-preferring compensation consultants.

[25] One compensation consultant, during our interview, suggested that such linear interpolations are the norm. We verified this by collecting a random sample of 20 contracts from our main sample (across both specific peer and index benchmarking firms) and found two standard contract forms, both of which are effectively linear in rTSR and payout in shares. Three of the 20 contracts are explicitly linear in rTSR. The remaining 17 are linear in the firm's TSR as a percentile of benchmark peers' stock-return distribution. (In untabulated simulations, we find that linearity in the percentiles of peers' TSR distribution also implies linearity in rTSR.) We infer linearity from firms' proxy-statement disclosures of "linear interpolation" in performance between pre-specified target levels.

[26] Morse et al. (2011) shows that, prior to 2006, ex-post cherrypicking of performance metrics was prevalent because contracts were not disclosed ex-ante.

4.4.3.3. Peer Selection for Inter-Firm Tournaments

The choice of performance peers could also serve to motivate CEOs via an implicit inter-firm tournament (Hvide 2002). However, such a selection should also, at least theoretically, be consistent with the informativeness principle. Given the large number of firms included in the index benchmarks we observe (e.g. S&P500, S&P1500, Russell 3000), it is unlikely that they are characterized by some type of "optimal tournament" involving hundreds or thousands of heterogeneous firms. In particular, Lazear and Rosen (1981) show that absent handicaps (which we do not observe in the performance contracts), heterogeneity among firms, which should be increasing in the peer set size, decreases the effectiveness of tournaments. Furthermore, the tournament mechanism requires both agents to be aware that they are competing in the same tournament, i.e., to be each other's mutually chosen peers. However, as noted in De Angelis and Grinstein (2016) and Shin (2016), a significant fraction of compensation and performance peers are one-sided (i.e., not mutually designated).

4.4.3.4. Peer Selection for Aspiration

Another possibility is that the choice of performance peers is aspirational (Francis, Hasan, Mani, and Ye 2016; Hemmer 2015). Under such a theory, the selection of peer benchmarks would push managers at the firm to generate performance commensurate with or superior to that of well-performing firms.[27] This reasoning suggests a rationale for boards' choice of aspirational firms as compensation-benchmarking peers (Hemmer 2015), but its applicability to the selection of an index benchmark in lieu of specific peers in relative-performance contracts is less obvious.

[27] Other models, such as that of Hayes and Schaefer (2009), argue that when there is asymmetric information about the match surplus between the manager and the firm, and when boards care about short-run perceptions of firm value, boards may inflate the CEO's wage as a costly signal of the match surplus (Scharfstein and Stein 1990).

More broadly, we do not believe that aspiration is the primary motivation behind the determination of performance peers. For example, the aspirational theory does not predict that we would observe a return-regression slope of 1. The findings in Table 4.4 are consistent with RP peers chosen to eliminate common shocks. Similarly, such an explanation is also inconsistent with the evidence in Table 4.8, which suggests that the primary objective of rTSR (in the eyes of compensation consultants) is to provide a measure of firm-specific performance or shareholder value improvement that removes the effect of common shocks, rather than as a measure that motivates performance via aspiration.

4.4.3.5. Managers' Abilities to Self-Insure

Another possible explanation, offered by Garvey and Milbourn (2003) is that managers with greater ability to self-insure against the common factor benefit less from better benchmarks. If so, it is possible that selection of index-based benchmarks on the part of certain firms may reflect non-risk-sharing motivations. However, our empirical results are not consistent with these predictions. In particular, we do not find significantly positive associations (at the 10% level) between *CEO Age*, a common proxy for the ability to self-insure, and the choice of index-based benchmarks in Table 4.8; the coefficients are also weak in economic magnitudes.

4.4.3.6. Persistence of Common Shocks

It is also possible that perfect filtration of common shocks is not optimal under other circumstances. The dynamic agency models of Hoffmann and Pfeil (2010) and DeMarzo, Fishman, He, and Wang (2012) predict that boards will imperfectly filter observable and *persistent* common shocks to performance. This prediction implies that the difference in efficacy between firms that use specific peers and those that use indexes as benchmarks can be driven by greater common-shock persistence in the latter group. However, in untabulated results, we do not find evidence consistent with this theory. Using SBPs to proxy for the common factor, we do not find a difference in the persistence of the common factor between firms that benchmark against specific peers and those that benchmark against indexes.

4.4.3.7. Implementation Costs

It is also possible that implementation costs of computing rTSR are lower conditional on selecting an index rather than specific peers. One compensation consultant remarked, however, that the typical relative-performance incentive contract requires tracking of the individual performance of firms within an index, not just the overall performance of the index itself. Thus, from an implementation perspective, computing rTSR using index-based benchmarks is in fact costlier than using a narrower set of specific peers.

Overall, our analyses suggest that compensation-consultant tendencies and governance-related frictions best explain empirical patterns in benchmark inadequacy and in the choice of index-based benchmarks. Boards of directors that do not carefully scrutinize compensation consultants' recommendations may result in evaluations of executives based, to some degree, on the systematic component of performance. An implication of this finding is that the properties of performance metrics chosen by boards to evaluate managers could be novel indicators of governance quality that is not subsumed by existing metrics familiar in the literature.

4.5. Reduced-Form Performance Implications in the Cross-Section

We conclude our analysis by seeking to determine whether there are observed realized operating performance differences between firms that choose index-based benchmarks and those that choose specific peers, as predicted in equation (7). Table 4.9 reports reduced-form regressions of firm performance, in terms of ROA, on the choice of index-based benchmarks and the set of CEO, board, and firm characteristics examined in our prior analyses. The three specifications differ based on the fixed effects included. Consistent with our calibration exercise, we find evidence in each of the three specifications that the choice of index-based benchmarks is associated with worse operating performance. Interpreting column 3, which includes year-, industry-, and compensation-consultant fixed effects, the choice of index-based benchmarks is associated with a 70 basis point decline in ROA, which is statistically significant at the 5% level. Relative to our sample median ROA of 490 basis points, this is a 14.3% proportional decline.[28]

As in the prior analyses, we test for the joint significance of compensation consultant fixed effects and fail to reject the null of no compensation-consultant effects at the 10% level in the ROA regression of column 3. This alleviates the potential concern that unobservables correlated with performance drive firms' selection of compensation consultants. Thus, the joint significance of compensation consultant fixed effects in the selection regression of Table 4.8 probably reflects compensation consultants' systematic tendencies. We also find that, consistent with poor relative-performance metrics being associated with poor performance, among the 26 firms that switched benchmark types during our sample period (from index to specific peers or vice-versa), 19 (or 73%) switched from index to specific peer benchmarks. Based on a binomial test, we reject the null of equal likelihood against the alternative that benchmark-type switchers are more likely to have

[28] In untabulated tests, we also find similar results using annual stock returns.

Table 4.9. Performance Consequences of Inadequate Benchmarks

This table reports OLS regressions of firms' ROA on an indicator of having chosen an index as the relative performance benchmark along with CEO, board of directors, and firm characteristics. Observations are at the annual firm-benchmark level and all variables are defined in Table A3. All specifications include time fixed effects. Column 2 includes industry-fixed effects using the 2-digit Global Industry Classification Standard definitions. Column 3 includes compensation consultant fixed effects. The reported *p*-values refer to joint F-tests of the significance of the compensation consultant fixed effects. Standard errors are clustered at the firm level and reported below the point estimates in brackets. Significance levels are indicated by *, **, *** for 10%, 5%, and 1% respectively.

	ROA		
	(1)	(2)	(3)
Index	-0.006	-0.008**	-0.007*
	[0.004]	[0.004]	[0.004]
CEO Characteristics			
Log CEO Pay	-0.003	-0.008*	-0.008*
	[0.004]	[0.004]	[0.004]
CEO Tenure	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]
CEO Age	0.001	0.001**	0.001**
	[0.000]	[0.000]	[0.000]
Board and Firm Characteristics			
% Busy Directors	-0.006	-0.010	-0.011
	[0.035]	[0.034]	[0.034]
Board Size	-0.004***	-0.003***	-0.003***
	[0.001]	[0.001]	[0.001]
Director Workload	-0.008	-0.004	-0.003]
	[0.005]	[0.005]	[0.005]
% Age 65+ Directors	0.003	0.007	0.004
	[0.12]	[0.011]	[0.011]
Log Market Cap	0.012***	0.011***	0.011***
	[0.003]	[0.003]	[0.003]
Return Volatility	-0.128**	-0.235***	-0.228***
	[0.052]	[0.056]	[0.056]
Book-to-Market	-0.061***	-0.051***	-0.053***
	[0.009]	[0.008]	[0.008]
Staggered Board	0.007*	0.005	0.006*
	[0.004]	[0.003]	[0.003]
Dual-Class Shares	0.023**	0.015	0.013
	[0.010]	[0.010]	[0.010]

Table 4.9. Performance Consequences of Inadequate Benchmarks (Continued)

Industry Characteristics			
Census-based HHI Index	0.028	-0.004	-0.003
	[0.037]	[0.035]	[0.036]
Time FE	Yes	Yes	Yes
Industry FE	No	Yes	Yes
Comp Consultant FE	No	No	Yes
p-value of F-test			0.5355
Observations	1,175	1,175	1,175
Adjusted R-squared	0.2993	0.3519	0.3529

initially chosen an index (p-value of 0.014). Moreover, consistent with our main cross-sectional results, we find in untabulated results that (index-based) benchmarks performed worse in measurement error variances than SBPs prior to the switch to specific peers, which perform nearly equally to SBPs after the switch.

Together, these results can be interpreted in one of two ways. One view is that these findings suggest that the selection of relative performance benchmarks is important for the inducement of executive effort, and the inability to capture systematic shocks has significant performance implications for firms. An alternative interpretation of our findings is that the quality of relative performance benchmarks provides incremental information about the quality of a firm's board and governance, which affects firm performance.

4.6. Conclusion

Market participants have increasingly looked to relative performance metrics such as rTSR to evaluate the performance of firms and managers. Such attention has coincided with a growing trend toward tying executive performance-based compensation contracts to rTSR. However, a central challenge to the application of rTSR persists: the selection of peers to measure and filter the systematic component of performance.

This paper tests the extent to which boards' choices of rTSR measures evaluate managers on the basis of the unsystematic component of TSR, following the informativeness principle. This is an objective that many compensation consultants invoke to justify the use, and increasing popularity, of rTSR. Consistent with this objective, we find that over half of firms that tie CEO performance-based contracts to rTSR – those that choose specific peers as benchmarks – do a remarkable job of measuring and filtering for systematic risk in TSR. This finding may be particularly surprising in light of the prevailing view that the executive pay-setting and evaluation process is broken and compromised by powerful executives (Bebchuk et al. 2011; Morse et al. 2011). However, we find that nearly half of such firms – those that choose index-based RP benchmarks – use rTSR measures that contain a significant amount of systematic risk and thus do not abide by the informativeness principle. These results provide more direct and novel evidence – made possible by the SEC's 2006 compensation disclosure reform – on the longstanding question of whether and to what extent boards evaluate managers on the basis of systematic or non-systematic performance (Antle and Smith 1986).

We also examine the frictions that could rationalize the observed benchmarking inefficacy, and show that a firm's choice of index-based benchmarks, in lieu of specific peers, is associated with its compensation consultant's systematic tendencies and with board-level monitoring weaknesses. Benchmarking inefficacy is also associated with lower realized annual ROA. Our analyses suggest a simple fix for firms that choose index-based benchmarks: their chosen rTSR measures could be significantly improved by using their compensation-benchmarking peers as measures of systematic risk. Indeed, we find that the vast majority of firms that altered their performance incentive plan structure during our sample period switched from index-based to specific peers, resulting in improvements to their measurements of systematic risk. Jointly, these findings provide new evidence on the role that compensation consultants play in the managerial evaluation process, and suggest that the properties of performance metrics chosen by boards to evaluate executives can serve as novel indicators of board-level governance quality.

Although our findings are restricted to a sample of firms that tie CEOs' performance-based contracts to rTSR, our analysis may also apply to firms whose executives have implicit rTSR-based incentives. For example, if shareholders, board members, or the executive labor market evaluate managers' competence at least partly on the basis of rTSR, managers' reputational, career, or prestige concerns could be tied to such relative-performance metrics. We believe that distinguishing the role of such implicit incentives from formal incentives for understanding the growing use of relative-performance metrics represent a challenging but promising avenue for future research.

BIBLIOGRAPHY

Abernethy, M. and J. Bouwens, C. Hofmann, and L. van Lent, Laurence. 2015. “Social Norms, Agents’ Choices and Incentive Contract Design” *SSRN Working Paper* 2067040

Abernethy, M., J. Bouwens, and L. van Lent. 2013. "The role of performance measures in the intertemporal decisions of business unit managers." *Contemporary Accounting Research* 30.3: 925-961.

Abernethy, M., and P. Brownell. 1997. "Management control systems in research and development organizations: the role of accounting, behavior and personnel controls." *Accounting, Organizations and Society* 22.3-4: 233-248.

Abernethy, M., H. Dekker, and A. Schulz. 2015. "Are employee selection and incentive contracts complements or substitutes?." *Journal of Accounting Research* 53.4: 633-668.

Aggarwal, R., and A. Samwick. 1999a. "Executive compensation, strategic competition, and relative performance evaluation: Theory and evidence." *The Journal of Finance* 54.6: 1999-2043.

Aggarwal, R., and A. Samwick. 1999b. "The other side of the trade-off: The impact of risk on executive compensation." *Journal of Political Economy* 107.1: 65-105.

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

Aghion, P., Bloom, N., Lucking, B., Sadun, R. and Van Reenen, J., 2017. Turbulence, Firm Decentralization and Growth in Bad Times (No. w23354). *National Bureau of Economic Research.*

Ai, C., and E. Norton. 2003. "Interaction terms in logit and probit models." *Economics Letters* 80: 123-129.

Akerlof, G., and R. Kranton. 2000. “Economics and Identity.” *The Quarterly Journal of Economics*
115: 715–54.

Akerlof, G., and R. Kranton. 2005. "Identity and the Economics of Organizations." *The Journal of Economic Perspectives* 19.1: 9-32.

Albuquerque, A. 2009. "Peer firms in relative performance evaluation." *Journal of Accounting and Economics* 48.1: 69-89.

Albuquerque, A. 2013. "Do growth-option firms use less relative performance evaluation?." *The Accounting Review* 89.1: 27-60.

Ali, A., S. Klasa, and E. Yeung. 2008. "The limitations of industry concentration measures constructed with Compustat data: Implications for finance research." *The Review of Financial Studies* 22.10: 3839-3871.

American Institute of Certified Public Accountants (AICPA). 1978. The Commission on Auditors Responsibilities: Report, Conclusions and Recommendations. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). 1992. Statement of Position Regarding Mandatory Rotation of Audit Firms of Publicly Held Companies. New York, NY: AICPA.

Anand, J., and H. Singh. 1997. "Asset redeployment, acquisitions and corporate strategy in declining industries." *Strategic Management Journal*: 99-118.

Angrist, J., and J. Pischke. 2009. "Mostly harmless econometrics: An empiricist's companion." Princeton NJ: Princeton University Press 2009.

Antle, R., and A. Smith. 1986. "An empirical investigation of the relative performance evaluation of corporate executives." *Journal of Accounting Research*: 1-39.

Aranda, C., J. Arellano, and A. Davila. 2014. "Ratcheting and the role of relative target setting." *The Accounting Review* 89.4: 1197-1226.

Armenakis, A., and A. Bedeian. 1999. "Organizational change: A review of theory and research in the 1990s." *Journal of Management* 25.3: 293-315.

Arya, A., and B. Mittendorf. 2004. Using job rotation to extract employee information. *Journal of Law, Economics, and Organization* 20: 400-414.

Ashford, S., C. Lee, and P. Bobko. 1989. "Content, causes, consequences of job insecurity: A theory-based measure and substantive test." *Academy of Management Journal* 32: 803-829.

Ashkanasy, Neal M., and Scott Holmes. 1995. "Perceptions of organizational ideology following merger: A longitudinal study of merging accounting firms." *Accounting, Organizations and Society* 20.1: 19-34.

Baiman, S., D. Larcker, and M. Rajan. 1995. "Organizational design for business units." *Journal of Accounting Research*: 205-229.

Baiman, S., and M. Rajan. 1995. "The informational advantages of discretionary bonus schemes." *The Accounting Review*: 557-579.

Baker, G. 2000. "The use of performance measures in incentive contracting." *American Economic Review* 90.2: 415-420.

Baker, G., R. Gibbons, and K. Murphy. 1994. "Subjective performance measures in optimal incentive contracts." *The Quarterly Journal of Economics* 109.4: 1125-1156.

Baker, G., R. Gibbons, and K. Murphy. 2002. "Relational Contracts and the Theory of the Firm." *The Quarterly Journal of Economics* 117.1: 39-84.

Baldenius, T., J. Glover, and H. Xue. 2016. Relational contracts with and between agents. *Journal of Accounting and Economics* 61: 369-390.

Banerjee, A., and E. Duflo. 2000. Reputation effects and the limits of contracting: A study of the Indian software industry. *The Quarterly Journal of Economics* 115: 989-1017.

Baiman, S., D. Larcker, and M. Rajan. 1995. Organizational design for business units. *Journal of Accounting Research*: 205-229.

Baiman, S., and M. Rajan. 1995. The informational advantages of discretionary bonus schemes. *The Accounting Review*: 557-579.

Baiman, S., and M. Rajan. 2002. Incentive issues in inter-firm relationships. *Accounting, Organizations and Society* 27: 213-238.

Baiman, S., and M. Rajan. 2002. The role of information and opportunism in the choice of buyer-supplier relationships. *Journal of Accounting Research* 40: 247-278.

Banker, R., and S. Datar. 1989. "Sensitivity, precision, and linear aggregation of signals for performance evaluation." *Journal of Accounting Research* 27: 21-39.

Banker, R., G. Potter, and D. Srinivasan. 2000. "An empirical investigation of an incentive plan that includes nonfinancial performance measures." *The Accounting Review* 75.1: 65-92.

Barro, J., and R. Barro. 1990. "Pay, performance, and turnover of bank CEOs." *Journal of Labor Economics* 8.4: 448-481.

Barron, D., R. Gibbons, R. Gil, and K. Murphy. 2015. Relational Adaptation under Reel Authority. *MIT Working Paper*.

Bastien, D. 1987. "Common patterns of behavior and communication in corporate mergers and acquisitions." *Human Resource Management* 26: 17-34.

Bebchuk, L., K. Cremers, and U. Peyer. 2011. "The CEO pay slice." *Journal of Financial Economics* 102.1: 199-221.

Bennett, B., J. Bettis, R. Gopalan, and T. Milbourn. 2017. Compensation goals and firm performance. *Journal of Financial Economics* 124.2: 307-330.

Bernheim, B.D., and M.D. Whinston. 1998. Incomplete Contracts and Strategic Ambiguity. *American Economic Review* 88: 902-32.

Bertrand, M., and S. Mullainathan. 2001. "Are CEOs rewarded for luck? The ones without principals are." *The Quarterly Journal of Economics* 116.3: 901-932.

Bettis, J., J. Bizjak, J. Coles, and B. Young. 2014. The presence, value, and incentive properties of relative performance evaluation in executive compensation contracts. *SSRN Working Paper* 2392861.

Bizjak, J., S. Kalpathy, Z. Li, and B. Young, B. 2017. The role of peer firm selection in explicit relative performance awards. *SSRN Working Paper* 2833309.

Board, S. 2011. Relational Contracts and the Value of Loyalty. *American Economic Review* 101: 3349-67.

Bol, J. 2008. "Subjectivity in Compensation Contracting". *Journal of Accounting Literature* 27: 1-24.

Bol, J. 2011. "The determinants and performance effects of managers' performance evaluation biases." *The Accounting Review* 86.5: 1549-1575.

Bol, J., and S. Smith. 2011. "Spillover effects in subjective performance evaluation: Bias and the asymmetric influence of controllability." *The Accounting Review* 86.4: 1213-1230.

Bouwens, J., and L. Van Lent. 2007. "Assessing the performance of business unit managers." *Journal of Accounting Research* 45.4: 667-697.

Bowman, E., and H. Singh. 1993. "Corporate restructuring: Reconfiguring the firm." *Strategic Management Journal* 14.S1: 5-14.

Brav, A., W. Jiang, F. Partnoy, and R. Thomas. 2008. Hedge fund activism, corporate governance, and firm performance. *The Journal of Finance* 63.4: 1729-1775.

Bull, C. 1987. The existence of self-enforcing relational contracts. *The Quarterly Journal of Economics* 102: 147-159.

Buono, A, J. Bowditch, and J. Lewis. 1985. "When cultures collide: The anatomy of a merger." *Human Relations* 38: 477-500.

Burney, B. 2016. Relative tsr prevalence and design of s&p 500 companies. *Exequity White Paper*

Cai, C., O. Kini, and R. Williams. 2016. Do compensation consultants have distinct styles*? SSRN Working Paper* 2724072.

Calzolari, G., L. Felli, J. Koenen, G. Spagnolo, and K. Stahl. 2015. Trust, competition and innovation: theory and evidence from German car manufacturers. *CESifo Working Paper*

Campbell, D. 2008. "Nonfinancial performance measures and promotion-based incentives." *Journal of Accounting Research* 46.2: 297-332.

Campbell, D. 2012. "Employee selection as a control system." *Journal of Accounting Research* 50: 931-966.

Campbell, D., S. Datar, and T. Sandino. 2009. "Organizational design and control across multiple markets: The case of franchising in the convenience store industry." *The Accounting Review* 84.6: 1749-1779.

Cardinaels, E., and H. Yin. 2015. "Think Twice Before Going for Incentives: Social Norms and the Principal's Decision on Compensation Contracts." *Journal of Accounting Research* 53.5: 985-1015.

Carter, M., C. Ittner, S. and Zechman. 2009. Explicit relative performance evaluation in performance-vested equity grants. *Review of Accounting Studies* 14 (2-3): 269-306.

Cohen, Y., and J. Pfeffer. 1986. “Organizational Hiring Standards.” *Administrative Science Quarterly*
31: 1–24.

Coles, J., M. Lemmon, and J. Meschke. 2012. "Structural models and endogeneity in corporate finance: The link between managerial ownership and corporate performance." *Journal of Financial Economics* 103.1: 149-168.

Core, J., W. Guay, and R. Verrecchia. 2003. "Price versus non-price performance measures in optimal CEO compensation contracts." *The Accounting Review* 78.4: 957-981.

Datar, S., S. Kulp, and R. Lambert. 2001. "Balancing performance measures." *Journal of Accounting Research* 39.1: 75-92.

Davila, T. 2005. “An exploratory study on the emergence of management control systems: Formalizing human resources in small growing firms.” *Accounting, Organizations and Society* 30: 223-248.

De Angelis, D. and Y. Grinstein. 2016. Relative performance evaluation in CEO compensation: A non-agency explanation. *SSRN Working Paper* 2432473.

DeMarzo, Peter M., M. Fishman, Z. He, and N. Wang. 2012. "Dynamic agency and the q theory of investment." *The Journal of Finance* 67.6: 2295-2340.

Dikolli, S. 2001. "Agent employment horizons and contracting demand for forward-looking performance measures." *Journal of Accounting Research* 39.3: 481-494.

Dikolli, S., C. Hofmann, and T. Pfeiffer. 2013. "Relative performance evaluation and peer-performance summarization errors." *Review of Accounting Studies* 18.1: 34-65.

Drexler, A., and A. Schoar. 2014. Do relationships matter? Evidence from loan officer turnover. *Management Science* 60: 2722-2736.

Dutta, S. and S. Reichelstein. 2005. "Stock price, earnings, and book value in managerial performance measures." *The Accounting Review* 80.4: 1069-1100.

Ederhof, Merle. 2010. "Discretion in bonus plans." *The Accounting Review* 85.6: 1921-1949.

Eriksson, T., and J. Ortega. 2006. The adoption of job rotation: Testing the theories. *ILR Review* 59: 653-666.

Fee, C. E., & Thomas, S. 2004. "Sources of gains in horizontal mergers: evidence from customer, supplier, and rival firms." *Journal of Financial Economics*, *74*:423-460.

Feltham, G., and J. Xie. 1994. "Performance measure congruity and diversity in multi-task principal/agent relations." *The Accounting Review* 69: 429-453.

Francis, B., I. Hasan, S. Mani, and P. Ye. 2016. "Relative peer quality and firm performance." *Journal of Financial Economics* 122.1: 196-219.

Garvey, G., and T. Milbourn. 2003. "Incentive compensation when executives can hedge the market: Evidence of relative performance evaluation in the cross section." *The Journal of Finance* 58.4: 1557-1582.

Gayle, G., and R. Miller. 2009. "Has moral hazard become a more important factor in managerial compensation?." *American Economic Review* 99.5: 1740-69.

Gayle, G., and R. Miller. 2015. "Identifying and testing models of managerial compensation." *The Review of Economic Studies* 82.3: 1074-1118.

Gibbons, R. 1997. An Introduction to Applicable Game Theory. *Journal of Economic Perspectives* 11: 127-49.

Gibbons, R., and R. Henderson. 2012. Relational contracts and organizational capabilities. *Organization Science* 23: 1350-1364.

Gibbons, R., and K. Murphy. 1990. "Relative performance evaluation for chief executive officers." *ILR Review* 43.3: 30-S.

Gibbons, R., and K. Murphy. 1992. "Optimal incentive contracts in the presence of career concerns: Theory and evidence." *Journal of Political Economy* 100.3: 468-505.

Gibbs, M., K. Merchant, W. Van der Stede, and M. Vargus. 2004. "Determinants and effects of subjectivity in incentives." *The Accounting Review*, 79(2): 409-436.

Gil, R. 2013. The interplay of formal and relational contracts: Evidence from movies. *Journal of Law, Economics, and Organization* 29: 681-710.

Gil, R., and J. Marion. 2013. Self-enforcing agreements and relational contracting: evidence from California highway procurement. *The Journal of Law, Economics, and Organization* 29: 239-277.

Gipper, Brandon, L. Hail, and C. Leuz. 2017. On the Economics of Audit Partner Tenure and Rotation: Evidence from PCAOB Data. *Stanford University Graduate School of Business Research Paper* No. 17-56.

Godfrey, D. and J. Bourchier. 2015. Discussion of tsr alternatives. *GRG Remuneration Insight* 77.

Gong, G., L. Li, and J. Shin. 2011. "Relative performance evaluation and related peer groups in executive compensation contracts." *The Accounting Review* 86.3: 1007-1043.

González-Uribe, J. and M. Groen-Xu. 2017. "CEO contract horizon and innovation." London School of Economics Working Paper.

Goo, J., R. Kishore, H. Rao, and K. Nam. 2009. The role of service level agreements in relational management of information technology outsourcing: an empirical study. *MIS Quarterly*: 119-145.

Grossman, Sanford J., and Hart, Oliver D. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy* 94: 691-719.

Halac, M. 2012. Relational Contracts and the Value of Relationships. *American Economic Review* 102: 750-79.

Hart, Oliver D. 1995. *Firms, Contracts, and Financial Structure*. Oxford: Clarendon.

Haubrich, J. 1994. "Risk aversion, performance pay, and the principal-agent problem." *Journal of Political Economy* 102.2: 258-276.

Hayes, R., and S. Schaefer. 2000. Implicit contracts and the explanatory power of top executive compensation for future performance. *RAND Journal of Economics* 31: 273-293.

Hayes, R., and S. Schaefer. 2009. "CEO pay and the Lake Wobegon effect." *Journal of Financial Economics* 94.2: 280-290.

Hemmer, T. 2015. Optimal dynamic relative performance evaluation. *Rice University Working Paper*.

Hoffmann, F., and S. Pfeil. 2010. "Reward for luck in a dynamic agency model." *The Review of Financial Studies* 23.9: 3329-3345.

Hogg, Michael A., and Deborah J. Terry. 2014. *Social identity processes in organizational contexts*. Psychology Press.

Holmstrom, B. 1979. "Moral hazard and observability." *The Bell Journal of Economics*: 74-91.

Holmstrom, B. 1982. "Moral hazard in teams." *The Bell Journal of Economics*: 324-340.

Holmstrom, B. 1999. "Managerial incentive problems: A dynamic perspective." *The review of Economic studies* 66.1: 169-182.

Holmstrom, B., and P. Milgrom. 1987. "Aggregation and linearity in the provision of intertemporal incentives." *Econometrica* 55.2: 303-328.

Holmstrom, B., and P. Milgrom. 1991. "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design." *Journal of Law, Economics, & Organization* 7: 24-52.

Hoppe, F., and F. Moers. 2011. "The choice of different types of subjectivity in CEO annual bonus contracts." *The Accounting Review* 86.6: 2023-2046.

Hugessen. 2016. Assessing relative tsr for your company: A brief overview. *Hugessen Consulting White Paper.*

Hvide, H. 2002. "Tournament rewards and risk taking." *Journal of Labor Economics* 20.4: 877-898.

Indjejikian, R., and M. Matějka. 2009. "CFO fiduciary responsibilities and annual bonus incentives." *Journal of Accounting Research* 47.4: 1061-1093.

Indjejikian, R., and M. Matějka. 2011. "Accounting decentralization and performance evaluation of business unit managers." *The Accounting Review* 87.1: 261-290.

Indjejikian, R., M. Matějka, K. Merchant, and W. Van der Stede. 2014. Earnings targets and annual bonus incentives. *The Accounting Review* 89(4): 1227-1258.

Indjejikian, R., M. Matějka, and J. Schloetzer. 2014. "Target ratcheting and incentives: Theory, evidence, and new opportunities." *The Accounting Review* 89.4: 1259-1267.

Ittner, C., and D. Larcker. 1997 "Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction." *Journal of Accounting Research* 36: 1-35.

Ittner, C., and D. Larcker. 2001. "Assessing Empirical Research in Managerial Accounting: A Value-Based Management Perspective." *Journal of Accounting and Economics* 32: 349-410.

Ittner, C., and D. Larcker. 2002. "Determinants of performance measure choices in worker incentive plans." *Journal of labor economics* 20.S2: S58-S90.

Ittner, C., D. Larcker, and M. Rajan. 1997. "The choice of performance measures in annual bonus contracts." *The Accounting Review*: 231-255.

Janakiraman, S., R. Lambert, and D. Larcker. 1992. "An empirical investigation of the relative performance evaluation hypothesis." *Journal of Accounting Research*: 53-69.

Jemison, D., and S. Sitkin. 1986. "Corporate acquisitions: A process perspective." *Academy of Management Review* 11.1: 145-163.

Jensen, M., and W. Meckling. 1992. "*Knowledge, control and organizational structure: Parts I and II.*" Contract economics: 251-74.

Jensen, M., and W. Meckling. 1998. Specific and general knowledge and organizational structure. Published in *Foundations of Organizational Strategy*. Harvard University Press.

Johnson, S., J. McMillan, and C. Woodruff. 2002. Courts and relational contracts. *Journal of Law, Economics, and Organization* 18: 221-277.

Joshi, M. P., R. Kathuria, and S. J. Porth. 2003. Alignment of strategic priorities and performance: An integration of operations and strategic management perspectives. *Journal of Operations Management* 21: 353-369.

Kaplan, R. S., and D. P. Norton. 1996. *The balanced scorecard: translating strategy into action*. Boston, MA: Harvard Business School Press

Keil, J. 2017 "The trouble with approximating industry concentration from Compustat." *Journal of Corporate Finance* 45: 467-479.

Kim, J., and S. Finkelstein. 2009. "The effects of strategic and market complementarity on acquisition performance: Evidence from the US commercial banking industry, 1989-2001." *Strategic Management Journal*: 617-646.

Klein, B. 1996. Why hold-ups occur: the self-enforcing range of contractual relationships. *Economic Inquiry* 34: 444-463.

Lambert, R., and D. Larcker. 1987. "An analysis of the use of accounting and market measures of performance in executive compensation contracts." *Journal of Accounting Research*: 85-125.

Larsson, R., and S. Finkelstein. 1999. "Integrating strategic, organizational, and human resource perspectives on mergers and acquisitions: A case survey of synergy realization." *Organization Science* 10.1: 1-26.

Lazear, E., and S. Rosen. 1981. "Rank-order tournaments as optimum labor contracts." *Journal of political Economy* 89.5: 841-864.

Lee, C., P. Ma, and C. Wang. 2015. "Search-based peer firms: Aggregating investor perceptions through internet co-searches." *Journal of Financial Economics* 116.2: 410-431.

Lee, C., P. Ma, and C. Wang. 2016. "The search for peer firms: When do crowds provide wisdom?" *Harvard Business School Working Paper*.

Levin, J. 2003. Relational incentive contracts. *American Economic Review* 93: 835-857.

Lewellen, S. 2015. Executive compensation and industry peer groups. *London Business School Working Paper*.

Lewellen, S., and A. Metrick. 2010. Corporate governance and equity prices: Are results robust to industry adjustments. *Yale University Working Paper*.

Loehmann, G. 2016. Revisiting relative tsr. *Compensia White Paper*.

MacLeod, B., and J. Malcomson. 1989. Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica* 57: 447-480.

Macneil, I. 1980. *The New Social Contract: An Inquiry into Modern Contractual Relations*. Yale University Press: New Haven, CT.

Malcomson, J. 2013. "Relational Incentive Contracts." Chapter 25 in *The Handbook of Organizational Economics*. Princeton, NJ: Princeton University Press.

Margiotta, M., and R. Miller. 2000. "Managerial compensation and the cost of moral hazard." *International Economic Review* 41.3: 669-719.

Matějka, M., K. Merchant, and W. Van der Stede. 2009. "Employment horizon and the choice of performance measures: Empirical evidence from annual bonus plans of loss-making entities." *Management Science* 55.6: 890-905.

McMillan, J., and C. Woodruff. 1999. Interfirm relationships and informal credit in Vietnam. *The Quarterly Journal of Economics* 114: 1285-1320.

Medland, C., A. McElheran, A. Stancel, A. Conradi, and J. Anderson. 2016. Choosing the right performance peer group. *Meridian Compensation Partners White Paper*.

Mengistae, T., and C. Lixin. 2004. "Agency theory and executive compensation: The case of Chinese state-owned enterprises." *Journal of labor Economics* 22.3: 615-637.

Mercer. 2013. Total shareholder return: Making it work. *Mercer Compensation White Paper.*

Milbourn, T. 2003. "CEO reputation and stock-based compensation." *Journal of Financial Economics* 68.2: 233-262.

Merchant, K. 1985. “Control in Business Organizations”. Cambridge, MA: Ballinger, 1985.

Merchant, K., W. Van der Stede, and L. Zheng. 2003. “Disciplinary Constraints on the Advancement of Knowledge: The Case of Organizational Incentive Systems.” *Accounting, Organizations and Society* 28: 251-86.

Moers, F. 2006. "Performance measure properties and delegation." *The Accounting Review* 81.4: 897-924.

Morse, A., V. Nanda, and A. Seru. 2011. "Are incentive contracts rigged by powerful CEOs?." *The Journal of Finance* 66.5: 1779-1821.

Milgrom, P., and J. Roberts. 1992. *Economics, Organization, and Management*. Englewood Cliffs, N.J.: Prentice-Hall.

Myers, J., L. Myers, and T. Omer. 2003. Exploring the term of the auditor-client relationship and the quality of earnings: A case for mandatory auditor rotation?. *The Accounting Review* 78: 779-799.

Nagar, V. 2002. "Delegation and incentive compensation." *The Accounting Review* 77.2: 379-395.

Nelson, R., and S. Winter. 1982. An Evolutionary Theory of Economic Change. Cambridge, MA: Harvard University Press 1982.

Ortega, J. 2001. Job rotation as a learning mechanism. *Management Science* 47: 1361-1370.

Pakela, S., B. Lane, and B. Scheiring. 2017. Effectively administering a relative tsr program: Learning and best practices. *Pay Governance LLC White Paper.*

Pasmore, W., and M. Fagans. 1992. "Participation, individual development, and organizational change: A review and synthesis." *Journal of Management* 18.2: 375-397.

Patel, H. and K. Edwards. 2012. A more thoughtful approach to using relative tsr in performance plans. *Towers Watson White Paper*.

Poppo, L., and T. Zenger. 2002. Do formal contracts and relational governance function as substitutes or complements?. *Strategic Management Journal* 23: 707-725.

Prendergast, C. 2002. The tenuous trade-off between risk and incentives. *Journal of Political Economy* 110: 1071-1102.

Prendergast, C. 2008. “Intrinsic Motivation and Incentives.” *American Economic Review: Papers & Proceedings* 98: 201–5.

Radford. 2016. Relative tsr is the top metric for performance-based equity compensation. *Aon Hewitt Radford White Paper.*

Raith, M. 2008. "Specific knowledge and performance measurement." *The RAND Journal of Economics* 39.4: 1059-1079.

Rajan, M., and S. Reichelstein. 2006. Subjective performance indicators and discretionary bonus pools. *Journal of Accounting Research* 44: 585-618.

Rajan, M., and S. Reichelstein. 2009. Objective versus subjective indicators of managerial performance. *The Accounting Review* 84: 209-237.

Reda, J., and M. Tonello. 2015. The conference board CEO and executive compensation practices 2015 edition key findings. *SSRN Working Paper* 270256.

Ross, S. 1973. The economic theory of agency: The principal's problem. *American Economic Review* 63: 134-139.

Sandino, T. 2007. "Introducing the First Management Control Systems: Evidence from the Retail Sector." *The Accounting Review* 82: 265-293.

Sashkin, M., and W. Burke. 1987. "Organization development in the 1980's." *Journal of Management* 13.2 (1987): 393-417.

Scharfstein, D., and J. Stein. 1990. "Herd behavior and investment." *The American Economic Review:* 465-479.

Schweiger, D., J. Ivancevich, and F. Power. 1987. "Executive actions for managing human resources before and after acquisition." *Academy of Management Executive* 1: 127-138.

Schweiger, D., and Y. Weber. 1989. "Strategies for managing human resources during mergers and acquisitions: An empirical investigation". *Human Resource Planning* 12(2): 69-86.

Securities Exchange Commission (SEC). 1994. Staff Report on Auditor Independence. Washington, D.C.: *Government Printing Office*

Shavell, S. 1979. "Risk sharing and incentives in the principal and agent relationship." *The Bell Journal of Economics*: 55-73.

Shin, J. 2016. An evaluation of compensation benchmarking peer groups based on mutual peer-designating behaviors. *SSRN Working Paper* 2823592.

Simons, R. 1987. "Accounting Control Systems and Business Strategy: An Empirical Analysis." *Accounting Organizations and Society* 12: 357-374.

Simons, R. 1990. "The role of management control systems in creating competitive advantage: New perspectives." *Accounting, Organizations and Society* 15: 127-143.

Simons, R. 1994. "Levers of Control: How Managers Use Innovative Control Systems To Drive Strategic Renewal." Boston, MA. Harvard University Press 1994.

Simons, R. 2000. "Performance Measurement and Control Systems for Implementing Strategy." Upper Saddle River, NJ: Prentice Hall, 2000.

Sinkular, J. and J. Kennedy. 2016. Analysis of recent annual incentive financial goals. *Pay Governance LLC White Paper.*

Sirras, T., and B. Sullivan. 2012. The problem with relative total shareholder returns. *Workspan Magazine.*

Sues, J., A. Ulaj, and E. Graskamp. 2016. The 2015 top 250 report: Long-term incentive grant practices for executives. *Frederick W Cook & Co White Paper*.

Swinford, D. 2015. The limits of using tsr as an incentive measure. *Pearl Meyer Partners White Paper*.

Topkis, D. 1978. "Minimizing a submodular function on a lattice." *Operations research* 26.2: 305-321.

Tran, J., S. Wong, and M. Goforth. 2016. Compensation consultant market share rankings 2015. *Equilar White Paper*.

U.S. Senate. 1976. The Accounting Establishment. Prepared by the Subcommittee on Reports, Accounting and Management of the Committee on Governmental Affairs. Washington, D.C.: *Government Printing Office*.

Van den Steen, E. 2005. "Organizational Beliefs and Managerial Vision." *Journal of Law, Economics, and Organization* 21: 256–83.

Van den Steen, E. 2010a. "On the Origin of Shared Beliefs (and Corporate Culture)." *RAND Journal of Economics* 41: 617–48.

Van den Steen, E. 2010b. "Culture Clash: The Costs and Benefits of Homogeneity." *Management Science* 56: 1718–38.

Vnuk, M., S. Bisman, and C. Carter. 2012. Long-term incentive plans with relative performance measurements. *Compensation Advisory Partners White Paper*.

Vuolteenaho, T. 2002. "What drives firm-level stock returns?" *The Journal of Finance* 57.1: 233-264.

Walker, D. 2016. The way we pay now: Understanding and evaluating performance-based executive pay. *Journal of Law, Finance, and Accounting* 1.2: 395-449.

Weber, Yaakov. 1996. "Corporate cultural fit and performance in mergers and acquisitions." *Human relations* 49.9: 1181-1202.

Williamson, O. 1975. Markets and Hierarchies: Analysis and Antitrust Implications. New York: Free Press.

Williamson, O. 1979. Transaction-cost economics: the governance of contractual relations. *The Journal of Law and Economics* 22.2: 233-261.

Williamson, O. 1985. The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting. New York, Free Press.

Woodman, R. 1989. "Organizational change and development: New arenas for inquiry and action." *Journal of Management* 15.2: 205-228.

Zanarone, G. 2013. Contract adaptation under legal constraints. *Journal of Law, Economics, and Organization* 29: 799-834.

APPENDIX

Appendix Table I. Sample Selection

This table reports the selection criterion used to generate the final samples used in Tables 4.5 and 4.6.

Main Sample Selection	Firm-year Observations	Firm-year-month Observations	Unique Firms
(1) Firms in ISS Incentive Lab data that include CEO grant data between fiscal year 2004 and 2013	12,216		1,668
(2) Less firms without CEO grants based on an RP component	(8,998)		
	3,218		751
(3) Less firms whose relative benchmark cannot be identified	(685)		
	2,533		645
(4) Less firms that do not use stock price as the relevant RP performance measure	(486)		
	2,047		554
(5) Less firms without CIK-GVKEY matches	(226)		
	1,821		487
(6) Merged with monthly return data from CRSP		21,710	
(7) Less observations with missing SBP data	(635)	(6,654)	(131)
(8) Less observations before calendar year 2006	(50)	(764)	(4)
(9) Less observations that use both, index and specific peers, in a given fiscal year	(85)	(1,107)	(11)
(10) Less observations with fewer than 10 monthly returns in the time-series regressions	(13)	(77)	(11)
Final Sample	1,038	13,108	330

Appendix Table II. Summary Statistics of Compensation Benchmarking Peers

This table reports the number of relative performance (RP) and compensation benchmarking (CB) peers, and the extent to which these two peer sets overlap for each fiscal year using the ISS Incentive Labs data prior to any sample selection restrictions. Columns 1 and 2 report the mean number of chosen peers in the performance benchmarking peer group and the compensation benchmarking peer group, respectively. Column 3 reports the mean number of firms in both peer groups. Columns 4 and 5 report the fraction of overlapping peers in the performance benchmarking peer group and the compensation benchmarking peer group, respectively.

Fiscal Year	# of RP Peers	# of CB Peers	# of Overlap Peers	Proportion of RP Peers Also in CB Peers	Proportion of CB Peers Also in RP Peers
	(1)	(2)	(3)	(4)	(5)
2006	15.18	20.55	8.66	0.63	0.47
2007	15.90	25.44	9.18	0.67	0.47
2008	16.81	26.93	8.54	0.67	0.49
2009	16.56	25.55	8.68	0.63	0.46
2010	17.30	26.06	8.66	0.62	0.46
2011	18.21	25.79	8.70	0.62	0.46
2012	17.24	21.09	7.78	0.62	0.43
2013	17.84	21.83	6.83	0.60	0.40
2014	17.37	19.60	6.25	0.55	0.37

Appendix Table III. Compensation Consultants' Description of the Objective of Relative TSR

This table presents passages from prominent compensation consulting firms' white papers about the objectives and implementation of relative total shareholder returns (rTSR) in executive performance-based incentive plans. The Market Share column draws on Equilar's 2015 Market Share Rankings, which are based on S&P 500 board engagements (Tran et al. 2016). Note that the concept of removing common noise and shareholder alignment are one and the same. Without loss of generality, under a linear one-factor structure, $p = a + c + e$, where p is returns, c is the common component of returns, and e is idiosyncratic noise in returns, relative returns $(p - c)$ is equivalent to the excess returns of the firm $(a + e)$, i.e., the manager's "alpha".

Firm	Market Share	Description
Frederic Cook	25.8%	TSR, specifically relative TSR, has emerged as the metric of choice under Say-on-Pay. For shareholders, there is an elegance to TSR in that it demonstrates the return relative to alternative investments. It is also the singular definition of corporate performance used by ISS, as well as the sole performance metric required by the SEC for pay and performance disclosure under Dodd-Frank. As such, some companies view relative TSR as a means to satisfy shareholder, ISS and SEC preferences (Sues et al. 2016).
Meridian Compensation	13%	Defining performance in terms of relative ranking against peers is easier - all a committee has to do is express a philosophical preference about what portion of peers a company has to outperform in order to earn a PSU payout at some level. The market does all of the heavy lifting, taking into account the macroeconomic and other market factors affecting the business (Medland et al. 2016).
Pay Governance	12.8%	Relative TSR is a performance metric most often used in LTI performance plans. Its use as a metric has nearly doubled over the past 5 years and is now used by approximately 50% of companies spanning all sizes and industries. ...the appeal of this metric for shareholders and directors alike is its alignment with shareholder value creation and the absence of having to establish long-term performance goals (Pakela et al. 2017).
Pearl Meyer Partners	7.9%	Measuring TSR on a relative basis levels the playing field by removing overall market movements and industry cycles from the evaluation of executive performance (Swinford 2015).
Semler Brossy	5.1%	The theory behind relative shareholder return as an incentive metric is sound: Executives earn rewards only when shareholders experience above-market returns (Sirras and Sullivan 2012).
Tower Watson	5.1%	There are numerous theoretical and pragmatic arguments for the use of relative TSR as a long-term measure. Relative TSR is viewed as objective, transparent and easy to communicate to participants. It supports shareholder alignment and incorporates relative measurement (Patel and Edwards 2012).
Exequity	4.2%	[rTSR] is intended to align executive wealth with the shareholder experience (Burney 2016).
CA Partners	4%	Companies tend to use industry peer groups or indices when economic factors have a unique impact on the industry. Companies that use a broader market group, such as a general industry index like the S&P 500, believe that companies compete broadly for investor dollars (Vnuk et al. 2012).
Mercer	3.5%	In an ideal world, companies would use a peer group of like organizations that are subject to the same external influences so that share price movements genuinely reflect decisions made by management (Mercer 2013).
Compensia	2.6%	[rTSR] provides an unambiguous link to shareholder value, with outcomes that are not driven by overachievement against low expectations (Loehmann 2016).

Appendix Table III. Compensation Consultants' Description of the Objective of Relative TSR (Continued)

Hugessen Consulting	NA	[rTSR] counterbalances the general market movement "wind-fall" problems associated with stock options. Satisfies motivation and retention objectives in both up and down markets. [rTSR] may result in a closer measure of management performance (measuring "alpha"), at least theoretically (Hugessen 2016).
Radford (Aon Hewitt)	NA	Performance-based equity incentive plans with Relative TSR metrics offer companies a wide range of important benefits, including: Reduced forecasting requirements for long-term performance goals, especially in an uncertain macro-economic environment. Clearer links between final executive compensation payouts and shareholder value creation. Reduced use of redundant performance metrics between annual cash incentive plans and long-term equity incentive plans, improving the overall risk profile of executive compensation programs (Radford 2016).

Appendix Table IV. Descriptive Statistics

Panel A reports summary statistics on the variables used in Tables 4.7, 4.8, and 4.9. Panel B reports the correlation matrix of the same variables. Observations are at the annual (fiscal) firm-benchmark level. Significance levels are indicated by *, **, *** for 10%, 5%, and 1% respectively.

Variables are defined as follows: the variable names from the relevant databases are reported in brackets. Using Compustat, we define the following variables on firm characteristics: *ROA* is the ratio of net income to total assets [ni/at]; *Log Market Cap* is the log of the firm's market capitalization ($Millions) as of the fiscal year end [mkvalt]; and *Book-to-Market* is the book value of common equity ($Millions) [ceq] divided by market capitalization ($Millions) [mkvalt]. Census-based HHI Index is the US census-based Herfindahl-Hirschman Index available from Keil (2017). Using Execucomp, we define the following variables on CEO characteristics: *Log CEO Pay* is the log of the CEO's total compensation (in $Thousands) [tdc1]; *CEO Tenure* is the current year minus the year in which the CEO joined the firm [becameceo]; and *CEO Age* is the age of the CEO [age]. Using MSCI GMI's databases on companies and directorships, we define the following variables on board characteristics: *% Busy Directors* is the percentage of the firm's directors with more than four board seats at public firms; *Board Size* is the number of directors on the board; *Director Workload* is the number of full board meetings held over the prior fiscal year [BDMTGS] divided by the number of directors, and *% Age 65+ Directors* is the fraction of board members who are aged 66 or greater. Using the ISS Governance database, we define the following variables on a firm's governance characteristics: *Staggered Board* is a variable that equals 1 if the firm holds staggered director elections and 0 if the firm has a unitary board; *Dual-Class Shares* is an indicator variable that equals 1 if the firm has multiple classes of voting shares and 0 if it has a single class of voting shares. Using CRSP, we define *Return Volatility* as the standard deviation of monthly cum-dividend returns [ret] of a firm over the fiscal year. Finally, *Index* is a dummy variable that equals 1 if the firm uses an index as its relative performance benchmark in a given fiscal year.

Panel A: Distributional Statistics

	Obs	*Mean*	*Std.Dev.*	*P25*	*Median*	*P75*
ROA	1175	0.049	0.053	0.024	0.044	0.078
Index	1175	0.341	0.474	0.000	0.000	1.000
Log CEO Pay	1175	8.894	0.681	8.452	8.874	9.357
CEO Tenure	1175	5.415	4.593	2.000	4.000	7.000
CEO Age	1175	56.445	5.192	53.000	57.000	60.000
% Busy Directors	1175	0.022	0.048	0.000	0.000	0.000
Board Size	1175	10.597	2.06	9.000	10.000	12.000
Director Workload	1175	0.807	0.35	0.583	0.727	1.000
% Age 65+ Directors	1175	0.311	0.319	0.214	0.333	0.500
Log Market Cap	1175	9.038	1.272	8.123	8.886	9.754
Cenusus-based HHI Index	1175	0.072	0.038	0.051	0.060	0.082
Return Volatility	1175	0.079	0.047	0.047	0.068	0.098
Book-to-Market	1175	0.52	0.312	0.302	0.483	0.683
Staggered Board	1175	0.344	0.475	0.000	0.000	1.000
Dual-Class Shares	1175	0.031	0.172	0.000	0.000	0.000

Appendix Table IV. Descriptive Statistics (Continued)

Panel B: Correlation Matrix

ROA	1.00														
Index	-0.04	1.00													
Log CEO Pay	0.17***	0.14***	1.00												
CEO Tenure	0.02	0.07**	0.01	1.00											
CEO Age	0.07**	0.03	0.06**	0.37***	1.00										
% Busy Directors	0.01	0.05*	0.06**	-0.03	-0.07**	1.00									
Board Size	-0.01	0.06**	0.24***	-0.07**	0.05*	0.05*	1.00								
Director Workload	-0.08***	0.09***	0.00	-0.06**	-0.09***	0.06**	-0.35***	1.00							
% Age 65+ Directors	0.03	0.09***	0.09***	0.07**	0.12***	-0.09***	0.08***	-0.09***	1.00						
Log Market Cap	0.31***	0.05	0.68***	-0.05*	0.06**	0.06**	0.44***	-0.09***	0.12***	1.00					
Cenusus-based HHI Index	0.07**	0.05	0.18***	-0.04	0.07**	0.03	0.16***	-0.07**	-0.09***	0.11***	1.00				
Return Volatility	-0.26***	-0.00	-0.09***	0.04	-0.08***	0.02	-0.21***	0.12***	-0.03	-0.34***	-0.01	1.00			
Book-to-Market	-0.46***	-0.07**	-0.16***	-0.03	0.03	-0.06**	0.01	0.09***	0.03	-0.25***	-0.13***	0.23***	1.00		
Staggered Board	0.03	0.01	-0.13***	0.01	0.02	-0.04	-0.13***	-0.01	0.16***	-0.19***	-0.10***	0.08***	0.04	1.00	
Dual-Class Shares	0.13***	0.01	0.11***	-0.04	0.01	-0.02	0.09***	-0.06**	0.01	0.07**	0.04	-0.03	-0.12***	-0.01	1.00