Title
Essays on the Economics of Organizations, Productivity and Labor

Author
Cowgill, Bradford Lee

Publication Date
2015

Peer reviewed|Thesis/dissertation

Essays on the Economics of Organizations, Productivity and Labor

by

Bradford Lee Cowgill, Jr.

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor John Morgan, Chair
Professor Noam Yuchtman
Professor Stefano DellaVigna
Professor David Card

Spring 2015

Essays on the Economics of Organizations, Productivity and Labor

Abstract

Essays on the Economics of Organizations, Productivity and Labor

by

Bradford Lee Cowgill, Jr.

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor John Morgan, Chair

This dissertation is about how firms use incentives and information in internal personnel and management practices, in particular relating to hiring and innovation. In the first chapter, I study competition between workers inside of firms. Why do firms use incentives that encourage anti-social behavior among employees? Rank-based promotion schemes are among the most widespread forms competition and incentives, despite encouraging influence-peddling, sabotage and anti-social behavior. I study a natural experiment using rich administrative data from a large, white collar firm. At the firm, competitors for promotions depend partly on dates-of-hire. I utilize the date-of-hire assignment as a source of exogenous variation in the intensity of intra-worker competition. I use the firm's multidimensional timestamped productivity logs as "time diaries" to study the amount, character and allocation of output across tasks. I find that competition has significant incentives for effort and efficiency – as well as lobbying- and sabotage- like behaviors – without affecting the quality and innovativeness of output. I also find that employees facing high competition are more likely to quit and join other companies, particularly higher-performing employees. Lastly, I show that competition induces workers to differentiate and specialize by concentrating effort into a smaller set of tasks. These results show that while workers respond to incentives from competition, they also seek to avoid it through sorting and differentiation strategies. The productivity gains from differentiation and specialization may partly explain the common use of these incentives by firms.

In the second chapter, I study how firms use social networks in hiring. Using personnel data from nine large firms in three industries (call-centers, trucking, and high-tech), we empirically assess the benefit to firms of hiring through employee referrals. Compared to non-referred applicants, referred applicants are more likely to be hired and more likely to accept offers, even though referrals and non-referrals have similar skill characteristics. Referred workers tend to have similar productivity compared to non-referred workers on most measures, but referred workers have lower accident rates in trucking and produce more patents in high-tech. Referred workers are substantially less likely to quit and earn

slightly higher wages than non-referred workers. In call-centers and trucking, the two industries for which we can calculate worker-level profits, referred workers yield substantially higher profits per worker than non-referred workers. These profit differences are driven by lower turnover and lower recruiting costs for referrals.

In the third and final chapter, I study the use of betting markets inside of firms. Despite the popularity of prediction markets among economists, businesses and policymakers have been slow to adopt them in decision making. Most studies of prediction markets outside the lab are from public markets with large trading populations. Corporate prediction markets face additional issues, such as thinness, weak incentives, limited entry and the potential for traders with biases or ulterior motives raising questions about how well these markets will perform. We examine data from prediction markets run by Google, Ford Motor Company and an anonymous basic materials conglomerate (Firm X). Despite theoretically adverse conditions, we find these markets are relatively efficient, and improve upon the forecasts of experts at all three firms by as much as a 25% reduction in mean squared error. The most notable inefficiency is an optimism bias in the markets at Google. The inefficiencies that do exist generally become smaller over time. More experienced traders and those with higher past performance trade against the identified inefficiencies, suggesting that the markets' efficiency improves because traders gain experience and less skilled traders exit the market.

Chapter 1: Competition and Productivity in Employee Promotion Contests

1 Introduction

Competition between employees within a firm is among the most widespread forms of formal incentives. In America, intra-worker competition at the same firm is a feature of wage growth for 77% of the workforce.[1] Worker-vs-worker competition has been publicly championed as a management technique by high-profile business leaders including Jack Welch, Marissa Mayer and Steve Ballmer, and is a leading application for a growing formal literature on contest theory (Siegel, 2009).

Despite its prevalence, workplace competition has been widely criticized for encouraging sabotage (Dye, 1984) and influence-peddling (Milgrom and Roberts, 1988) in both academic and popular (Eichenwald, 2005; Swisher, 2013; Carlson, 2014) management circles. Are the potential productivity gains worth the risk of encouraging lobbying and anti-social behavior?

This paper studies the extent to which intra-worker competition affects workplace performance and behavior using a natural experiment. At the firm I study, competitors for promotions depend partly on dates-of-hire. I utilize this as a source of exogenous variation in the intensity of intra-worker competition in an instrumental variables setup.[2]

The outcomes I study come from the firm's administrative data. The firm logs a wide variety of timestamped, labeled worker productivity data to facilitate collaboration and debugging. I use these activity logs as "time diaries" (Hamermesh et al., 2005) to analyze how competition affects the level and composition of worker effort. The data includes not only estimates of time spent, but also the quantity of output per-dimension using the firm's internal task descriptions.[3]

[1] This figure comes from a survey conducted for this paper of over 15,000 employed Americans, reweighed to match the census tracts. The survey and its results are described in Section 2.1 and reported in Table 1.

[2] The first stage has an $F-$ statistic of $\sim$31.

[3] By contrast, traditional time use datasets usually measure only inputs (time) and not outputs. An additional advantage of this data is that it involves no self-reporting. Lastly, my data contains labels of activities designed by the firm, rather than by an external researcher. As such, the division of tasks I study in this pa-

I use this productivity and time use data to measure the effect of competition on worker effort – including not only the amount of effort, but also its efficiency, quality and allocation across multiple dimensions of productivity. For example, I study effects on innovation, productive on-the-job cooperation with colleagues, providing public goods for one's peers and other economic characteristics of output (besides its amount). The multidimensional nature of output also permits analysis of specialization and the division of labor.

This paper has four main findings. First, I find that competition creates strong incentives for effort, output and efficiency, without decreases (or increases) in the quality or innovativeness of output.[4]

Second, the competition does encourage several negative behaviors by employees. Workers report lower job satisfaction in competitive settings, and are less likely to engaged in productive cooperation and organizational citizenship on several dimensions.

Third, I find that employees respond to the competition through sorting and retention decisions. Employees facing high competition sort into different competitive pools via quitting. Workers in high competitive situations are more likely to quit and join other companies. This is especially true for high-performing workers.

Fourth, I find that higher competition induces a different division of labor across workers. Competition affects not only the amount of output, but also its composition and division between workers. I find that competition induces differentiation and specialization by workers. As competition increases, workers increase the concentration of their effort into a subset of tasks.

This result has two interpretations. As early as Smith (1776), economists have viewed specialization and the division of labor within firms as a source of efficiency gains and a rationale for the existence of firms. In intra-worker competition, workers are likely to specialize in tasks where they hold comparative advantages against other contest opponents. The resulting productivity gains from specialization create additional efficiency benefits for the firm.

Specialization also makes comparisons between employees harder. A second interpretation of these results are that differentiation thus resembles a contest-theoretic version of differentiation (Hotelling, 1929; Tirole, 1988) or "obfuscation" (Ellison and Ellison, 2009).

per aligns with the firm's production function (or at least its self-perception thereof). They also assist with interpreting my results in light of the economic theory on incentives, contests and internal labor markets. Despite these advantages, these logs share some disadvantages with traditional time use data, described in Section 4.

[4] Although this result is consistent with most theory, some models of heterogeneity in workplace competition (i.e., Gürtler and Gürtler, 2013) predict the opposite results.

By differentiating, workers can affect the noisiness of evaluation – a parameter set by the principal in most contest-theoretic models. Increases in evaluation noise decrease players' equilibrium effort and increase player welfare.

Specialization thus benefits the worker and the firm, and dampens incentives for sabotage and influence-peddling. It thus represents a heretofore unrecognized benefit of rank-based competition, improving productivity and mitigating antisocial incentives. The effects of competition on the division of labor offer an additional and novel explanation for the widespread use of contests inside firms in multi-tasking environments, despite their negative side-effects on sabotage and politics.

This paper builds on several literatures. I use a methodological approach from the research on the economics of time use (Aguiar et al., 2012) to study a widespread phenomena at the intersection of labor economics and contract theory. Although I describe my data as "time diaries," they differ in important ways from classical time use datasets and I discuss these differences in Section 4.

This paper is also the first to attempt to measure the prevalence of a tournament-like form of wage growth and incentives. Through a large survey, I collect new data to measure the pervasiveness of workplace tournaments: The results in Table 1 suggests they are a part of compensation for 77% of American workers. If anything, these figures may understate the prevalence of tournament-like rewards.[5]

The contract theory literature has embraced tournaments as a descriptive model of incentives within firms. Workplace contests are among the most common motivating applications of the theoretical tournament literature. An analysis of a natural experiment in stock option pricing by Cowgill and Zitzewitz (2009) suggests that promotion contests are a cheaper, more powerful form of incentives than stock options, even using the most generous estimates of the effects of options.

However, despite the central component of tournaments, the existing theoretical literature is mostly concerned with effort, sabotage and (to a lesser extent) "influence activities." This paper shows the central role that sorting, differentiation, specialization and obfuscation may have in the context of workplace tournaments. One recent paper (Morgan et al., 2012) models self-selection into tournaments, but its theoretical predictions are ambiguous. There are few contest-theoretic papers with multidimensional effort at all.[6]

[5]Behavioral economists and psychologists have argued that even without formal contest incentives, workers care about rank-order performance as they directly affect self-image (Benabou and Tirole, 2003; Köszegi, 2006; Maslow, 1943; McClelland et al., 1953) and convey status (Besley and Ghatak, 2008; Moldovanu et al., 2007; Frank, 1985).

[6]Two rare exceptions include Garicano and Palacios-Huerta (2005) and a very a brief discussion in Acemoglu and Jensen (2010). A related literature on auctions has studied multidimensional bidding, for exam-

In addition, the empirical literature on contracting has narrowly focused on situations whereby we observe clear unidimensional measures of output (e.g., sports[7] and manual labor[8]). In contrast, this paper focuses on the allocation of effort across a multidimensional set of tasks, some of which involve more uncertainty (e.g. innovative activities). These are arguably more common workplace environments. These jobs require workers to not only follow manager instructions, but also to form judgments, strategize, take risks, innovate and specialize. The multidimensional nature of output introduces contracting issues around multitasking (Holmstrom and Milgrom, 1991a; Baker, 1992a), specialization, innovation and complementarities that aren't captured in farming or athletic settings.

White-collar work also exhibits contracting and measurement features that naturally lend themselves to tournaments. One of the heralded benefits of contests is "it is not necessary to determine how much better one worker is than another; all that is needed is rank order information" (Prendergast, 1999a). Managers in sports and fruit-picking can easily measure differences in employee output. It is thus unclear why industries such as sports and farms need tournament incentives at all, rather than alternative contracts.

By contrast, managers in white collar work cannot easily measure differences in the market value of software output, patentable ideas or consulting advice – but can make reasonable rank comparisons. As such, these settings embody the information and contracting details that give rise to tournament incentives.

This paper also incorporates ideas from the larger literature on how firms mitigate competition, price wars and other forms of costly competition. In different contexts, economists and strategy researchers have addressed maneuvers to avoid or "soften" competition through differentiation (Hotelling, 1929; Tirole, 1988), tacit or explicit collusion (Stigler, 1964), capacity restrictions, erecting barriers to entry (Caves and Porter, 1977) and/or "obfuscation" (Ellison and Ellison, 2009; Ellison and Wolitzky, 2012).

This paper shows the applicability of these ideas from industrial organization in a labor and organizational economics setting. Like other competitive settings, contests are characterized by zero-expected profits through rent- dissipation. In most laboratory studies, rents are actually over- dissipated in contests (Sheremeta, 2014).

Lastly, this paper contributes to the literature on peer effects (e.g. Sacerdote, 2011). Much of the peer effects literature measures the homogenizing influence of peers on each other. However, the contest literature predicts a different type of peer effects in com-

ple, Che (1993), Yoganarasimhan (2013) and Krasnokutskaya et al. (2013).

[7]For example, Brown (2011); Garicano and Palacios-Huerta (2005); Balafoutas et al. (2012); Becker and Huselid (1992).

[8]For example, Bandiera et al. (2013), Drago and Garvey (1998), Knoeber and Thurman (1994), Knoeber (1989).

petitive settings. This paper demonstrates competition creating specialization and differentiation between peers, rather than homogenization. I also show evidence of endogenous self-sorting (Carrell et al., 2013) motivated by these competition-related peer effects. Tournament-like incentives exist in many of the settings wherein economists care about peer-effects; not only in firms, but also in classrooms where grades are often awarded on relative performances "curves" that create contest-like incentives.

The remainder of this paper is organized as follows. Section 2 describes the setting and details of the firm's promotion practices. Section 3 describes the natural experiment, and Section 4 describes the data. Section 5 describes the empirical specifications and identification. Section 6 covers reduced form results. We conclude with a discussion in Section 6.

2 Setting and Institutional Details

The data in this paper comes from a large white-collar technology and services company. Employees at this firm are mostly software engineers, product managers, sales, marketing, and support staff. At this firm, promotions are a major component of incentives and compensation. After every quarter of performance, employees were evaluated subjectively and assigned a score on a one to five scale. Carlson (2014) describes a similar, five-point employee evaluation process used at Yahoo. To limit managerial favoritism, scores were decided by a committee using input from the manager and peers.

Career progress and promotions at the firm were managed using job levels. Within each type of job, employees occupied a level which generally spanned from one (most junior) to nine (most senior). A promotion constituted a vertical move upwards on this ladder and was associated with a permanent increase in base salary of roughly 20%. This increase in base salary was the primary benefit of being promoted. Although the firm set higher performance expectations for each level,[9] duties often do not significantly change after promotions. The promotion system at this firm is thus mostly about providing incentives, and not selecting the best candidate for a different role.

Promotion decisions were made twice per year. To initiate a promotion, a worker must agree to be nominated. Workers can self-nominate, or be nominated by a manager. The manager's support is a valuable but not required condition for promotion. Like subjective performance evaluations, the promotion decision is made by a committee to avoid man-

[9]For each combination of job level and type, the firm published descriptions of expected responsibilities and contributions for employees to review.

agerial bias and conflicts of interest. According to analysts at the firm, a small number of employees succeed in gaining promotion without their manager's support.[10]

On average about 20% of employees are nominated per cycle. The committee makes decisions on the basis of three primary sources. The first source is a collection of letters of evaluation from peers and managers. The promotion candidate and his manager can decide who these evaluators are. The second source is a statement by the candidate, outlining his or her accomplishments and case for promotion. The third source is the employee's history of numeric and written performance reviews.

The numeric performance scores were not mechanically linked to promotion decisions through a rigid, inflexible formula.[11] Nonetheless, these subjective performance metrics represent the closest empirical analogue to the "scores" used in Siegel (2009) and other contest theoretic models to select winners and losers in tournaments.

Company policy requires that at most 10% of all workers be promoted. In the data, the 10% figure is exactly realized nearly always. This policy is the basis for the characterization of the incentive scheme as a contest or tournament.[12]

Why did the firm organize a tournament? In interviews with managers, the main reason provided was a concern about collusion between middle managers and employees. "If we didn't limit promotions," said one HR professional, "Then managers would simply promote everyone, as a means of buying the cooperation and loyalty of their workers at the expense of the company."[13] In addition, some employees mentioned the simplicity of a 10% rule. This required only that managers involved in promotion decision simply rank the promotion candidates and select the top 10%, rather than decide both the threshold and ranking (or making compensation decisions for each employee).

Table 2 describes the relationship between performance scores and promotions. All regressions are linear probability models of the promote/not decision. In Panel A, we see that worker fixed effects and tenure controls alone have a large explanatory power over promotion decisions. With both sets of fixed controls included and nothing else, R^2 is 0.37.

[10] This paper does not include data about nominations or their sources, so I cannot report how many.

[11] In other settings, companies have implemented a formulaic relationship between subjective performance scores and promotion contests. For example, Carlson's 2014 account of Yahoo's employee ranking system says that, "Eligibility for bonuses, promotions, and transfers within the company would be based on their average score for the past four quarters. *[...]* Under the new system, the only way to get a promotion or raise at Yahoo was to have an average score of three for the past four quarters."

[12] I interviewed managers at the firm to question the characterization of these incentives as contests. Given several potential shortcomings of tournaments, would the firm truly create a contest-like promotion system? "These are most definitely contests,"said one HR representative.

[13] This differs from the standard reason for relative performance evaluation, which was to remove common productivity shocks (Green and Stokey, 1983).

In Panel B, we see the role played by the subjective performance scores. A single standard deviation increase in subjective performance in a given quarter increases the probability of promotion by 5%. Performance scores from prior quarters are also positively correlated with performance scores. However, once worker fixed effects and tenure controls are included, only the most recent performance score has significant weight.[14]

Table 2, Panel C compares two transformations of the performance score. The first is a normalization of the performance score, and the second is the percentile rank of the scores. Contests and auctions are decided by discontinuities in ranks, and indeed the percentile version of the regressions have the highest explanatory power. This is particularly true when the regression includes a dummy when an employee's score is above a certain percentile. A flexible algorithm chose the 87th percentile as the likely location of a kink. This is very close to the threshold of 90% implied by the firm's institution of promoting 10% of workers per cycle.

These findings are consistent with the characterization of the firm's promotion system as a tournament in which the top 10% are rewarded with promotion. In the final panel D, we see results about the scope of comparison in promotion decisions. Performing well compared to employees within the same business unit is a significant predictor of promotion. However, there is additional predictive power of performing well compared to local peers who have the same manager.

The remainder of this section discusses how common promotion systems like this firm's are (2.1) and a few comparative features of the firm's model (2.2).

2.1 How common are workplace tournaments?

How common are contest-like systems for promotions, raises and other real-wage increases? The consensus among economists and business scholars appears to be that workplace tournaments are widespread.[15] The practice of workplace competition has several high-profile advocates in business[16] who refer to the practice using various nicknames.[17]

[14]Table 2, Panel B looks only at the first lag, but similar qualitative results exist for the second, third and fourth lags.

[15]For example, Prendergast (1999a) writes that promotion tournaments are, "Perhaps the most common means of rewarding white-collar workers for effort," and asks "[W]hy are they so popular?"

[16]This includes former CEO and Chairman Jack Welch, Former Microsoft CEO Steve Ballmer, current Yahoo-CEO Marissa Mayer, Enron executives Jeffrey Skilling and Kenneth Lay as well as executives at McKinsey & Co. (Michaels et al., 2001).

[17]For example, the tournament systems are sometimes called "rank-and-yank" (Gladwell, 2002), "force-rank" (Blume et al., 2006), "stack-ranking" (Eichenwald, 2012), "the vitality curve" (Welch and Byrne, 2001),

A list of well-known firms currently or previously using tournaments (according to journalistic and academic sources) is reported in Panel B of Table 1, and includes a mix of old- and new-economy firms.[18]

Despite the growing theoretical and empirical literature about contests – much motivated by workplace tournaments – data on the abundance of promotion contests is rare. No paper attempts to measure the empirical prevalence of contests as employee incentives.

This paper attempts to fill that gap in that data by conducting two surveys of over 15K Americans. Results (reported in Table 1) suggest that "Performing well compared to peers at job," is the most common reason for a real-wage increase, and is almost twice as common as earning a real-wage increase from leaving for a new employer, enjoying a market wide wage increases or asking an employer to match an outside offer.

The results also suggest that over 77% of Americans work in a company with a tournament-like system of promotions that feature restrictions on the number of candidates who can be promoted.

2.2 Important Details and Comparisons to other firms' tournaments

An important detail of this firm's tournament system are the consequences for employees in the bottom of the distribution. At Microsoft, Enron, GE and many other firms using contests (see Section 2.1), low-ranking employees were subject to harsh punishment such as dismissal, demotions or sometimes pay reductions.

By comparison, consequences for low-ranking employees in this paper were milder. Such employees were denied wage growth, income and career opportunities. Low rankings could have reputation effects, particularly because the ranking would become part of the employee's permanent personnel record that can be viewed by future managers at the firm. However, workers were not automatically fired or reduced in pay. Some

"differentiation" (Welch and Byrne, 2001), "the 70/20/10 rule" (Welch and Byrne, 2001), "performance review committee" (Thomas, 2002).

[18]Because of negative publicity from media stories such as Eichenwald's 2012 study of Microsoft's promotion system, some firms are now denying the use of "yank-and-rank" or claim to have abolished it. The author has investigated, and it appears that these firms have not entirely abandoned rank-based incentives. Instead, they have weakened the incentives at the bottom of the distribution, for example, by no longer automatically firing the lowest ranking workers and providing a weaker punishment instead. Meanwhile, incentives at the top of the distribution remain strong and provisioned by rank-based methods. As discussed in Section 2.2, this distinction makes no difference for the behavioral predictions of standard contest theory models.

workers were able to recover from low rankings. Carlson (2014) writes of Yahoo, "Employees ranked poorly would see their lives materially turn for the worse as they lost out on raises, promotions, and bonuses. *[...]* Employees with low enough scores would be asked to leave the company."

From the lens of standard contest theory models, framing rewards as gains or avoided losses makes no difference. The theoretical predictions regarding effort, sabotage and other variables are the same however the problem is framed, so long as utility is distributed via the tournament mechanism. However, journalistic accounts of Microsoft and GE's method by Eichenwald (2012) and others suggest these punishments were especially motivating, possibly because of loss aversion (Kahneman and Tversky, 1984). We discuss our results in light of this difference in the conclusion (6) of this paper.

A second detail is that the 10% budget is within business units consisting of workers of many different ranks and types. This system creates competition between dissimilar employees for promotion slots, such as junior and senior employees. Although lower level employees face lower performance expectations, the firm's institutions place them into competition with senior employees through the 10% quota. A high-performing senior employee may face competition from high-performing junior employees even if he outperforms his other senior colleagues.[19]

3 Identifying a Causal Effect of Competition

Despite creating a contest for promotions, the firm paid little attention to contest-theoretic issues when assigning workers into larger departments and teams. The firm did not make a conscious decision to aggregate workers into units based on optimal levels of intra-worker competition.

Doing so would have been very hard to achieve, even had the firm attempted it. Many of the firm's employees were new, and the firm lacked much experience with these employees to use as the basis for an optimal match for a new employee (about whom it knew even less than the current employees).

Interviews with the human resource professionals confirm that contest-theoretic issues – even informally understood – were not the basis of consideration for the formation of business units. These issues were ignored both with regards to initial unit assignments as well as reorganizations. By contrast, matching decisions were made somewhat hap-

[19]This is not only an issue in theory – one senior employee interviewed discussed concern about his own prospects based on the strong performance of junior employees.

hazardly. A discussion of how and why the firm (and its recruits) would accept such haphazard matching follows later in this section.

A common way of assigning new employees to units at the firm was based on date-of-hire. In this method, the firm assigned employees in batches on certain days. Workers were given wide latitude to select the timing of arrival, and the timing was generally affected by landlord, spousal, vacation and relocation-related preferences. The mapping of dates to business units was unknown to the firm itself until a few days in advance.

As such, similar employees who arrived in slightly different dates would be assigned to different units and different contemporaneous (and future) competition. Even if competition dynamics were perfectly observable and predictable, this mechanism made it very hard for employees to strategically time arrival to join a particular group. The discussion and analysis of identification in Section 5.2 shows that arrival date was indeed a strong predictor of unit assignment.

We see very little evidence of selection, or workers endogenously conspiring to manipulate assignments. As discussed below, predicting worker performance is difficult and would be necessary for strategic matching on the basis of contest considerations. However, to account for undetectable types of endogeneity, the results in this paper include instrumental variable specifications attempting to identify estimates from the date-of-hire variation. The resulting estimates are not qualitatively different than the OLS results, and the estimates' confidence intervals mostly overlap.

The remainder of this section discusses the practice of assigning workers to business units based on the date-of-hire. Why would a firm organize itself this way (3.1)? And why would job candidates agree to job offers with uncertain assignments (3.2)?

3.1 Reasons for the firm's assignment mechanism

It may seem odd that the firm attempted so little optimization of employee assignments. A rudimentary attempt at optimal matching could improve welfare for both the firm and its workers. If workplace competition indeed has the effects documented in this paper, it is strange that the firm wouldn't try to exploit them more directly through contest design.

A rudimentary form of optimal matching could also improve welfare for candidates, and thus the firm's recruiting. In many companies, job candidates know their potential destination inside the firm at the time they accept/reject a job offer. However, to quote from an internal document from the firm, "Candidates were hired into [*Company*] without knowing their intended project focus."

There are several reasons for the firm's assignment institution. We discuss a few below. The first reason is that the firm's HR function was capacity constrained. The firm was growing rapidly in nearly every department except HR. The firm lacked a mechanism for performing the match, that didn't involve HR employees reading resumes and attempting a match.

The firm's and industry's fast growth created other problems for optimizing matching. The company's executives were philosophically opposed to "overfit" job assignments. They anticipated technological change in the product market that would change demands on employees' work.

The resulting recruiting and hiring strategy placed a heavy emphasis on general human capital, rather than task-specific human capital. This emphasis was meant to avoid "overfit" candidates who may have been a great match for current needs in a particular department, but who may have been unable to adapt. Hiring was therefore centralized. The centralization was meant to remove hiring from the hands of business units, who were likely to hire based on short-term rather than long-term needs. This philosophy is not uncommon in the technology industry. For example, a management book by Google executives Schmidt and Rosenberg (2014) says, "The urgency of the role isn't sufficiently important to compromise quality in hiring."

This is clearly not a firm that believes strongly in the benefits of optimal short-run matching of employees to business units.

Lastly: Organizing workers into optimal contests would be hard to implement for reasons anticipated by the employer learning literature (Kahn and Lange, 2014; Radner, 1992). Organizing workers optimally would have required advanced knowledge of the long-term career trajectories of employees at the time of hire and assignment. This would be difficult under any circumstances, but especially because the firm's production technology was new and different than others' in the industry. As such, inferences made on the basis of past performances would have been less robust.[20]

Nonetheless, this firm repeatedly attempted to make these inferences through analysis of employees' interview scores (taken from the time of hire) with realized performance on the job after being hired (measured by subjective performance scores and other metrics). Working independently and together, the firm's own statisticians were unable to find any robust relationship. Outside statisticians were also consulted and were also unable to find a predictive relationship.

An executive from the firm's industry discussed similar research findings in a *New York*

[20]The relationship between technological change, obsolescence and labor markets is a well-chronicled issue in the technology industry. See, for example, Brown and Linden (2009).

Times interview in 2013 (Bryant, 2013), saying:[21]

> "Years ago, we did a study to determine whether anyone at [our company] is particularly good at hiring. We looked at tens of thousands of interviews, and everyone who had done the interviews and what they scored the candidate, and how that person ultimately performed in their job.
>
> We found zero relationship. Its a complete random mess[.]"

These findings are consistent with a large literature in industrial psychology summarized by Macan (2009), which show the low- or non-existent predictiveness of interviews on job performance.[22]

The imperfection of employer learning and predicting employee quality has also been a theme in economics. Slow employer learning is a leading hypothesis for the correlation between wage dispersion and age/experience. According to this line of research, wage dispersion increases with tenure because employers need time to distinguish high- and low- talent employees.

This literature suggests that the employees in this study have exactly the age and experience profile that would make them most difficult to organize into optimally designed contests.

Employer learning is further made difficult because employee productivity evolves over time. Kahn and Lange (2014) summarize and extend this literature, finding that employer learning continues throughout the life cycle and that, "Imperfect learning [...] means that wages differ significantly from individual productivity all along the life-cycle because firms continuously struggle to learn about a moving target in worker productivity." This would make organization into optimal contests hard, even for higher-tenure employees.

Firms' struggle to estimate worker productivity is, in some sense unsurprising. At their core, worker measurement issues are causal inference questions ("How much did worker X 'cause' outcomes to improve?"). Firms rarely create experiments necessary for causal inference at an individual-level. Even if they did, they would often lack statistical power due to the extreme noisiness of outcomes.[23]

For example, some economists have suggested sales as a white-collar job whose output can be easily measured in dollars. However, most firms do not experimentally rotate

[21] URL for interview and article: http://www.nytimes.com/2013/06/20/business/in-head-hunting-big-data-may-not-be-such-a-big-deal.html?pagewanted=all. Last accessed October 24, 2014.

[22] Macan (2009) reports particularly poor results for "unstructured interviews," in which interviewers are relatively free to lead the discussion in any direction. The firm in question used unstructured interviews.

[23] Team production also complicates measuring individual-level talent.

salespeople to measure the causal effect of individual salespeople. Furthermore, many firms assign entire teams of salespeople to important clients, thus confounding the measurement problem.

Even if firms ran such experiments, sales data is extremely noisy. Lewis and Rao's 2012 study of advertising shows that even optimally-designed sales experiments would lack statistical power and fail to produce useful standard errors. In a labor productivity context, Barankay (2012) mentions the large variance in sales performance as an obstacle to designing experiments with firms.[24]

3.2 Reasons the uncertainty in assignments was acceptable to job candidates

For my instrumental variables approach, I show in Table 4 that employees would not have known what start date would get them assigned to more or less competitive groups. Variables in Table 4 known at the time of hire are uncorrelated with my instrument.

A separate question is why a worker would agree to join a firm without more details about his/her job assignment. During the sample period, the firm enjoyed a strong reputation as an employer, and experienced positive balance sheet growth.

As such, the firm was able to recruit effectively without making promises about job destinations. One example of this strategy came from Facebook COO Sheryl Sandberg. In 2012, Sandberg told Harvard Business School graduates the story of her recruitment to Google in 2001. She discussed her unease about the role she was offered, and Google CEO Eric Schmidt allegedly replied:

> "Don't be an idiot. [...] If you're offered a seat on a rocket ship, don't ask what seat. Just get on."

Sandberg forwarded this advice to the Harvard graduates at her speech.[25] The strategy worked well for Sandberg, who accepted the Google's offer and was a sales executive

[24] Barankay (2012) writes, "For reasons of statistical power, I had to span the experiment due to the large variance in sales performance: As I wanted to be able to test the difference across treatments and by gender, I could only have three treatment groups per year to achieve the required statistical power."

[25] Text of the speech: http://www.businessinsider.com/sheryl-sandbergs-full-hbs-speech-get-on-a-rocketship-whenever-you-get-the-chance-2012-5.
Video of the speech: https://www.youtube.com/watch?v=2Db0_RafutM.
Both last accessed October 24, 2014.

from 2001 to 2008 during a period of strong growth.[26]

She became a multibillionaire in Facebook's 2012 IPO. Sandberg and Schmidt's stories provide evidence of how and why job applicants may accept ambiguous job offers in order to join firms with strong reputations and growing balance sheets. Schmidt summarizes:

> "When companies are growing quickly and they are having a lot of impact, careers take care of themselves. And when companies aren't growing quickly or their missions don't matter as much, thats when stagnation and politics come in."

4 Data

The data from the firm includes roughly 60,000 workers over the period of 2000-2009. The workers are mostly sales and engineering staff who have a high level of education. The median salary is over $100,000 per year.

For the analysis in this paper, I organized the data into a quarterly panel of individuals. Unless otherwise stated, the variables below are aggregated at the individual $\times$ quarter level. I organize. this section into three sections: Competition measures (4.1), outcome measures (4.2) and controls (4.3).

4.1 Measures of Competition

The objective of this paper is to measure how changes in competition for promotion affect worker behavior. As such, measuring the level of competition intelligently is important. Unlike hard information such as lines of code written or reviewed, competition represents a more abstract and amorphous concept to be handled circumspectly. Rather than viewing a single competition measure as "right," and the others as wrong, a comprehensive approach seems more fruitful and convincing. Below, I describe my preferred measure of competition. However, recognizing the potential for disagreement, the online appendix reproduces the analysis with several plausible alternatives. These other measures produce the same qualitative results, so little of what follows hinges on the the quirks of this specific measure.

[26]She later repeated the strategy. In 2008 Sandberg turned down offers to become CEO of other companies (including *The Washington Post*) to accept a COO job, reporting to a 23-year old. She once again accepting restrictions and uncertainty on her responsibilities and job title in order to join a faster-growing firm.

The primary measure of competition used in this paper is derived from subjective performance scores of workers in the previous quarter. Employees whose peers score similarly are coded as facing high competition, and employees whose peers score very differently (either higher or lower) face lower competition.

A formalization follows in Equation 1. I will first offer a defense of this measure motivated by the literature on formal contest theory models. These models feature a collection of rivals, each with a parameter representing their talent (or other form of advantage) that affects their odds of winning. Workers then choose effort and create output, which is observed with noise.

In this setup, homogeneity in the odds of winning (through equal talents or other handicaps) motivates effort. In a battle between unequal rivals, neither has any incentive to work. The stronger will prevail with minimal effort, even if the weaker supplies maximal effort. Seeing this in advance, neither side chooses to work. By contrast: If rivals are roughly equal, then the returns to effort are greater and thus agents supply more of it.

My measure of competition is thus based on subjective performance scores, which are analogous to each employee's ex-ante probability chance of promotion. As discussed in the institutional details (Section 2), subjective performance scores are closely tied to promotions. This is shown in the data in Table 2. One of the primary reasons the firm scores candidates is to guide the promotion decision. The performance score can be interpreted as an employee's proximity to promotion.

To measure competition, I thus calculate how much each focal worker is equidistant to promotion compared with his/her peers. If they are equidistant, and if only a finite number can be promoted, then this would place the peers into greater competition with each other. At equidistance, an increase in effort would have a greater impact on the probability of success than if peers are disperse for the reasons discussed in my example above.

The measure is formally defined below. An employee i working in business unit j in period t would face competition defined as follows:

$$x_{i,j,t} = -1/N \sum_{k \in j, \neq i} |PerformanceScore_{i,t-1} - PerformanceScore_{k,t-1}| \quad (1)$$

... where $k \in j, \neq i$ refers to the other workers in unit j besides i. This formula codifies the concept of the average absolute distance in the previous quarter between an employee's subjective performance score and those of his or her contest rivals.

I use a lagged measure in order to capture the competitive standing at the last time the firm formally scored candidates before choices are made in the current period. Would

workers know their competitive standing? I address this question in 4.1.2, below.

Using the specification in Equation 1, a low distance implies lots of close competition and strong incentives. A high distance implies little competition or incentives.

The rest of this section proceeds as follows. First, I compare my measure of competition and incentives with a few others used in the empirical contest literature. Next, I discuss how workers might know their competitive standing (4.1.2). Lastly, I discuss measurements of counterfactual competition. This measure of counterfactuals will assist in the instrumental variables strategy discussed in Section 5.2.

Lastly, several alternative measures of competition are also possible based on job levels and other functions of performance scores. I study these for robustness, and these are defined and discussed in the online appendix.

4.1.1 Comparison to Other Measures of Contest Intensity

Much of the extant empirical work on contests exploits naturally occurring or experimentally induced variation in prize structure or prize allocation rules. By contrast, variation in incentives in this setting derives from changes in the composition of contestants.

In the empirical contest literature, various other variables have been used to measure the strength of competition and incentives. One common measure of incentive strength is to utilize variation in the size of the contest prize (Garicano and Palacios-Huerta, 2005; Ehrenberg and Bognanno, 1990a).

Rather than studying variation in contest size, this paper studies variation in the composition and relative strength of opponents. In this way, the paper is similar to Brown's 2011 study of contestants in golf tournaments. However, my measure is different from Brown's 2011 competition variable in that Brown (2011) was entirely about the presence or absence of a single superstar contestant (Tiger Woods). This is partly because Brown's 2011 paper is motivated by a different research question regarding superstar effects. Even within the setting of professional golf, it is clear that incentives can also come from the proximity of non-superstar opponents who may be able to offer competition. This is true in a workplace setting as well.

In addition: In a workplace setting, single superstar opponents are less important sources of competition. This is because the contest allocates more than one identical top prize (in our firm's case, 10%). This widens the space for the competitive influence of non-superstar opponents. In fact, the strongest competition may be in fact near the 10% percentile threshold – very far from Brown's 2011 elite swinger.

4.1.2 Do Workers Know their Competition?

A necessary condition for competition to motivate is that workers know the degree of competition they face. A worker knows his or her own performance but is not privy to the performance ratings of others, so these must be inferred. Is such inference likely, or even possible?

Workers have many ways in which to evaluate peers' performance and relative positions. Since expectations and accomplishments are often announced, acknowledged and celebrated at regular meetings and shared informally, workers know who contributed significantly within their team. Such "wins" are key drivers of performance scores.

Many of these accomplishments directly affect the work of other peers. Thus it is not possible to hide significant accomplishments or lack thereof. If a software engineer develops a new feature for his/her product, other engineers will have to interact with that engineer and his feature in order to integrate it. In the process, these peers will be able to observe the quality of the contribution and how it was received by its consumers. Similar dynamics exist in a sales context. If a salesperson has recruited a large client, others will be assigned to that client to assist in the additional work of servicing the account.

In addition, the firm has a system of quarterly tactical planning in which quantifiable goals and deadlines are set in unit-wide meetings along with owners. Explicit performance expectations are set in these meetings. For example, the minutes of one such meeting include the statement: "By January 2003, we aim to launch a feature codenamed XYZ that will enable clients to accomplish X with a single click. [*Employee Names*] are in charge of this implementing and launching this feature by January 2003."

These goals are quantitatively scored afterwards in similar unit-wide meetings. While not identical to performance scores (since they are scored by goal rather than by individual), they are nonetheless indicative of the responsible individual's standing among his peers.

There are other direct ways to evaluate peers' performance. As discussed in Section 4.2, the "objective" performance data in this paper are taken from internal production infrastructure. These systems log worker activity mainly to facilitate followup, debugging, collaboration and auditing by peer workers.

Workers may readily observe these measures of peers' activity. Although they may not directly know how these measures are interpreted by management, this does give additional insight into how competitors are performing. Regressions in table X show that these "objective" measures are correlated with subjective performance and promotion outcomes.

A final way that employees know their relative standing is through direct sharing and discussion. Anecdotal and interview evidence suggests that workers discuss impressions of colleagues' career prospects, and sometimes share private information about their own evaluations. Such discussions are much like those that Lewis (1989) describes as common in the finance industry.

The setting is not unlike academia: Although formal feedback on tenure progress is private, colleagues can form impressions based on publicly available signals. These may include conference and/or journal acceptances, editorships, refereeing, conferral with others and direct evaluation of research and teaching. In addition, some colleagues share their private information in ways that partially or fully reveal competitive standing information.

4.2 Outcome Measures

The most novel data in this paper are the variables relating to the many dimensions of productivity and effort. The nature of production at this firm generates a large amount of data. Employee use of the firm's production infrastructure is often tagged and timestamped with worker-identifying information. The firm logs this data primarily for debugging and followup purposes,[27] and not for performance evaluation.[28]

My data includes roughly 18 measures of output from the activity logs. Below, I give a sample of outcome measures for the two most common job categories (software engineers and sales or marketing professionals). The full set of underlying variables are described in detail in the online appendix. Next, I itemize how these measures are summarized into measures of interest to the literature.

For software engineers, important measures of productivity include lines of code submitted or changed in the repository, the amount of compiles and the number of code reviews performed for other employees. For sales and support staff, important variables include data how often each employee contacts clients, and how often employees perform

[27] For example: Additions and modifications to the firm's software corpus are tagged and timestamped. Through this setup, managers and employees are able to identify the original authors of code or documents that need followup. The firm's production systems (such as its software compilers) also track usage. Should one of these systems temporarily break, these logs are useful for debugging.

[28] The firm's employees and managers are opposed to using much of this data for performance evaluation because of the possibility of gaming. Using the example in the previous footnote: If the firm evaluated performance based on lines of code contributed to the codebase, then employees would have incentive to contribute lengthy, inefficient code.

small tasks for clients such as reviewing and approving requests for account changes.[29]

These underlying variables are summarized into into several outcome measures in a worker $\times$ quarter granularity. These are generally the dependent variables in regressions described in Section 5. These variables are discussed below.

4.2.1 Output and Effort

Dating back to Lazear and Rosen (1981), economists studying contests have primarily focused on "choice of effort" in theoretical and empirical research (CITES). We measure these choices in the following way.

Hours Worked. As previously mentioned, use of the company infrastructure is often logged with worker-IDs and timestamps. To measure hours of work, I count unique hours with a timestamp.

This measure has two benefits compared with typical measures of working hours used elsewhere. First: It is not-self reported. Second, it does not include working hours in which the worker may be physically at work but not doing anything.

This measure no doubt excludes some productive activity while at the firm that cannot be logged. To control for the possibility that these holes in the data affect some groups differentially, my econometric specifications include fixed effects for individual workers, time-periods, job types, job levels and business-units.

In addition, in some specifications I measure hours as the difference between each day's first and last hour in local time. When this version of the variable is used, it is noted in the table.

Total Output. Total output is measured as the sum of all activities across all productivity dimensions. To ease interpretation of coefficients, I standardize the variable in regressions by subtracting the mean and dividing by the standard deviation. To account for certain jobs and business units working more on particular activities, my specifications include worker-, job-level-, job-type- and business-unit fixed effects. Additional detail on these specifications is in Section 5.1.

[29] For both of these measures for the sales staff, there is enough demand for this labor such that workers never run out of tasks.

4.2.2 Efficiency and Quality

Promotion contests have been criticized for encouraging "influence activities" (Milgrom and Roberts, 1988; Milgrom, 1988, Gubler et al., 2013) rather than productivity. An example of this comes from Eichenwald's 2012 journalistic account of Microsoft's promotion system:[30]

> "The best way to guarantee a higher ranking, executives said, is to keep in mind the realities of those behind-the-scenes debates– every employee has to impress not only his or her boss but bosses from other teams as well. And that means schmoozing and brown-nosing as many supervisors as possible."

Similarly at Enron, Bodily and Bruner's 2002 writes that "[R]ank-and-yank turned into a more political and crony-based system." Carlson (2014) of Yahoo, "As employee ratings got passed up the management ladder, individual scores sometimes had to be adjusted at the department level so that the right amount of employees were in each bucket. This lead to favor trading between managers. It also meant that employees felt they had to brownnose their boss' peers and their boss' boss."

Even if workplace competition increases worker effort, much of that effort may go into lobbying and influence activity. As Eichenwald (2012) describes Microsoft:

> "I asked Cody whether his review was ever based on the quality of his work. He paused for a very long time. 'It was always much less about how I could become a better engineer and much more about my need to improve my visibility among other managers."'

In the data, such activities would appear as decreases in per-hour efficiency or quality. Hence, I address these hypotheses through the following variables:

Efficiency: The ratio of output to hours. This is also reported in normalized terms. Regressions with efficiency as the outcome variable include controls described above.

Quality: The amount of code submitted to the code repository that was subsequently, in future periods, withdrawn or reverted because of problems or bugs discovered in the code.

[30] http://www.vanityfair.com/business/2012/08/microsoft-lost-mojo-steve-ballmer

4.2.3 Cooperation and Sabotage

Promotion contests also potentially create incentives for sabotage and other forms of anti-social behavior, as well as discouraging productive cooperation between employees. This possibility has been much studied theoretically (Dye, 1984, Lazear, 1989 and Rob and Zemsky, 1997), and memorably dramatized in the fictional real-estate sales tournament in *Glengarry Glen Ross*. While David Manet's play illustrates an extreme form of sabotage, more subtle forms of sabotage seem to be common at some companies. For example, Eichenwald (2012) writes of Microsoft:

> "Staffers were rewarded not just for doing well but for making sure that their colleagues failed. As a result, the company was consumed by an endless series of internal knife fights. [...]
>
> People responsible for features will openly sabotage other people's efforts. One of the most valuable things I learned was to give the appearance of being courteous while withholding just enough information from colleagues to ensure they didn't get ahead of me on the rankings."

Similarly, Carlson (2014) wrote that at Yahoo, "Workers would prioritize tasks that got them closer to their personal goals over doing anything else. This made sense. Collaborating and helping out on a project that wasn't going to get you close to an 'exceeds' was just a stupid thing to do."

The empirical literature on sabotage has mostly taken place in labs (for example, Carpenter et al., 2010; Harbring and Irlenbusch, 2011). The rare exceptions from outside the lab include studies of professional soccer by Garicano and Palacios-Huerta (2005) and judo by Balafoutas et al. (2012). Drago and Garvey (1998) use survey data from Australian manufacturing to shed some light on cooperation – the survey results show that when promotion incentives are strong, workers are less likely to provide "helping" behavior such as allowing others to use their equipment.

In this paper, we lack direct measures of sabotage. However, we have several measures of productive cross-worker cooperation and collaboration. These are:

Peer bonuses. Employees at the firm are able to reward each other with roughly $200 bonuses for excellent, otherwise unseen or unrewarded contributions or assistance. These rewards are accompanied by a laudatory email to the recipient's manager and sometimes the executive and/or entire group. The outcome metric used in my regressions is the

number of peer bonuses received by each employee in each quarter.[31]

Writing documentation. Inside a software company, documentation of the firm's infrastructure is an important facilitator of collaboration and teamwork between employees. Using documentation, a software engineer can quickly gain the knowledge necessary to contribute to a co-worker's project.

4.2.4 Innovation

Incentives for innovation are also a longstanding theme in the contest literature. One of the primary applications for contests to encourage innovation (for example, Baye and Hoppe (2003) and Fullerton and McAfee (1999)). A related literature shows the power of specially-designed contests for innovation and research (see for example Boudreau et al. (2011)).

In theory, incentives based on output might give employees incentives to invest in projects. In a slightly different context, Aghion et al. (2005) discusses the "inverted U-shaped" relationship between competition and innovation, with an interior optimum.

However, the same factors that lead workers to produce "influence activity" instead of productive output may also lead them to eschew innovation in favor of more visible forms of output. Azoulay et al. (2011) show that weaker incentives allows longer term investment that fosters high impact and creative work. As stated in Eichenwald's 2012 account of Microsoft's promotion system:

> "[B]ecause the reviews came every six months, employees and their supervisorswho were also rankedfocused on their short-term performance, rather than on longer efforts to innovate."

To address this topic, our data contains many measures of innovation.

Patenting: Employees at the firm are rewarded bonuses for submitting ideas that lead to a patent application. Most of these applications lead to a patent, but the evaluation process at the government usually takes several years. This data on patenting productivity data comes from the firm's internal record of who made the patents.

Contributions to ideas board: In order to solicit and evaluate new product, business and feature ideas, the firm created an internal "ideas board" to collect innovative proposals from employees. This application was accessible via a web browser, and allowed

[31] My data does not specify who the bonus was from. Although it is possible for peer bonuses to be given to worker from someone outside of his or her set of contest opponents, the firm says this is relatively rare.

employees to suggest new ideas and receive ratings and feedback on them. I include not only the number of these ideas as an outcome variable, but also their average rating.

4.2.5 Specialization and Differentiation

Specialization and differentiation are relatively unstudied aspects of the literatures on contest theory, organizations and specialization. However, the relationship between incentives and specialization is discussed somewhat in Holmstrom and Milgrom (1991a), and in a variety of other analogous setting such as MacDonald and Marx (2001).

My calculation of worker specialization uses the a formula for the Herfindahl-Hirschman Index ("HHI") from the IO literature on market concentration (Hirschman, 1945; Herfindahl, 1950; Hirschman, 1964).[32] The formula is below.

$$Specialization_{i,t} \in (0,1] = \sum_{l \in L} s_{ljt}^2 \tag{2}$$

... where $s_{i,l,t}$ refers to the share of the worker's total activity counts from activity in dimension l, and L refers to the set of all dimensions/activities.[33] The variable is has an upper bound of 1, representing complete specialization into one activity.

For differentiation, I create a measure of how similar workers' outputs are to each other. To implement this, I represent each worker's productivity as a point in multidimensional Cartesian space, and measure the average distance between the focal employee and his or her contest opponents. My preferred measure of distance is Euclidean, although many alternative measures are possible.[34]

$$Differentiation_{i,t} = \frac{1}{N} \sum_{j \in J} \sqrt{\sum_{y \in Y} \left(p_{i,t,y} - p_{j,t,y} \right)^2} \tag{3}$$

As with the variables above, these variables are normalized to ease interpretation. Regressions include controls to account for different roles. Alternate measures of specialization or differentiation produce very similar results.

[32] The measure is known as the "Simpson index" in population biology (Simpson, 1949) and is known as the "participation ratio" in physics (Eliazar and Sokolov, 2010).

[33] For example: Suppose a worker performed 4 total activities activities in a month. One of the activities was a code review, and the other three were patents. This worker would have a specialization score of $0.625 = 0.25^2 + 0.75^2$.

[34] Results from alternative definitions are available by email.

4.2.6 Sorting and Selection

Finally, an important function of contests is selection and sorting. The nature of competition may not only effect how hard employees work, but also whether and where they choose to work at all.

The limited theoretical research on this by Morgan et al. (2012) shows that in large contests, high-types clustering together, despite the stiffer competition. However, this contradicts the intuition that high types will want to separate in order to mitigate competition, and thus each become the "king" of a smaller, local hill. Morgan et al. (2012) shows this behavior is possible.

The selection and sorting aspects of workplace competition are an important practical consideration in contests. Carlson (2014) writes of Yahoo:

> As 2013 rolled on, Mayer's system made life particularly difficult for Yahoo's middle managers. It was hard to get talented people to work in the same group. Not only did people not want to compete against other talented employees, they also worried that if they transferred in the middle of a quarter, they'd whiff on their goals, get a mere "achieves [expectations, a type of rating]," and lose out on a chance for a raise anytime in the next twelve months.

My primary measure of sorting decisions is quitting, which I can measure for all workers as well separately for different levels of subjective performance scores.

4.2.7 Subjective Well-Being and Job Satisfaction

The final set of outcome variables comes from a confidential job satisfaction survey asked by the firm's HR department. This happens only once in the middle of our sample, and is only available for a subset of employees who answered the firm's survey. The survey generally measures satisfaction with the company, one's job and manager and peers.

4.2.8 Use of the Outcome Measures by the Firm

The activity logs I use for measurement are common practices in software engineering and other knowledge-based work. Such logs are proliferating as more production takes place through remote "cloud" servers, even for word-processing and other non-engineering tasks. For software engineering, these logging features are offered in the most popular version control systems and are used throughout the software development industry. Other

scholars (for example, Lerner et al., 2006) have used similar data on code contributions to study the economics of open source software communities.

Workers in these industries generally know that their productive activities are logged for debugging and auditing purposes. In interviews at this firm, workers did not consider this logging as performance measurement at all, but rather as useful productivity, development and coordination tools for an inherently team-based production environment. Much of the logging is deliberately made available to other workers at the firm in order to assist collaboration. For a given piece of code or document, it is easy to see who altered which parts and when. In theory, workers could compile these logs into the daily aggregations about each other, but this would require considerable additional effort without any clear benefit.

Importantly, this data is generally not used for performance evaluations and never contracted upon. In fact, most of this data was in fact created and systematically compiled by this author for this research agenda.

The idea of using these metrics for performance evaluation is generally repugnant to workers, managers and HR executives in white-collar industries.[35] Incentivizing metrics such as "lines of code" or "documents edited" would create perverse incentives for inefficient writing and software.

Although the firm was not systematically aggregating this data, the metrics in this paper do capture aspects of performance that are observable to the managers. For example: Although the firm did not collect data about hours of work for employees at the firm, individual managers would be aware of it through direct observation. In Section X, I show that the measures I collected about worker effort and performance are predictive of subjective performance reviews and promotions.

4.3 Controls

Much of the data in this paper are typical personnel-economics variables from inside the firm. This includes the role, rank and salary of each worker. [Unfinished]

[35] In Cringely's 1996 1996 documentary *Triumph of the Nerds*, Microsoft executive Steve Ballmer described IBM contracting on "lines of code" as a performance metric: "In IBM there's a religion in software that says you have to count K-LOCs (and a K-LOC is a thousand lines of code). [...] IBM wanted to sort of make it the religion about how we got paid. And we kept trying to convince them - hey, if developer's got a good idea and he can get something done in 4K-LOCs instead of 20K-LOCs, should we make less money? Because he's made something smaller and faster? Less K-LOC? Ugh! Anyway, that always makes my back just crinkle up at the thought of the whole thing." Ballmer's animated retelling of this story appears in the the second episode of the series at 38:56 (https://www.youtube.com/watch?v=PWylb_5IOw0&t=38m56s).

5 Estimation and Identification

5.1 Specifications

The specifications in this paper are panel data models with worker, period and business -unit fixed effects. The specifications take the following form:

$$y_{i,t} = \alpha_i + \delta_t + \gamma_j + \Omega c_{i,t} + \beta x_{i,j,t} + \epsilon \tag{4}$$

with:

y_{it}: Performance outcome by worker i at time t (see Section 4.2).
α_i: Worker fixed effect.
δ_t: Period-fixed effect (quarterly).
γ_j: Business-unit (contest) fixed effects.
$c_{i,t}$: Vector of additional employee and business-unit level controls.
$x_{i,j,t}$: Measure of competition faced by worker i in unit j at time t (see Section 4.1).
ϵ: Error term.

Standard errors are clustered at the business-unit level. The coefficient of main interest is β, which is predicted to be positive. Next I describe conditions under which the β coefficient from the specification above can be interpreted as causal and an instrumenting strategy.

5.2 Identification

As discussed in Section 3, assignments to units were strongly influenced by the date-of-hire. However, even without this source of variation, it may have been hard for employees or managers to conspire ex-ante to select levels of competition endogenously. This would require the ability to predict performance trajectories, which the firm (as well as the academic literatures) has found difficult.

Nonetheless, to account for unforeseen forms of selection, the results in this paper include instrumental variable specifications attempting to isolate variation from the date-of-hire variation. The resulting estimates are not qualitatively different than the OLS results, and the estimates mostly overlap.

A brief discussion of the empirical properties of the $x_{i,j,t}$ variable follows. Readers interested only in identification from the date-of-hire variation can skip to section 5.3 for the discussion of the that specification.

Brief discussion of $x_{i,j,t}$: Show that this variable is a random walk. Show that this variable is smoothly distributed. Show that the diff of this variable is smooth.

5.3 Instrumental Variable Approach

In this section I discuss the explicit use of time-of-hire based variation in an instrumental variables approach. This approach utilizes variation in the competitive pool created by the assignment mechanism which assigns workers to projects based on date of hire. To measure this variation using ex-ante characteristics, I use an instrument based on interview scores.

Creating an instrumental variable from the date-of-hire requires some additional work since time-of-hire affects employees differently. For "Worker B+," arriving on January 5 rather than January 6 may have resulted in facing greater competition. For a different employee ("Worker B-") – arriving early may have resulted in lower competition. For some employees, the arrival date would have made little difference.

I first show that the timing of arrival does influence project assignments. In Table 5 contains regressions of match realization. Each row in this regression is a possible match between an employee and a business unit. The dependent variable is a binary variable representing which matches were realized (assigned). The explanatory variables include interactions between worker characteristics (such as their job types and quarters of hire) and particular assignments.

The data show that these interactions do have explanatory power over assignments. Certain business units are hiring more in certain months. However, even on top of that – variables relating to micro-timing have explanatory power over assignment outcomes. For example: If assignments were based on hire-dates, then arriving on an odd or even day of the month would affect which business unit a new employee joined. As Table 5 shows, interactions between even/oddness of hire date and business-unit fixed affects have highly statistically significant explanatory power, measured by an F-test of these interactions.

I next define an instrument that effects $x_{i,j,t}$ (the measure of competition each worker faces). I create an instrument based on the level of competition implied by the average pre-hire interview scores for each employee and his/her contest opponents. To define this

instrument, I first create a measure of competitiveness based on average interview scores. This is analgous to my measure of competition $x_{i,j,t}$, defined in Equation 1 – except it uses interview scores rather than performance scores.

$$x^{pre}_{i,j,t} = -1/N \sum_{k \in j, \neq i} |InterviewScore_i - InterviewScore_{k,t}| \tag{5}$$

To examine the causal effects of competitiveness on performance, we must calculate the counterfactual situation where individual i went to team k as opposed to her true assignment, team j. The measure in Equation 5 above permits this. These counterfactual estimates allow us to utilize econometric strategies that exploit differences in competition across groups, as a result of quasi-randomization of assignment.

From here, I create an instrument that measures the difference in x^{pre} caused by local differences in date-of-hire. Formally, the instrument is defined as:

$$z_{it} = \underbrace{1/N_K \sum_{k \in K} x^{pre}_{i,k,t}}_{\substack{\text{Average competition } i \\ \text{faces by joining one of the} \\ \text{units that accepted workers} \\ \text{on his hire date.}}} - \underbrace{1/N_J \sum_{j \in J} x^{pre}_{i,j,t}}_{\substack{\text{Average competition } i \\ \text{would have faced by joining} \\ \text{one of the other units that} \\ \text{received workers in the} \\ \text{surrounding two weeks.}}} \tag{6}$$

... where:

-K refers to the set of units that received new workers of i's type on the same day as i
-J refers to the set of units NOT in K that received new workers of i's type in the surrounding two weeks.
-N_K and N_J refers to the number of units in of sets K and J, respectively.
-$x^{pre}_{i,j,k}$ calculated as in Equation 5.

This instrument is the difference in the realized $x^{pre}_{i,j,t}$ compared to the $x_{i,j,t}$ had the worker been counterfactually assigned to the units of similar workers who arrived in the two weeks surrounding his actual hire date.[36] This formulation exploits the fact that two workers with the same performance score may face different competition depending on peers.

[36]Two weeks is an arbitrary restriction, and the results were generally robust to different windows.

The first stage of the IV is very strong, with an F-statistic of 30.71. This shows that there are meaningful differences in competition between contests generated by the assignment mechanism.

6 Results and Discussion

Table 6 presents normalized results on hours and output. We see a relatively large and statistically significant effect of the competition on overall effort and output. In Table 7, I show the effects of competition on the efficiency and quality of output. We see an increase in efficiency without effects on the quality of code. In Table 9, I shows effects on innovation, in which there are no statistically significant results in either direction.

Table 8: Effects on cooperation. This shows some truth to the concerns about sabotage. Table 11: Effects on quitting. These results are largely consistent with the existing literature. However, a large body experimental and field research in psychology suggests that incentives may actually undermine performance in tasks that require innovation or creativity. McGraw (1978), McCullers (1978), Kohn (1993), and Amabile (1996) suggest that incentives encourage repetition of past success strategies with greater effort, rather than exploration of novel untested ideas.

In my empirical setting, the core job responsibilities require some amount of creativity. Developing efficient software code, marketing plans and sales pitches are not "rote" activities akin to fruit-picking or athletics and instead require innovation and risk-taking. The results on my output measures in 6 contain many output measures that require creativity. Contrary to the aforementioned psychology, workers respond to contest incentives.

The variables in Table 9 (patents and ideas) attempt to isolate a few variables that particularly measure pure "innovation," such as patenting and idea generation. Even in this setting – the outcomes most related to the psychologists' theories – the effect of competition incentives is ambiguous.

Tournament incentives are also said to have negative side effects, particularly in the form of lobbying and sabotage (or decreased productive cooperation). I find some evidence of this in Table 8, which shows a lower level of cooperation and organizational citizenship between promotion competitors.

However, commentators and theorists discussing these negative side effects suggest that they may impact the efficiency or quality or output. In particular, effort spent on lobbying may detract time and effort from productive output. Similarly, sabotage and lobbying may harm the quality of output. My results on efficiency and quality in Table 7 suggest

the overall effect on efficiency is strongly positive, and the effects on product quality can't be detected.

In addition, I see lower job satisfaction across a variety of dimensions in Table 3. The survey holds the key to interpreting a number of other findings. Many economists believe competitive workplaces are naturally more unpleasant. The literature predicts greater effort, greater sabotage and decreased probability of promotion in competitive settings. In addition, if workers facing competition are more likely to quit, this suggests a distaste for the experiencing competition.

However, an alternative interpretation is possible. Many workers, particularly in a high -achiev-ing environment, may be drawn to the presence of other high-achievers. Despite the competition, workers may find this work satisfying or challenging. They may enjoy working alongside close competition. They may benefit from learning or peer-effects from other high performers. If they indeed benefit from these peer effects, they may find themselves with greater outside options (which would explain the results about quitting).

The subjective satisfaction data shows that do indeed feel a distaste for the workplace competition. Across a variety of measures, workers are less satisfied with their jobs, projects, teammates and managers if they are facing greater promotion competition.

This finding may be surprising to some readers. The level of credentials of the firm's employees are known prior to joining the competition. In some sense, employees may be able to estimate some level of competition. Why would they be surprised or unhappy to face lots of promotion competition and strong opponents for promotions? Overoptimism about one's chances may be one reason. An additional reason comes from anecdotal evidence that many employees are joined the firm in a position beneath their ideal level, and hoped to work their way up through promotion into something more satisfactory. Such workers may aspire to enjoy peer effects from future colleagues, but may be frustrated by competition against peers at their current level.

Table 10 shows effects on specialization, or concentration of effort into a small set of tasks. We see that as competition increases, worker effort becomes concentrated into a smaller set of tasks. In addition to specializing, workers differentiate by choosing orthogonal specializations.

I suggest two interpretations of this behavior. These interpretations are not mutually exclusive; it is likely that both interpretations explain the specialization/differentiation phenomena. One interpretation, inspired by Ellison and Ellison (2009), is that this form of differentiation constitutes "obfuscating." The goal of this behavior is to make comparisons between employees more difficult.

In a tournament setup, this behavior would be rational for players. In contest theory, changes that make comparisons harder could be interpreted as an increase in the "noisiness" of evaluation in the contest. Greater uncertainty in evaluation decreases equilibrium incentives and effort from players and increases player welfare welfare.

In most formal models of tournaments, the "noise" is either an exogenous parameter (Lazear and Rosen, 1981) or a parameter set by a contest designer (Morgan et al., 2012). The results of this paper suggest that players in these models – rather than the principle or the environment – may be able to choose their own noise parameter as a function of their specialization decisions in a multi-tasking setup.

A second possible interpretation is that the orthogonal specialization is meant to help the players win the contest through pursuit of comparative advantage. One could imagine that for each dimension of productivity, a) the principle has a "weight" ("β_j") of value that it places on that type of output, and b) each contestant has a "talent" ("θ_{ij}") governing his/or her ability to produce that form of output compared to others. To win a close competition, players in this case may "specialize" by putting more effort into comparative advantage in accordance with a relationship between β_j and θ_{ij}.

In this model of comparative advantage, specializing helps the contestant win. This specialization and differentiation may also benefit the firm by allocating effort into comparative advantages which are positive for the firm's efficiency and productivity. This may be especially good for the firm if there are complementarities between the set of tasks.

The "specialization," "obfuscation" and "comparative advantage" effects of contests are relatively unexplored aspect of the game-theoretic literature on tournaments. These result suggest that these topics are important components of contests in a real-world setting. Because of the possibility that that differentiation creates gains for both the firm and its workers, it may be one unexpected reason this form of incentives are commonly used.

Tables and Figures

Table 1: Most wage increases come from promotions

Panel A: Google Consumer Survey, CPS weighted	
Source of Most Recent Real Wage Increase (n=15,540)	
Leaving your employer for a new job	24.1%
Stayed at firm; market-wide increase	26.1%
Asking your employer to match another offer	8.2%
Performing well compared to peers at job	41.6%
"At your job or workplace, are promotion slots limited?" (n=16,377)	
Limited number of promotion slots, even if all workers perform well.	77.4%
Unlimited number of promotion slots. All can be promoted who qualify.	22.5%

Panel B: Firms using tournaments for promotions
Adobe, AIG, Amazon, American Express, Cisco Systems, Conoco, Dow Chemical, Enron, Expedia, Facebook, Ford, General Electric, GlaxoSmithKline, Goldman Sachs, Goodyear Tire, Google, Hewlett-Packard, IBM, Intel, LendingTree, Lucent, Microsoft, Motorola, Sun Microsystems Valve and Yahoo.

Notes: **Google Consumer Survey** questions were asked by through a survey created by the author of 10,000 respondents each. The survey was conducted in August 2014. Additional details of the author's survey are discussed in the online appendix. Google published a description and comparative analysis of their methodology in McDonald et al. (2012). To this author's knowledge, the only comparative analysis of GCS' accuracy came from election statistician Nate Silver. Silver's 2012 post-election analysis of polls ranked Google Consumer Surveys the second most accurate poll in the sample of 21 used in his forecasts. Silver (2012) concluded, "Perhaps it wont be long before Google, not Gallup, is the most trusted name in polling."

Table 2: Decisions to Promote

Panel A: Non-score predictors of promotion

	(1) Promoted	(2) Promoted	(3) Promoted
Basic Controls	Yes	Yes	Yes
Worker FEs	No	Yes	Yes
Tenure Controls	No	No	Yes
R^2	0.0704	0.214	0.344
Observation	532272	532272	532272

Panel B: Score-based predictors of promotion

	(1) Promoted	(2) Promoted	(3) Promoted	(4) Promoted
Score	0.0532***		0.0455***	0.0508***
	(0.00122)		(0.00136)	(0.00152)
Lag 1 Score		0.0418***	0.0154***	0.00619***
		(0.00121)	(0.00129)	(0.00124)
Basic Controls	No	No	No	Yes
Worker FEs	No	No	No	Yes
Tenure Controls	No	No	No	Yes
R^2	0.0193	0.0123	0.0201	0.367
Observations	532272	532272	532272	532272

Panel C: Rank, Score and optimal threshold predictors of promotion

	(1) Promoted	(2) Promoted	(3) Promoted	(4) Promoted
Score	0.0557***		0.0418***	0.0103**
	(0.00145)		(0.00466)	(0.00482)
Percentile Score		0.181***	0.0482***	0.111***
		(0.00454)	(0.0146)	(0.0146)
In 87th Percentile				0.0683***
				(0.00549)
Controls	All	All	All	All
R^2	0.366	0.366	0.366	0.367
Observations	532272	532272	532272	532272

Panel D: Local comparisons predict promotion

	(1) Promoted	(2) Promoted	(3) Promoted	(4) Promoted
Score (Global)	0.0557***		0.0248***	0.0633***
	(0.00145)		(0.00231)	(0.00156)
Score (Local)		0.0615***	0.0386***	
		(0.00154)	(0.00257)	
Diff of Local & Global Score				0.0384***
				(0.00256)
Controls	All	All	All	
R^2	0.366	0.367	0.367	0.367
Observations	532272	532272	532272	532272

Notes: This table presents regressions of promotions (represented as one or zero) as a function of the level of absolute and relative productivity, measured through subjective performance scores. The unit of analysis is the employee-quarter. All regressions include controls for team size, ..., as well as fixed effects for quarter, job type, job level and a twelve-degree polynomial in the worker's tenure at the firm and tenure in his or her current job.

Table 3: Promotion Competition and Satisfaction

Job Satisfaction Survey (n∼5000)	
Satisfaction with: Company overall.	-0.155**
Satisfaction with: Your manager.	-0.098*
Satisfaction with: Your projects.	-0.174**
Satisfaction with: Your workload.	-0.211***
Satisfaction with: Ability to manage/balance work and personal life.	-0.117**
"There is a climate of trust within company."	-0.226***
"The people in my work group cooperate to get the job done."	-0.141***
"I understand how my performance is evaluated."	-0.341*
"I think my performance on the job is fairly evaluated."	-0.153**

Notes: This table presents results of instrumental variable regressions on a cross sectional measure of job satisfaction and survey responses. For interpretation, I have used normalized independent and dependent variables. The unit of observation in these regressions is a worker approximately halfway through the sample. The sample size is rounded to the nearest hundred for confidentiality reasons. Standard errors are clustered at the business unit containing the contest. A full discussion of these specifications is included in Section 5.

The instrumented variable "Proximity of Competition" is based on subjective performance scores of each employee's contest peers in the previous quarter. This measure is described in depth in Section 4.1. The instrument utilizes date-of-hire variation and is described in Section 5.3.

All regressions control for worker, quarter, business unit, and basic job-type and job-level fixed effects. In addition, all regressions control for the focal worker's tenure at the firm, tenure at his current position performance score in the previous quarter.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Observable Characteristics are Uncorrelated with the Instrument

	(1) Instrument	(2) Instrument	(3) Instrument
Normalized Interview Score	-0.0863	-0.0868	-0.0867
	(0.105)	(0.104)	(0.105)
Log(Starting Salary)		0.000398	0.000290
		(0.000729)	(0.000752)
Job Level (1-9)			-0.00211
			(0.00277)
N	532314	532314	532314

Notes: This table presents regressions of my instrument (described in Section 2) on observable characteristics at the time of hire. Standard errors are clustered at the business unit level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Which Possible Assignments are Realized for New Workers?

Panel A: All Possible Assignments

	(1)	(2)	(3)	(4)	(5)	(6)
Quarter & Team FEs	Yes	Yes	Yes	Yes	Yes	Yes
Team x Quarter FEs		Yes	Yes	Yes	Yes	Yes
Team x Job Type FEs			Yes	Yes	Yes	Yes
Job Type x Quarter FEs				Yes	Yes	Yes
Team x Even Day FEs					Yes	Yes
Team x Quarter x Even Day FEs						Yes
F-test of new FEs=0	$p<0.001$	$p<0.001$	$p<0.001$	$p\sim 1$	$p<0.001$	$p<0.001$
R^2	0.0333	0.0643	0.103	0.103	0.103	0.108
Observations	8834760	8834760	8834760	8834760	8834760	8834760
Clusters	41	41	41	41	41	41

Panel B: Removing Team x Quarters with No Hires

	(1)	(2)	(3)	(4)	(5)	(6)
Quarter & Team FEs	Yes	Yes	Yes	Yes	Yes	Yes
Team x Quarter FEs		Yes	Yes	Yes	Yes	Yes
Team x Job Type FEs			Yes	Yes	Yes	Yes
Job Type x Quarter FEs				Yes	Yes	Yes
Team x Even Day FEs					Yes	Yes
Team x Quarter x Even Day FEs						Yes
F-test of new FEs=0	$p<0.001$	$p<0.001$	$p<0.001$	$p<0.001$	$p<0.001$	$p<0.001$
R^2	0.0419	0.0591	0.119	0.119	0.119	0.123
Observations	2705099	2705099	2705099	2705099	2705099	2705099
Clusters	41	41	41	41	41	41

Notes: This table presents regressions using a dataset of all possible initial business unit assignments for each employee. Each employee is assigned at most one initial assignment, which is the dependent variable (1=assigned, 0=not). The independent variables are interactions between employee's characteristics characteristics upon entry (the date of hire and his type of position) and each team.

The table shows that employees of certain job types are more likely to be assigned to certain teams; this is because of some amount of functional separation within the firm's business units. It also shows that certain professional roles were more popular with with certain teams at certain times.

Importantly, the table also shows in the latter columns that arriving on an even or odd day predicts teams assignment. This is because, as described in Section 3, a common way of assigning new employees to units at the firm was based on date-of-hire. In this method, the firm assigned employees in batches on certain days. Workers were given wide latitude to select the timing of arrival, and the timing was generally affected by landlord, spousal, vacation and relocation-related preferences.

Standard errors are clustered at the quarter-of-hire level.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Effects on Effort and Productivity

	(1) Hours	(2) Hours	(3) Output	(4) Output
Competition	0.742***	0.758***	0.171*	0.187**
	(0.125)	(0.133)	(0.0930)	(0.0892)
Controls	No	Yes	No	Yes
Worker FEs	Yes	Yes	Yes	Yes
Quarter FEs	Yes	Yes	Yes	Yes
Unit FEs	Yes	Yes	Yes	Yes
Tenure Controls	Yes	Yes	Yes	Yes
R^2	0.619	0.619	0.385	0.385
Observations	532272	532272	532272	532272
Clusters	418	418	418	418

Notes: This table presents results of instrumental variable regressions of effort- and productivity- related out comes in a panel data setting. For interpretation, I have used normalized independent and dependent variables. The unit of observation in these regressions is a worker $\times$ quarter. Standard errors are clustered at the business unit containing the contest. A full discussion of these specifications is included in Section 5.

"Output" refers to the sum of all activity measures. "Hours" refers to the number of hours with productive activity during the period; I use this as a measure of effort. A related variable, "Efficiency" (output per hour) is studied in Table 7 below. The dependent variables "Output" and "Hours" are described in greater detail in Section 4.2.1.

The instrumented variable "Proximity of Competition" is based on subjective performance scores of each employee's contest peers in the previous quarter. This measure is described in depth in Section 4.1. The instrument utilizes date-of-hire variation and is described in Section 5.3. The F-statistic in the first stage of is 31.

All regressions control for worker, quarter, business unit, and basic job-type and job-level fixed effects. In addition, all regressions control for the focal worker's tenure at the firm, tenure at his current position performance score in the previous quarter.

The additional "Controls" above refer to controls for higher and lower granularity controls for job type and level, as well as controls for business unit size and composition. These are included and excluded as robustness checks of different plausible specifications.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Effects on Efficiency

	(1) Efficiency	(2) Efficiency	(3) Rollbacks	(4) Rollbacks	(5) % Rollbacks	(6) % Rollbacks
Proximity to Competition	0.511***	0.525***	0.00204	0.00111	-0.0241	-0.0236
	(0.0710)	(0.0762)	(0.00641)	(0.00604)	(0.0174)	(0.0176)
Controls	No	Yes	No	Yes	No	Yes
Worker FEs	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FEs	Yes	Yes	Yes	Yes	Yes	Yes
Unit FEs	Yes	Yes	Yes	Yes	Yes	Yes
Tenure Controls	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.376	0.376	0.123	0.123	0.0863	0.0863
Observations	532272	532272	532272	532272	532272	532272
Clusters	418	418	418	418	418	418

Notes: This table presents results of instrumental variable regressions of effort- and productivity- related out comes in a panel data setting. For interpretation, I have used normalized independent and dependent variables. The unit of observation in these regressions is a worker $\times$ quarter. Standard errors are clustered at the business unit containing the contest. A full discussion of these specifications is included in Section 5.

"Efficiency" refers to the ratio of output and hours. Similar results for these variables are in Table 6. "Rollbacks" is a measure of engineering quality; a "rollback" is when code is later withdrawn. I include both total lines of code rolled back, as well as percentage of lines of code rolled back. Additional discussion of my measures of efficiency and quality are in 4.2.2.

The instrumented variable "Proximity of Competition" is based on subjective performance scores of each employee's contest peers in the previous quarter. This measure is described in depth in Section 4.1. The instrument utilizes date-of-hire variation and is described in Section 5.3. The F-statistic in the first stage of is 31.

All regressions control for worker, quarter, business unit, and basic job-type and job-level fixed effects. In addition, all regressions control for the focal worker's tenure at the firm, tenure at his current position performance score in the previous quarter.

The additional "Controls" above refer to controls for higher and lower granularity controls for job type and level, as well as controls for business unit size and composition. These are included and excluded as robustness checks of different plausible specifications.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8: Effects on Cooperation

	(1) Peer Bonuses	(2) Peer Bonuses	(3) Doc Edits	(4) Doc Edits
Competition	-0.745***	-0.741***	-0.389***	-0.392***
	(0.149)	(0.147)	(0.149)	(0.151)
Controls	No	Yes	No	Yes
Worker FEs	Yes	Yes	Yes	Yes
Quarter FEs	Yes	Yes	Yes	Yes
Unit FEs	Yes	Yes	Yes	Yes
Tenure Controls	Yes	Yes	Yes	Yes
R^2	0.227	0.227	0.471	0.471
Observations	532272	532272	532272	532272
Clusters	418	418	418	418

Notes: This table presents results of instrumental variable regressions of cooperation- related out comes in a panel data setting. For interpretation, I have used normalized independent and dependent variables. The unit of observation in these regressions is a worker $\times$ quarter. Standard errors are clustered at the business unit containing the contest. A full discussion of these specifications is included in Section 5.

"Peer bonuses" refer to roughly $200 bonuses that employees can give each other publicly to highlight exceptional work by peers. "Edits" refers to edits to internal documentation pages that others within a team; I use this as a measure of hoarding knowledge or sharing it. The dependent variables "Peer Bonuses" and "Edits" are described in greater detail in Section 4.2.3.

The instrumented variable "Proximity of Competition" is based on subjective performance scores of each employee's contest peers in the previous quarter. This measure is described in depth in Section 4.1. The instrument utilizes date-of-hire variation and is described in Section 5.3. The F-statistic in the first stage of is 31.

All regressions control for worker, quarter, business unit, and basic job-type and job-level fixed effects. In addition, all regressions control for the focal worker's tenure at the firm, tenure at his current position performance score in the previous quarter.

The additional "Controls" above refer to controls for higher and lower granularity controls for job type and level, as well as controls for business unit size and composition. These are included and excluded as robustness checks of different plausible specifications.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 9: Effects on Innovation

	(1) Patents	(2) Patents	(3) Ideas	(4) Ideas	(5) Rating	(6) Rating
Proximity to Competition	-0.188	-0.190	0.0351	0.0265	0.0795	0.0481
	(0.152)	(0.154)	(0.0219)	(0.0220)	(0.0659)	(0.0568)
Controls	No	Yes	No	Yes	No	Yes
Worker FEs	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FEs	Yes	Yes	Yes	Yes	Yes	Yes
Unit FEs	Yes	Yes	Yes	Yes	Yes	Yes
Tenure Controls	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.162	0.162	0.0817	0.0818	0.260	0.262
Observations	532272	532272	532272	532272	527240	527240
Clusters	418	418	418	418	418	418

Notes: This table presents results of instrumental variable regressions of effort- and productivity- related out comes in a panel data setting. For interpretation, I have used normalized independent and dependent variables. The unit of observation in these regressions is a worker $\times$ quarter. Standard errors are clustered at the business unit containing the contest. A full discussion of these specifications is included in Section 5.

"Patents" refer to patent applications rewarded to the focal worker. "Ideas" refers to the number of ideas an employee submitted to an internal system for proposing new ideas. "Rating" refers to the average score assigned to the ideas submitted by the focal worker by peer reviewers. Additional discussion of theses measures can be found in Section 4.2.4.

The instrumented variable "Proximity of Competition" is based on subjective performance scores of each employee's contest peers in the previous quarter. This measure is described in depth in Section 4.1. The instrument utilizes date-of-hire variation and is described in Section 5.3. The F-statistic in the first stage of is 31.

All regressions control for worker, quarter, business unit, and basic job-type and job-level fixed effects. In addition, all regressions control for the focal worker's tenure at the firm, tenure at his current position performance score in the previous quarter.

The additional "Controls" above refer to controls for higher and lower granularity controls for job type and level, as well as controls for business unit size and composition. These are included and excluded as robustness checks of different plausible specifications.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 10: Effects on Specialization

	(1) Specialization	(2) Specialization
Competition	1.178***	1.105***
	(0.310)	(0.268)
Controls	No	Yes
Worker FEs	Yes	Yes
Quarter FEs	Yes	Yes
Unit FEs	Yes	Yes
Tenure Controls	Yes	Yes
R^2	0.587	0.597
Observations	532272	532272
Clusters	418	418

Notes: This table presents results of instrumental variable regressions of effort- and productivity- related out comes in a panel data setting. For interpretation, I have used normalized independent and dependent variables. The unit of observation in these regressions is a worker $\times$ quarter. Standard errors are clustered at the business unit containing the contest. A full discussion of these specifications is included in Section 5.

Specialization is measured through Herfindahl-Hirschman Index ("HHI") (Hirschman, 1945) measure of the share of output from various activities. Discussion of the measurement of this variable is in Section 4.2.5.

The instrumented variable "Proximity of Competition" is based on subjective performance scores of each employee's contest peers in the previous quarter. This measure is described in depth in Section 4.1. The instrument utilizes date-of-hire variation and is described in Section 5.3. The F-statistic in the first stage of is 31.

All regressions control for worker, quarter, business unit, and basic job-type and job-level fixed effects. In addition, all regressions control for the focal worker's tenure at the firm, tenure at his current position performance score in the previous quarter.

The additional "Controls" above refer to controls for higher and lower granularity controls for job type and level, as well as controls for business unit size and composition. These are included and excluded as robustness checks of different plausible specifications.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 11: Effects on Quitting

	(1)	(2)	(3)
Proximity to Competition	0.105***	0.113***	0.113***
	(0.0163)	(0.0152)	(0.0146)
Performance		-0.182***	-0.169***
		(0.00337)	(0.00644)
Proximity x Performance			0.0376***
			(0.00901)
Controls	Yes	Yes	Yes
Period Controls	Yes	Yes	Yes
Unit FEs	Yes	Yes	Yes
Tenure Controls	Yes	Yes	Yes
Observations	532272	532272	532272
Clusters	418	418	418

Notes: This table presents the results of Cox regressions in panel setting. Each observation is a worker $\times$ quarter. All regressions controls for job type, job level, organizational depth, percentage full time, the department/business unit of the worker and quarterly period fixed effects. Standard errors are clustered at the business unit.

Chapter 2:
The Value of Hiring Through Referrals

This chapter is co-authored with Stephen Burks, Mitchell Hoffman and Michael Housman

1 Introduction

Firms often use referrals from existing employees to hire new workers: about 50% of US jobs are found through informal networks and about 70% of firms have programs encouraging referral-based hiring.[37] A large and growing theoretical literature seeks to understand hiring through referrals, as well as draw out implications of referrals for many central issues in labor economics, including wage inequality, duration dependence in unemployment, racial gaps in unemployment, and the quality of worker-firm matching over the business cycle.[38] While there is a rich and growing empirical literature on referrals (see Ioannides and Loury [2004] and Topa [2011] for excellent reviews), particularly for questions on how networks affect worker outcomes (such as job-finding), relatively little is known empirically about what firms may gain from referral-based hiring. There are two main data challenges. First, referrals are difficult to directly observe. Second, understanding what firms gain from referrals requires data on productivity, which are also rare.

We overcome these challenges by assembling personnel data from nine large firms in three industries: call-centers, trucking, and high-tech. Spanning hundreds of thousands of workers and millions of applicants, our data combine direct measurement of employee referrals; high-frequency measurement of worker productivity on multiple dimensions; and surveys conducted by the firms and by the authors on different aspects of worker ability, including aspects that are not observed by the firm at time of hire. We organize our findings around answers to three main questions:

[37] Granovetter (1974) showed that roughly 50% of workers are referred to their jobs by social contacts, a finding which has been confirmed in more recent data Topa (2011). A leading online job site estimates according to their internal data that 69% of firms have a formal employee referral program CareerBuilder (2012). Employee referral programs encourage referrals from existing employees, often by offering monetary bonuses for when referred candidates get hired.

[38] (Montgomery, 1991) and Calvo-Armengol and Jackson (2007), among many others, analyze how referrals affect wage inequality. (Calvo-Armengol and Jackson, 2004) analyze referrals and duration dependence in unemployment. (Holzer, 1987b) and (Zenou, 2012) study referrals and racial gaps in unemployment. (Galenianos, 2012) analyzes how referrals affect the quality of worker-firm matching over the business cycle. For a detailed treatment of the growing theoretical literature on social networks in economics in general, see Jackson (2008).

1. **How do firms treat referred vs. non-referred applicants?** Compared to non-referred applicants, referred applicants are substantially more likely to be hired, and, conditional on receiving an offer, they are more likely to accept it. This occurs even though on most characteristics, both observable and unobservable to the firm at time of hire, referred and non-referred applicants are similar, as are referred and non-referred workers.

2. **Are referred workers more productive than non-referred workers?** On many productivity measures, referred and non-referred workers have economically similar performance. There are two main exceptions: (a) In trucking, referred workers have fewer accidents than non-referred workers and (b) In high-tech, referred workers are more likely to invent patents than non-referred workers.

3. **How costly to the firm are referred vs. non-referred workers in terms of turnover, wages, or other aspects?** In all three industries, referred workers are significantly less likely to quit than non-referred workers. Referred workers have higher wages than non-referred workers only in high-tech, where the difference is relatively modest.

Having documented these results, we move toward quantifying differences in profits between referred and non-referred workers. We focus on call-centers and trucking, where the production process is relatively simple. Referred workers produce substantially higher profits than non-referred workers. In both call-centers and trucking, profit differences are driven by lower turnover and lower recruiting costs. Turnover is costly since quitting workers need to be replaced, and because new workers require training and time to reach peak productivity. Referrals also reduce recruiting costs, as it requires substantially fewer applicants to be screened to produce a hire among referrals compared to non-referrals. Higher productivity is not a significant driver of profit differences in both call-centers and trucking.

We close by considering two sources of heterogeneity in the value that firms gain from referrals. The first source we consider is the *referrer*, that is, the employee making the referral. In high-tech and trucking, referrers tend to have higher productivity than workers who don't make referrals. In trucking, where we know who referred whom, we find that referrers tend to refer people like themselves in productivity. Consequently, there are large differences in profits between referrals from high-productivity referrers compared to low-productivity referrers. The second source is local labor market conditions. In trucking, we find that differences between referrals and non-referrals in hiring rates, offer acceptance rates, and trucking accidents are larger when the local economy is strong instead of weak.

Theory has identified a number of reasons why firms may benefit from hiring through referrals. In learning theories (Simon and Warner 1992; Dustmann, Glitz, and Schoenberg 2012; Brown, Setren, and Topa 2013; Galenianos 2013), referrals reduce uncertainty about match quality for potential workers. With less uncertainty about match, referred workers will have higher reservation wages than non-referred workers, as well as higher productivity and wages. In homophily theories Montgomery (1991); Casella and Hanaki (2008); Galenianos (2012), firms seek referrals from their highest ability workers, which they do given a tendency of people to be socially connected with those of similar ability. Homophily is the pervasive tendency of people to associate with those like themselves McPherson, Smith-Lovin and Cook (2001). Referred workers will have superior unobservables and productivity, and will produce positive profits for firms. In a third class of theories, which we call peer benefit theories Kugler (2003); Castilla (2005); Heath (2013), referrals are valuable because of benefits that referrers and referrals derive from working in the same organization. For example, referrers may mentor referrals or monitor their behavior, or it may simply be more enjoyable for referrers and referrals to work together.[39]

In empirical work, a relatively small, but growing literature explores referral-based hiring from the perspective of the firm.[40] (Beaman and Magruder, 2012) and (Pallais and Sands, 2013) conduct field experiments using workers in India and in an online marketplace, respectively, to study whether and why referred workers are more productive. Turning to non-experimental studies, recent contributions include (,)DustmannEtAl12, (Heath, 2013), and (Hensvik and Skans, 2013).[41] Most similar to our paper in using direct measures of referral status and developed country workers is (,)BrownEtAl12, who use personnel data from one US firm to provide a rich analysis of wages, turnover, and hiring. Relative to (,)BrownEtAl12, we use a much larger sample over nine firms, direct measures

[39]Besides what we consider the three leading classes of theories, it could also be the case that referrals reflect favoritism (e.g., Beaman and Magruder, 2012). Paralleling Becker's taste-based model of racial discrimination, it could be that incumbent employees persuade firms to hire social contacts, even if these social contacts may not be the best-suited for the job. If referrals reflect favoritism, then referred applicants may receive a "lower bar" in getting hired, and referred workers may end up having lower productivity.

[40]Pioneering work on referrals was conducted by (Rees, 1966) in economics and by Granovetter (1973, 1974) in sociology. There is also substantial more recent work by sociologists, e.g., (,)FernandezEtAl00 and (Castilla, 2005). In economics, there is now a significant literature on the effects of worker social networks in individual job search; see (Ioannides and Loury, 2004) and (Topa, 2011) for surveys, as well as (Kramarz and Skans, 2014) and (Schmutte, 2015) for noteworthy recent examples. This literature is connected to, but, we believe, conceptually separate, from work on why firms use referral-based hiring.

[41](,)DustmannEtAl12 develop a dynamic search model of learning through referrals, which they test using co-ethnic hiring patterns in German matched employer-employee data. (Heath, 2013) develops a model where referrals reduce limited liability problems, which she tests using data on Bangladeshi garment workers. (Hensvik and Skans, 2013) investigate homophily models by studying past co-worker linkages between entering and incumbent workers using matched employer-employee data from Sweden.

of productivity, and detailed data on applicant and worker skill characteristics.

The main contribution of our paper is to assess the benefits that firms receive from referrals vs. non-referrals, and to quantify those benefits in terms of profits. Past work has compared referred and non-referred workers in terms of wages. However, wage differences alone are not enough to compute whether there are profit differences, since productivity may differ importantly from wages (e.g., Lazear, 1979; Medoff and Abraham, 1980, 1981; Flabbi and Ichino, 2001; Shaw and Lazear, 2008) and there are often multiple dimensions to productivity. Several of our findings would be overlooked with only wage data. For example, in high-tech, referrals have substantially higher levels of innovation than non-referrals, whereas they have only modestly higher wages.

Although there are differences across industries that we highlight along the way, the facts we document are surprisingly consistent across firms and industries. This suggests that our findings may be relevant for many firms in the economy, although we certainly acknowledge that questions of generalizability may remain, even with nine firms.

2 Data

For each of the three industries, we discuss (1) the nature of the data; (2) how productivity is measured, how workers are paid, and what survey data were collected; and (3) how referrals are measured and how the bonuses for making referrals (i.e., the employee referral programs) are structured. Some details about the firms cannot be given due to confidentiality restrictions. For brevity, we provide variable definitions in Appendix B (all appendix material is in the Web Appendix on the journal website).

Call-centers. We obtained our call-center data from a human resources (HR) analytics firm called Evolv, which provides call-center firms with job testing software.[42] The data provided to us is comprised mainly of data from seven large call-center firms and we restrict our sample to these seven firms.[43] The data run from July 2009 to July 2013. Across the seven firms, data coverage begins at different dates, reflecting when the firms

[42]Evolv was purchased in late 2014 by Cornerstone OnDemand. One of the paper's authors, Michael Housman, is the current Chief Analytics Officer for Cornerstone OnDemand.

[43]The seven large firms comprise 93% of the applicants and 97% of the workers in the data provided to us, and we restrict our sample to these firms. The average number of hires among the seven large firms is 10,514 workers per firm. In addition to the seven large firms, the data provided to us include six additional firms, hiring an average of only 400 workers per firm across all years of the data. All our results are very similar regardless of whether we restrict to the seven large firms or whether we include the additional firms, as seen in Appendix Table C.25.

begin using Evolv's job test. Five firms adopt Evolv by quarter 1 of 2011, whereas the other two adopt in 2012.[44] Restricting to the seven large firms, our sample is comprised of about 350,000 applicants and 74,000 hires. The large majority of the workers (about 85%) are located in different parts of the US, with a small number located abroad (primarily in the Philippines). Each of the seven call-center firms has multiple locations (in the data each firm has, on average, about 15 locations) and provides service to large end-user companies, e.g., large credit card or cellphone companies. Within each location, different workers may work for different end-user companies.

In the call-centers, the production process consists of in-bound and out-bound calls, with workers doing primarily customer service or sales work. Performance is measured using five industry-standard productivity measures (three objective, two subjective), though which productivity measures are available varies by firm.[45] The three objective productivity measures are: schedule adherence, measuring the share of work time a worker spends performing work; average handle time, with a lower average call time indicating higher productivity; and the share of sales calls resulting in a successful sale. The two subjective productivity measures are a manager's assessment from listening in of whether the service was of high quality (quality assurance) and the customer satisfaction score. Workers are primarily paid by the hour. Turnover is high in call-centers and is costly for firms—in our data, roughly half of workers leave within the first 90 days. A great deal of information on applicant skills is available from applicant job tests, including numerous questions on cognitive and non-cognitive ability (though some skill characteristics are not collected by the job tests at some call-center locations).

Referral status is measured via a self-report on the applicant's job test ("Were you referred to this job application by someone that already works for this company?") Referral bonuses vary by firm and by location within the firm, but are typically around $50-$150. The applicant must be hired for a referrer to be paid, and there is typically some tenure requirement (e.g., 30 or 90 days) that the referral must stay to yield a bonus for the referrer. A referral bonus of $100 is about 0.5% of annual earnings for our sample.

Trucking. The data are from a very large US trucking firm, covering all driver applicants and hires over the period 2003-2009. To preserve the firm's anonymity, we do not release

[44]The firms begin using the Evolv test (i.e., the firms enter the sample) in the following months: 2009/07, 2010/09, 2011/01, 2011/02, 2011/03, 2012/02, and 2012/09. Within firms, there is also cross-location variation in when Evolv's job testing is implemented; thus, different locations within a firm enter our sample at different times. Appendix A.12 gives further discussion and provides evidence that the presence of variation in when firms and locations enter our sample does not seem important for the interpretation of our results.

[45]For some workers, no productivity measures are available. In Appendix A.5, we provide evidence that this is unlikely to be a source of bias for our productivity analysis.

the exact total number of applicants, employees, or employee-weeks in the sample. The baseline data include weekly miles, accidents, quits, and a number of background characteristics, and are available for tens of thousands of workers. In addition, we collected very detailed survey data one week into training for a subset of roughly 900 new drivers starting work in late 2005 and 2006. Data collected include cognitive and non-cognitive ability, experimental preferences (collected through incentivized lab experiments), and more detailed information on worker background, all data which were not observed by the firm at time of hire.[46]

Production consists of delivering loads between locations. Drivers are paid almost exclusively by the mile (a piece rate), are non-union, and are away from home for long periods of time. The standard productivity measure in long-haul trucking is miles driven per week. Even though most drivers work about the same number of hours (i.e., the federal legal limit of about 60 hrs/week), there are substantial and persistent productivity differences across workers in miles per week, which are due to several factors, including speed, skill at avoiding traffic, route planning, and coordinating with people to unload the truck. Beyond miles, another important performance metric is accidents. Turnover is high, both in quits and fires, though quits outnumber fires by 3 to 1. Roughly half of workers leave within their first year. Workers who have poor performance, either in miles or accidents, risk getting fired.

Referral status for truckers is recorded both using a survey question in the job application (how the worker found out about the job) and using administrative data from the firm's employee referral program. These two measures of referral status are highly correlated, suggesting that both are reliable (see Appendix B.1). For our analysis, we use the job application survey question measure of referral status, since it is available for the entire sample period, whereas we only have data from the employee referral program for October 2007-December 2009. A limitation is that we are missing information from the job application survey question for 37% of the applicants and 33% of the workers. Observations with missing referral status information are excluded from the sample. Fortunately, referral status does not appear to be missing in a systematic way that would affect our findings.[47] Out of the three industries, only for trucking do we have matched data on

[46]The data were collected during commercial driver's license training at one of the firm's regional training schools. The participation rate was high; 91% of those offered the chance to participate in data collection chose to do so. See Burks et al. (2008) and Appendix B.3 for more on the data collection. Appendix Table C.1 compares drivers in the full dataset to drivers in the subset with very detailed data. Drivers in the subset have a higher share of being referred. They also have lower earnings, reflecting that they are new drivers, as well as somewhat different demographics (reflecting that they are primarily from one region of the US). From the subset of 895 drivers, referral status is observed for 628 drivers and we restrict attention to these drivers.

[47]The job application survey question on referral status is optional (as are a number of other questions on

who referred whom (via the administrative data from the employee referral program), and we only have the match for Oct. 2007-Dec. 2009. For referring an experienced driver, an incumbent worker generally receives $500 when the driver is hired and $500 if the referred driver stays at least 6 months. For referring an inexperienced driver, an incumbent worker generally receives $500 if the worker stays 3 months.[48] For our sample, a referral bonus of $1,000 is about 3% of annual earnings.

High-tech. We use data from a large high-tech firm. The data have about 25,000 workers and 1.4 million applicants. For 2003-2008, we have data on all new regular employee hires, as well as on all applicants that are interviewed for those positions. In addition, for June 2008-May 2011, we have data on all applicants who apply (instead of just those interviewed) for engineering and computer programmer positions.

Most of the high-tech workers are high-skill individuals with advanced education. The largest share are engineers, computer programmers, and technical operations personnel. In addition, some workers are in sales and customer support. Unlike in call-centers and trucking, much of production occurs in teams. Productivity measures include both subjective performance reviews and detailed objective measures of employee behavior, including the number of times one reviewed or debugged other people's code, built new code, or contributed to the firm's internal wiki. We also have data on worker innovation, an important aspect of performance in many high-tech fields. Worker innovation is measured primarily using patent applications since getting hired.[49] We analyze the objective performance and innovation measures at the monthly level, whereas the subjective performance reviews are given by quarter. The vast majority of workers are paid by salary. Unlike in our call-center or trucking data, turnover is low, with workers staying years

the job application). Appendix Table C.2 shows that while there are some slight demographic differences between those with and without referral status information from the job application survey question, there is no significant correlation between trucker productivity (miles or accidents) and whether referral status is missing. Furthermore, the paper's results are qualitatively similar if we measure referral status using information from the administrative employee referral program where there is no missing data (Appendix Table C.23). For the other 8 firms (the 7 call-center firms and the high-tech firm), we only have one measure of referral status, but it has essentially no missingness. See Appendix A.3 for further details.

[48] These referral bonuses were generally standard over 2003-2009, though there were occasional cases when drivers were paid different amounts (e.g., different regions paid slightly different amounts), as described in Appendix B.3.

[49] At the high-tech firm, employees who create an invention file an Invention Disclosure Form. Attorneys from the firm then decide whether to file a patent application. Most of these patent applications are later approved as patents, but the process usually takes several years. For the analysis, our variable of interest is patent applications per employee. This is advantageous in two respects. First, patent applications are observed right away (whereas actual patent award occurs usually multiple years later). Second, it allows us to compare referrals and non-referrals in terms of the ideas that the firm thought were most valuable to patent, instead of merely all the ideas that an inventor chose to disclose.

with the firm. Incumbent workers are surveyed occasionally by the HR department; a 2006 survey provides us with data on personality traits which are not directly observed by the firm at time of hire.[50]

Referral status at the high-tech firm is measured using administrative data from the company's employee referral program (i.e., a current employee provided an applicant's resume to the HR department). Referral bonuses have varied over time, but are usually a few thousand dollars, and are paid to referrers for an applicant getting hired (no tenure requirement). We have data on receipt of referral bonuses, so we know which employees made successful referrals and when; however, we do not know whom an employee referred.

Summary. Table 12 summarizes the data elements available for the three industries. Certain elements such as cognitive ability and personality are available for workers in all three industries. Other elements are only available for particular industries. For example, trucking is the only industry where we know who referred whom, and where we have measures of worker behavior in lab experiments. Table 13 provides sample means for workers. The share of workers referred is 36%, 20%, and 33% in call-centers, trucking, and high-tech, respectively. During our sample period, 51% of truckers are observed ever having an accident and 6% of high-tech workers are observed ever developing a patent.

3 Applicant quality and hiring

Table 14 shows that referred applicants are substantially more likely to be hired. We analyze linear probability models of being hired on referral status and observables. Throughout the paper, standard errors are clustered at the applicant or worker level (depending on whether we are analyzing applicants or workers).[51] In call-centers, referred applicants are 6.3 percentage points more likely to be hired (up from a base of 19% for non-referred applicants), and falling to 6.0 percentage points once demographics are controlled for. In

[50]The survey also provides us with SAT scores. We do not have personality or SAT score data for employees who joined the company after the survey was administered. The survey was voluntary. The participation rate was only about one-third, but the participation rate was very similar for referred and non-referred workers (regressing whether one responds to the survey on referral status and the controls in Panel C of Table 15, the response rates of referrals and non-referrals differ by less than 1 percentage point). Also, some people who answered the survey chose to answer the personality questions, but left the question about SAT scores blank.

[51]We do this because referral status, the main regressor of interest, varies at the individual level. When we analyze the interaction term of referral times the annual state unemployment rate, we will show results clustering by state. In addition, we include dummy variables for missing instances of control variables.

trucking, referred applicants are 10 percentage points more likely to be hired up from a base of 17%. In high-tech, referred applicants are 0.27 percentage points more likely to be hired, which is sizable relative to the base of 0.28%. Because the coefficients remain large after observables are controlled for, this suggests that firms recognize that referrals may be better along unobserved dimensions.

Table 14 also shows that referred applicants are more likely to accept offers, conditional on receiving them. In baseline specifications without demographic controls, referred applicants are 5.1, 7.3, and 2.7 percentage points more likely to accept an offer in call-centers, trucking, and high-tech, respectively. Results are similar after demographic controls are added.

These differences in hiring and offer acceptance are interesting because in terms of characteristics, referred and non-referred applicants look similar. In addition, referred and non-referred workers look similar. We focus our comparison here on schooling, cognitive ability, and non-cognitive ability, analyzing additional characteristics in Appendix C. Schooling is observed by firms at time of hire, whereas cognitive and non-cognitive ability are generally not directly observed by the firm at time of hire.[52] For applicants, data on applicant schooling, cognitive ability, and non-cognitive ability are mostly only available for call-centers, so we focus the analysis there.[53] For trucking, data on worker characteristics are only available for the subset of 900 new drivers described in the data section. For high-tech, data on SAT scores and Big 5 personality traits come from a survey by the HR department.

Results are in Table 15. Starting with comparisons among referred and non-referred applicants, column 1 of Panel A shows that, in call-centers, referred applicants have 0.10 fewer years of schooling, which is statistically significantly different from 0, but economically small in comparison to the standard deviation of schooling, which is 1.3 years for the column 1 sample. Referred applicants score 0.02 standard deviations (σ) lower in intelligence, and 0.02σ lower on the Big 5 Personality Index, our measure of non-cognitive

[52]With the exception of SAT scores in high-tech, our data on cognitive skills, non-cognitive skills, and experimental preferences were not directly observed by firms at time of hire. In the call-centers, data on cognitive and non-cognitive skills, as well as substantial other information about work-relevant skills and job fit, are collected by the job testing company, Evolv; however, only overall job test scores on each applicant are shared with the call-center firms. In trucking, the data were collected by the authors on workers during training. In high-tech, the data on non-cognitive skills were collected in a survey of existing workers by the HR department. Of course, firms may receive partial information about cognitive/non-cognitive skills during recruitment, so it is more correct to think of such variables as not easily observed instead of unobserved Altonji and Pierret (2001).

[53]For high-tech, we have data on applicant schooling for a small sample of applicants, comprised of applicants applying between 2008 to 2010 for entry-level jobs. In that sample, referred workers have 0.12 fewer years of schooling than non-referred workers.

ability, which is equal to the mean of the normalized Big 5 personality characteristics. Looking one by one at the different personality characteristics in Appendix Table C.3, referred applicants are less conscientious, less agreeable, and less open, but are more extraverted.[54]

Turning to workers instead of applicants, Table 15 shows that, compared to non-referred workers, referred workers have slightly fewer years of schooling in call-centers and trucking, and similar years of schooling in high-tech. Referred and non-referred workers have similar levels of cognitive ability—referred truckers score 0.12σ lower on an IQ test and referred high-tech workers score 12 points higher on the SAT test (neither difference is statistically significant). Turning to non-cognitive ability, there are a few interesting patterns in personality when we look one by one at Big 5 traits in Appendix Table C.4—referred workers tend to be slightly less agreeable and slightly more extraverted. However, on the overall Big 5 Index, differences are slight in all three industries between referred and non-referred workers.[55]

While referrals do not have higher levels of human capital and ability in Table 15, it could be they differ in terms of certain preferences, which we measure for truckers using lab experiments; e.g., referrals might be less likely to quit because they are more patient or have a greater risk tolerance for weekly swings in trucker income. Appendix Table C.6 does not support this. The one significant difference is that referrals are less trusting than non-referrals.

4 Productivity

Table 16 shows that referred and non-referred workers have similar productivity on many metrics, while referred workers have superior performance in terms of accidents and innovation. We regress productivity on referral status, normalizing the productivity variables when appropriate to ease comparisons across performance measures and industries.

In call-centers, there are no statistically significant differences between referred and non-referred workers on 4 of 5 productivity measures (all normalized), and on schedule adherence, referrals are slightly less productive (by 0.03σ). The number of observations varies by regression because which productivity measures are available varies by firm (and firms enter the sample at different dates), and because certain productivity measures are mea-

[54] Of the Big 5, conscientiousness, agreeableness, extraversion, and openness are usually considered desirable, whereas neuroticism is usually considered undesirable (e.g., Dal Bo, Finan and Rossi, 2013).

[55] Appendix Table C.5 shows that referred and non-referred workers also look similar in terms of additional measures of schooling and experience.

sured more frequently than others.[56] The estimates are precise. In Castilla's 2005 study of one call-center, referrals have 3.5% more phone calls per hour than non-referrals. Using a much larger sample, our 95% confidence interval allows us to rule out a referral performance advantage of more than 0.3%, meaning we can rule out differences 10 times smaller than in (Castilla, 2005).[57]

Turning to trucking, using miles as an outcome, the coefficient on referral is essentially 0, with a standard error of 0.01σ. Although miles is the main performance indicator, another very important measure of performance is driver accidents. Using a linear probability model, we estimate that referrals have a weekly accident probability that is about 0.14 percentage points below that of non-referrals. Given a baseline accident probability of around 2.4% per week, referrals have roughly a 6% lower risk of having an accident each week. One potential explanation for this result, separate from referrals having lower underlying accident risk, is that referrals may be assigned different roles in a firm than non-referrals. Although we control for the different types of work that different drivers are doing, it might be possible that referrals are receiving preferential treatment or work type assignment by the firm on some unobserved dimension. To address this, we take advantage of the fact that accidents are divided into "preventable," accidents the driver had control over, and "non-preventable," accidents the driver could not control. Referrals are 11% less likely to have preventable accidents, which is substantial, but only 1% less likely to have non-preventable accidents.

In high-tech, referrals have slightly higher subjective performance scores (by 0.04σ),[58] arguably the most important performance metric. For objective productivity, we create a single index equal to the average of six normalized objective productivity variables. Referrals and non-referrals have similar performance on this index, and the tight standard error means we can rule out small differences in either direction. Looking one by one at the different objective productivity measures in Appendix Table C.7, we also see little performance difference between referrals and non-referrals.

[56] E.g., sales conversion data are not available for firms that don't do sales work, and quality assurance data are only available on days where managers listen in to a worker's calls.

[57] (Castilla, 2005) finds that referrals have 0.7 more calls/hour (off a base of 20), or about 3.5% more. See Appendix A.5 for more on comparing our estimates with (Castilla, 2005). (Holzer, 1987a) and (Pinkston, 2012) show that referrals have higher subjective productivity ratings, using data from the Employment Opportunity Protection Project. (Pallais and Sands, 2013) show that referrals have higher productivity on oDesk.

[58] To put the 0.04σ magnitude into perspective, we re-did the regression in column 1 of Panel C of Table 16, using the logarithm of the performance rating instead of standardized performance. The coefficient on referral is 0.0037 (se=0.0011), indicating that referred workers have 0.4% higher subjective performance, which is an order of magnitude less than the referral differences observed for trucking accidents and for patents.

Column 3 of Panel C of Table 16 shows that referrals are significantly more likely to file patent applications than non-referrals. Patents are a standard measure of innovation in firms, and though relatively rare in patents per worker, are believed to be an important driver of firm performance Bloom and Van Reenen (2002). Given the skewed, count nature of patent production, we estimate negative binomial models. Referrals produce about 24% more patents than non-referrals.[59] To account for patent quality, we also study citation-weighted patents. Referrals produce 27% more citation-weighted patents than non-referrals (column 4).[60]

Our results in Table 16 include demographic controls, which are available for workers in trucking and high-tech. Appendix Table C.9 repeats Table 16 without demographic controls for trucking and high-tech, while adding job test score controls for call-centers. The resulting estimates are mostly similar.[61]

One potential concern in estimating the relationship between referral status and productivity is differential attrition. Among non-referrals, low-productivity workers might get "weeded out" after some period of time, whereas both low- and high-productivity referred workers may stick with the job. As a robustness check, we repeat our productivity regressions restricting to workers whose tenure exceeds some length, T, looking at productivity in the first T periods. As seen in Appendix Table C.8, the resulting estimates are relatively similar.

Why might referred workers be less likely to have accidents and more likely to develop patents? Why couldn't the trucking firm use past accidents to predict new accidents, and why couldn't the high-tech firm use past patents to predict who will develop new patents? In trucking, the firm requests state driving records for applicants, and applicants with past safety issues are removed from consideration. Managers believed that among driver applicants who are not excluded for safety issues, predicting who will be a safe driver is very difficult. Referrals may be providing additional information from social contacts about a driver's difficult-to-observe accident risk.

In high-tech, information about past patents is generally not requested by the firm on applications or in interviews, though applicants could potentially choose to report this information themselves. Managers highlighted to us that the workers are quite young.

[59] The overdispersion parameter, α, is 16.9 (se=1.64) in column 3, indicating a highly significant degree of overdispersion, suggesting use of a negative binomial instead of a poisson model (Cameron and Trivedi 2005).

[60] While patents are the most standard measure of innovation, we also have data on contribution of ideas to the firm's internal idea board. On this, referrals also have superior performance (see Appendix A.7).

[61] Our results in Table 16 also include tenure controls. Appendix Table C.10 repeats Table 16 without tenure controls. The resulting estimates are also mostly similar.

The median age at hire is 27, with many workers starting right out of college or graduate school. As we describe in Appendix A.7, most of these workers have no or little patenting history before joining the firm. Referrals might provide useful information about innovative potential, given that there is limited information on past innovation performance.

Productivity Spillovers for the Referrer. Another way that using referrals may increase productivity is if there is a productivity benefit to the *referrer* from making a referral. A referrer may feel empowered if the person they refer is hired, or they may become more productive because they have a friend to work with. We examine whether referrers become more productive after making referrals using data from trucking and high-tech, where we know which workers are making referrals and when. We regress productivity on a dummy for having already made a first referral, worker fixed effects, and time-varying controls.

Appendix Table C.12 shows that there are no significant gains in productivity or salary for referrers making referrals. In trucking, miles, accidents, and earnings do not change after a referral has been made. In high-tech, subjective performance ratings, patents, and salary do not change after a referral. The "zeroes" we estimate are fairly precise for earnings in both industries, for miles in trucking, and for subjective performance in high-tech. For accidents and patents, while the point estimates are close to zero, we are unable to rule out moderate-sized spillovers in either direction (reflecting that accidents and patents are relatively rare).

5 How Costly to the Firm are Referred vs. Non-referred Workers?

We consider whether referred and non-referred workers may differ in turnover, wages, and benefits, aspects which affect how costly workers are to firms.

Turnover. Despite similarities in observable characteristics, Table 17 shows that referred workers are substantially less likely to quit than non-referred workers. We estimate Cox Proportional Hazard models. In call-centers, referred workers are about 11% less likely to quit, both with and without job test score controls. In trucking, referred workers are also about 11% less likely to quit. Given the coefficient on driver home state unemployment rate of -0.07, the reduction in quitting among referred workers is of the same magnitude impact as that from a 1.5 percentage point increase in the driver's home state unemployment rate. In high-tech, referred workers are around 26% less likely to quit. The coefficients in trucking and high-tech are similar after controlling for demographics.

One explanation for why referrals are less likely to quit, which is unrelated to underlying quit propensities, is that referrals postpone quitting so as to help their referrer get a referral bonus. Specifically, for truckers and many call-center workers, there are bonuses for referrers where part or all of the bonus is contingent on the referral staying for some period of time. To examine this explanation, we exploit the sharp referral bonus thresholds in trucking at six months (for experienced drivers) and at three months (for inexperienced drivers) with a regression discontinuity design. As seen in Appendix Table C.13, the referral bonus appears to have little impact on quitting around the bonus tenure threshold, and the "zero effect" is precisely estimated. In addition, the largest quitting differences in Table 17 are for the high-tech firm, where referral bonuses are paid solely for the referral getting hired. These two pieces of evidence suggest (but do not prove) that differences in quit rates are unlikely to be driven by referral bonuses.

Although our analysis focuses on quits, which are more common than fires in all three industries, referred workers are also less likely to be fired.[62]

Wages. Table 18 shows regressions of log earnings on referral status and controls. In call-centers, referrals and non-referrals have similar earnings. In trucking, recall that earnings are closely related to miles since truckers are paid primarily by piece rate. As for call-centers, we find similar earnings for referrals and non-referrals. In high-tech, referred workers earn around 1.7% higher wages both with and without controlling for demographics. Referred high-tech workers are paid more even conditional on their characteristics, as we would expect when there is an important unobservable component to match quality.

As for the productivity results, we explore the importance of differential attrition for the wage results by repeating our wage regressions restricting to workers whose tenure exceeds some length, T, analyzing the first T periods. Appendix Table C.14 shows similar results.

Benefits. Another cost where referrals and non-referrals might differ is in terms of employee benefits. Speaking to managers at two of the call-center firms, at the trucking firm, and at the high-tech firm, benefit eligibility does not depend on referral status for any benefit. Despite this, there could still be differences in benefit utilization (conditional on benefit eligibility) between referrals and non-referrals. Unfortunately, we were unable to obtain comprehensive data on benefit utilization for any firm. Fortunately, for the trucking firm, we obtained information on usage of a few benefits: holiday time and vacation time.

[62] In all three industries, we can distinguish quits and fires in the data. Referred workers are 1%, 11%, and 37% less likely to be fired in call-centers, trucking, and high-tech, respectively. The difference is highly statistically significant for trucking, but not statistically significant for call-centers and high-tech.

Appendix Table C.16 shows that referred and non-referred truckers do not significantly differ on usage of these two benefits. In addition, at the four firms we spoke to, managers had no reason to believe that referrals and non-referrals differed in benefit utilization.

6 Profits

Having documented several differences in behaviors, we turn now to profits. We focus our profits analysis on trucking and call-centers because the production process is relatively simple. For high-tech, the production process is much more complicated than in call-centers or trucking, making it infeasible to perform a profits analysis.

We compare the average profits received when a firm hires a referred worker vs. when a firm hires a non-referred worker. When a position is posted, it lies vacant for $S \geq 0$ periods, during which a vacancy cost of c_V is incurred per period. Recruitment costs are incurred to hire the worker, including all the time and money required to process and consider the applicants. After getting hired, the worker begins production, during which he produces weekly profits of Z_t. For a worker, i, who stays with a firm for T periods, the profits from that worker are:

$$\pi_i = -H_i - \delta^S RB_i + \sum_{t=S+1}^{S+T} \delta^{t-1} Z_{it} \tag{7}$$

The first term, H_i, is the hiring cost, which is equal to the recruiting cost, R_i, plus the cost of the position being vacant. The second term, $\delta^S RB_i$, is the discounted referral bonus, where δ is the discount factor. RB_i includes referral bonuses paid at time of hire, as well as possibly paid later (if referral bonuses are contingent on the worker staying for some period of time). The third term, $\sum_{t=S+1}^{S+T} \delta^{t-1} Z_{it}$, is discounted profits from production.

The different terms in equation (7) will vary between call-centers and trucking. The profit formula is somewhat more complex for trucking than for call-centers, and we begin with that one first. For both industries, we assume an annual discount factor of 0.95.[63]

For trucking, the weekly profit function is $Z_t = y_t(P - mc - w_t) - FC - c_A A_t + (1 - E)\,\theta k_t q_t$. The first term, $y_t(P - mc - w_t)$, is per-mile earnings from operating a truck, where y_t is a driver's weekly miles, P is revenue per mile, mc is the non-wage marginal cost per mile (such as, truck wear and fuel costs), and w_t is the wage per mile. The second term, FC, is fixed costs per week (for example, back office support for the driver and the capital cost of

[63] Our results are robust to different discount factors; e.g., as seen in Appendix Table C.17, our results are similar if we assume an annual discount factor of 0.90.

the truck). The third term, $c_A A_t$, is weekly costs from trucking accidents, where c_A is the cost per accident and A_t is a dummy for having an accident. The fourth term, $(1 - E)\,\theta k_t q_t$, represents penalties collected by the firm through its training contracts when inexperienced workers quit Hoffman and Burks (2014).[64] Based on consultation with managers at the trucking firm, we assume that $P - mc = \$0.70$ per mile, $FC = \$450$ per week, $c_A = \$1,000$ for non-preventable accidents, and $c_A = \$2,000$ for preventable accidents.[65] In addition, for inexperienced drivers, we use a cost of \$2,500 for commercial driver's license training, plus five weeks of costly on-the-job training (details on training in Appendix A.1). Turning to the referral bonus, if the driver is an experienced referral, the firm pays \$500 when the driver is hired and an additional \$500 if he stays at least 26 weeks. If the driver is an inexperienced referral, the firm pays \$500 if he stays at least 13 weeks. We describe the recruiting cost further below.

For call-centers, our analysis is primarily based on cost and revenue information from one of the seven firms. We assume that the other firms have a similar cost and revenue structure. Workers are paid by the hour, and the weekly profit function is given by $Z_t = P_t - mc_t - w_t$. During training, the worker produces no revenues ($P_t = 0$) and has an average wage of \$9 per hour. Training lasts five weeks. After training, revenues are $P = \$26.70$ per hour and the wage is \$10 per hour. Both during and after training, there is overhead cost equal to 63% of the hourly wage (covering wages for trainers and supervisors, as well as worker benefits, building costs, and equipment costs). Weekly profits, Z_t, will solely be determined by how long a worker stays with the firm, and thus abstracts from productivity along the other dimensions (such as calls per hour or call quality). However, given that referrals and non-referrals did not significantly differ along those dimensions, this simplification should not affect our conclusions comparing profits from referrals vs. non-referrals. The referral bonus is set to \$50 and is paid upon the applicant being hired. Appendix A.1 provides further details.

We need to calculate R, the recruiting costs involved in making a hire, for both referred and non-referred workers. While it is common for firms to measure their average recruiting costs per hire, it is less common to do so separately for referred and non-referred workers. One strategy suggested by our conversations with the firms was that hiring

[64] For inexperienced drivers, the firm provides free commercial driver's license training, but workers must sign a contract specifying penalties if they quit too soon (see Appendix A.1 for details). E is a dummy for being an experienced worker; $\theta = 0.3$ is the approximate share of quit penalties collected by the firm; k_t is the quit penalty at a given tenure level; and q_t is a dummy for quitting.

[65] For profits analysis in trucking, we restrict attention to new hires during Oct. 2007-Dec. 2009, the period when we have information on who referred whom. We do this so we can analyze how profits vary by referrer productivity (see Table 19). Our conclusions are robust to using the full sample period (2003-2009) for profits analysis, as we discuss further in Appendix A.1.

costs generally scale linearly with the number of people being considered.[66] We make this assumption, allowing us to compute cost per hired for referred and non-referred workers. For call-centers, the average recruiting cost per hire is \$600; given the estimates in Table 14, this implies that the recruiting cost per hire for referred workers is about \$497, whereas that for non-referred workers is \$658. For trucking, the average recruiting cost per hire is about \$1,500, with a corresponding cost per hire for referred workers of \$1,063 and of \$1,609 for non-referrals.[67]

Last before computing profits, we need to account for the cost of vacancies. The per-period vacancy cost, c_V, is assumed to be the average profits earned by a randomly selected alternative worker (that is, average profits averaged over all periods and workers). For the number of vacant periods, S, recall that the call-center and trucking firms have high turnover. Rather than waiting for workers to quit, the firms are usually in hiring mode, with new candidates constantly moving through the pipeline. Thus, when workers quit, they are often replaced very quickly. Based on conversations with managers, we assume a vacancy duration of $S = 1$ weeks for call-centers and $S = 2$ weeks for trucking.[68]

Table 19 shows that referred workers produce substantially higher profits than non-referred workers. To compute profits, we add up profits for each worker, and then take an average over referred workers and over non-referred workers. In call-centers, referrals yield average discounted profit of \$1,453 per worker, whereas non-referrals yield average discounted profit of \$1,201 per worker. Likewise in trucking, referrals yield average discounted profit of \$3,547 per worker, compared to \$2,549 per worker for non-referrals. As we show in Appendix Table C.17, the results are relatively similar in robustness checks.

Decomposition. To help understand what is driving profit differences between referrals and non-referrals, we perform a simple decomposition. Profit differences between referred and non-referred workers can be divided into differences in recruiting costs, productivity, and turnover. To obtain the share for each category, we divide discounted profit differences from each category over the total difference in discounted profits between re-

[66] The assumption that recruiting costs per hire scale linearly with the number of people being considered is a strong one and could be violated in either direction. See Appendix A.1 for further discussion.

[67] Let ρ be the share of workers who are referred and c_H be the average recruiting cost per hire. Then, it follows that the recruiting cost per hire is $\frac{\Pr(Hire|r=0)}{\rho \Pr(Hire|r=0)+(1-\rho)\Pr(Hire|r=1)} c_H$ for a referred worker and $\frac{\Pr(Hire|r=1)}{\rho \Pr(Hire|r=0)+(1-\rho)\Pr(Hire|r=1)} c_H$ for a non-referred worker. See Appendix A.1 for a derivation, as well as for details on how we implement these formulas.

[68] The durations are broadly consistent with other studies for the US (e.g., Davis, Faberman and Haltiwanger, 2013; Wolthoff, 2012). We assume vacancy duration does not depend on referral status. See Appendix A.1 for discussion.

ferrals and non-referrals, excluding referral bonuses.[69] Let $\bar{\pi}_i = \pi_i + \delta^S RB_i$ be profits excluding referral bonuses. Given that hiring costs equal recruiting costs plus vacancy costs, and given we assume that vacancy durations are the same for referred and non-referred hires, the share of profit differences due to hiring costs is the same as the share of profit differences due to recruiting costs.

For call-centers, recall that referrals and non-referrals have very similar productivity (Panel A of Table 16); thus, profit differences can be decomposed into recruiting costs and turnover. The share of profit differences due to recruiting costs is $\frac{E(H|r=0)-E(H|r=1)}{E(\bar{\pi}|r=1)-E(\bar{\pi}|r=0)}$, which we calculate to equal roughly 53%. The remaining 47% of profit differences between referrals and non-referrals comes from referrals having lower turnover.

For trucking, recall that referred workers have similar miles to non-referred workers, but have fewer accidents (Panel B of Table 16). Thus, we measure the share of profit differences due to productivity differences using differences in accident rates. The share of profit differences due to differences in accident rates is $\frac{E(\sum_{t=S+1}^{S+T}\delta^{t-1}c_A A_t|r=0)-E(\sum_{t=S+1}^{S+T}\delta^{t-1}c_A A_t|r=1)}{E(\bar{\pi}|r=1)-E(\bar{\pi}|r=0)}$, which we compute to be roughly 2%. Differences in recruiting costs, defined the same way as in call-centers, are estimated to comprise 33% of profit differences. Differences in turnover comprise the remaining 65% of profit differences.

Why are differences in turnover important for profit differences between referrals and non-referrals? Part of this reflects that a worker's profit stream is carried out longer. In addition, lower turnover makes it so that a greater share of weeks worked are profitable. In both call-centers and trucking, most new workers require large initial investments by the firms in training. During call-center training, workers yield negative profits per week. In trucking, there is also training for new workers, as well as an initial period of increasing productivity.[70]

[69] Profit differences from turnover are defined as the profit differences remaining after subtracting out differences due to recruiting costs and due to productivity.

[70] If hiring were costless, and all workers yielded the same profit at all levels of tenure, then turnover would not be costly for firms. If this were the case, and if referred workers were less likely to quit, using profits per worker may overstate whether referred workers are actually more valuable to firms than non-referred workers. As an alternative to calculating profits per worker, we have also calculated profits per worker per week, defined as the total profit produced among all workers (referred or non-referred) divided by the total number of regular weeks worked for all workers (inclusive of weeks when the position is vacant). Using profits per worker per week, we continue to find that referred workers are significantly more profitable than non-referred workers. For call-centers, we calculate average profits per worker per week of \$85 for referred workers and \$72 for non-referred workers. For trucking, average profits per worker per week is \$94 for referred workers and \$74 for non-referred workers.

7 Heterogeneity

We now discuss how the value firms gain from hiring through referrals depends on two factors: the identity of the referrer and local labor market conditions. Unlike for our main results which were for all three industries, our heterogeneity analysis is performed primarily for trucking.

Referrers. Before analyzing how the identity of the referrer relates to the value of the referral hired, we first consider if there is a relationship between worker productivity and whether a worker makes referrals. Table 20 shows that workers with higher productivity are more likely to ever make a referral, both for truckers in miles (Panel A) and for high-tech workers in average subjective performance scores and patents per year (Panel B). We do not have information on who makes referrals for call-centers.

Table 21 shows that referred workers tend to have similar performance to their referrers on particular productivity metrics. We focus on trucking where we know who referred whom. We regress a driver's productivity in a given week on the average productivity of their referrer and controls. Panel A shows that if a referrer's average lifetime productivity is 100 miles per week above the mean, the person they refer is on average around 35 miles per week above the mean. Panel B shows that if the referrer has an accident at some point, the person they refer is roughly 17% more likely to have an accident. A confound to identifying behavioral homophily would be if there were a common shock affecting referrers and referrals (e.g., a shock to trucker productivity in a given area). We assuage this concern by including geographic controls for both referrers and referrals (see Appendix A.10 for further discussion).[71]

In Table 19, we see that truckers referred by above-median productivity drivers (measured in terms of miles) yield $6,490 in average profits, whereas referrals from below-median productivity workers yield $1,526, which is below average profits from non-referred workers.

Labor Market Conditions. For trucking, the data contain workers living all over the US over a 7-year time-frame, allowing us to examine referral differences in varied local labor market conditions. As seen in Appendix Table C.19, not only are referred applicants more likely to be hired and more likely to accept offers, but these differences are greater where unemployment is lower at time of application. Likewise, for trucking accidents, non-referred worker performance is negatively correlated with unemployment at time of hire, whereas for referred workers, there is less cyclical correlation.

[71] (Pallais and Sands, 2013) also find a correlation between the productivity of referrers and referrals.

For non-referred workers, our finding on accidents is consistent with asymmetric information models of firing and hiring (e.g., Gibbons and Katz, 1991; Nakamura, 2008), where those looking for work in good times tend to be of lower quality. As to why referred worker quality appears to be less countercyclical, trucking firm managers suggested that referred worker quality may be constrained by reputational concerns for incumbent workers. In the terminology of one manager, incumbent workers may be generally unwilling to refer a "doofus," even in booms when the average quality of those looking for work may be lower. To the extent that offer acceptance reflects match quality, that referrals differ in offer acceptance in booms is consistent with this interpretation. And, if firms anticipate these differences in match quality, referred applicants may be differentially more likely to be hired in booms.

8 Conclusion

Employee referrals are a topic of interest for many social scientists. While we know that referral-based hiring is common, relatively little is known about what firms gain from hiring referred vs. non-referred workers. Our paper takes a step toward filling this gap by combining personnel data from nine large firms in three industries.

In all three industries, referred applicants have a higher chance of getting hired than non-referred applicants, and referred workers are less likely to quit than non-referred workers. On a few productivity dimensions, most notably trucking accidents and high-tech innovation, referred workers have superior performance, but on many dimensions, referrals have similar productivity to non-referrals. In call-centers and trucking, referred workers produce significantly higher profits per worker than non-referred workers, with differences driven primarily by referrals having lower turnover and requiring less money to recruit. Productivity differences are either absent or do not play a first-order role in profit differences for these two industries.

For high-tech, it is not feasible to calculate worker-level profits, due to the complexity of high-tech production. Thus, we are unable to assess the relative importance of recruiting costs, productivity, and turnover for the value of hiring through referrals in high-tech. We speculate, though, that the relative importance of productivity may be higher for high-tech, due to the importance of innovation for high-tech production.

While it is not the goal of our paper to *test* between theories of referral-based hiring, our results are still relevant for theory. Consistent with learning, homophily, and peer benefit theories, referred applicants are more likely to be hired than non-referred applicants and referred workers are less likely to quit compared to non-referred workers. Consistent

with homophily theories, high-ability workers are more likely to make referrals, as is the tendency of workers to refer those of similar ability. Potentially consistent with all three classes of theories, referred workers yield higher profits per worker than non-referred workers.

However, we also find results that seem inconsistent with existing theories. Referred workers do not consistently have higher wages than non-referred workers, nor are referrals consistently more productive across different metrics. In addition, in seeming contrast with learning and homophily theories, referrals do not have superior scores on unobserved-to-the-firm dimensions of quality. Part of this could reflect that existing theories do not generally include referral bonuses, which are observed in all three industries. When workers receive bonuses for making referrals, they may sometimes recommend unqualified candidates, which may work against referred workers having higher wages or being more productive.[72]

Outside of learning, homophily, and peer benefit theories, another explanation which has been proposed to explain differences between referrals and non-referrals is that referrals may have worse outside options Loury (2006). The worse outside option explanation is consistent with referred workers being less likely to quit and with referred applicants being more likely to accept job offers; however, it does not explain why referred workers have fewer trucking accidents and are more innovative than non-referred workers.

Methodologically, we illustrate both the promise and limitations in combining large personnel datasets. Personnel data can provide large-scale, inside-the-firm information, which may be valuable for bringing data to bear on a whole host of economic questions. Although using personnel data often leads to questions of external validity, by *combining* data from different industries, we can examine whether results are consistent across industries, which is largely the case for our findings. Still, even with nine firms in three industries, we acknowledge that our results may not be generalizable to all firms in the economy, though we believe our methodology represents a significant advance relative to existing knowledge. A significant limitation is that personnel policies are rarely randomized by firms.[73] Further empirical research on referrals using natural or randomized experiments is also sorely needed, and should be complementary to our paper.

[72] It does seem possible, however, that having referral bonuses could increase referral quality relative to having no bonus, particularly if bonuses are conditional on being hired or on worker performance (e.g., if it is costlier for an incumbent worker to find a high-quality candidate to refer than to find a low-quality candidate to refer). The only theory we are aware of which incorporates referral bonuses is that in the recent field experiments of (Beaman and Magruder, 2012) and (Beaman et al., 2013). See Appendix A.11 for more discussion on the relevance of our results for existing theories.

[73] For field experimental research on referrals, see (Beaman and Magruder, 2012), (Beaman et al., 2013), and (Pallais and Sands, 2013).

Table 12: Summary of Data Elements

	Call-centers	Trucking	High-tech
Referral status	W,A	W,A	W,A
Productivity	W	W	W
Demographics	A	W,A	W,A
Cognitive ability	W,A	W	W
Personality	W,A	W	W
Experimental games		W	
Who makes referrals		W	W
Who referred whom		W	

Notes: This table summarizes the data elements from the three industries. "W" means a data element is available for workers. "A" means a data element is available for applicants. For details on data collection and more on why different elements are available for different industries, see the Data Appendix (Appendix B).

Table 13: Sample Means

	Call-center firms	**Trucking firm**	**High-tech firm**
Referred	36%	20%	33%
Years of schooling	13	13	17
Female	62%	8%	Confidential
Black	21%	18%	Confidential
Hispanic	21%	5%	Confidential
Age at hire	26	39	29
Log(Salary)	4.32	6.53	Confidential
Accidents		0.024	
Preventable accidents		0.011	
Patents			0.0047
Have accident or patent		51%	6%
Number of workers	73,595	N	25,282

Notes: This table provides sample means for workers, as well as the sample size for workers. For applicants, the sample size is 349,562 applicants for call-centers; A applicants for trucking; and 1,415,320 applicants for high-tech. Some information cannot be shown in the table due to confidentiality requirements. For the trucking firm, exact sample sizes are withheld to protect firm confidentiality, $A >> 100,000$, $N >> 10,000$. Entries are blank if a variable is not applicable. In all three industries, means for referral status, schooling, gender, race, and age are calculated at the worker level. For call-centers, we can calculate sample means of demographics at the worker level, but we cannot link demographic information to our main data on worker outcomes (explained in Appendix B.2). Call-center mean salary is calculated at the worker-day level. For trucking, mean salary and accidents are calculated at the worker-week level. "Accidents" is the share of worker-weeks where a driver has an accident. The company's definition of an accident is quite broad and includes serious as well as relatively minor accidents. "Preventable accidents" are accidents the driver had control over. For high-tech, mean patents is calculated at the worker-month level (i.e., "patents" is the average number of patent applications per worker-month). "Have accident or patent" equals one if the worker has at least one trucking accident or one patent in the sample.

Table 14: Referred Applicants are More Likely to be Hired and More Likely to Accept Job Offers

Panel A: Call-center	(1) Hired	(2) Hired	(3) Accept offer	(4) Accept offer
Referral	0.063***	0.060***	0.051**	0.050**
	(0.002)	(0.002)	(0.025)	(0.025)
Demographic controls	No	Yes	No	Yes
Observations	349,562	349,562	2,362	2,362
R-squared	0.062	0.066	0.208	0.210
Mean dep var if ref=0	0.19	0.19	0.56	0.56
Panel B: Trucking	(1) Hired	(2) Hired	(3) Accept offer	(4) Accept offer
Referral	0.101***	0.098***	0.073***	0.073***
	(0.002)	(0.002)	(0.003)	(0.003)
Demographic controls	No	Yes	No	Yes
Observations	A	A	$0.22A$	$0.22A$
R-squared	0.067	0.068	0.070	0.071
Mean dep var if ref=0	0.17	0.17	0.80	0.80
Panel C: High-tech	(1) Hired	(2) Hired	(3) Accept offer	(4) Accept offer
Referral	0.0027***	0.0027***	0.027**	0.026**
	(0.0003)	(0.0003)	(0.011)	(0.011)
Demographic controls	No	Yes	No	Yes
Observations	1,175,016	1,175,016	5,738	5,738
R-squared	0.586	0.586	0.597	0.598
Mean dep var if ref=0	0.0028	0.0028	0.74	0.74

Notes: This table presents linear probability models analyzing whether referred applicants are more likely to be hired and, conditional on receiving an offer, whether they are more likely to accept. An observation is an applicant. Robust standard errors in parentheses. In Panel A, regressions include month-year of application dummies and location fixed effects. In columns 1-2, demographic controls are race, age, gender, and years of schooling. For columns 3-4, the only available demographic control is years of schooling. In Panel B, regressions include month-year of application dummies, work type controls, and state fixed effects. Demographic controls are age and gender. The exact number of applicants, A, is withheld to protect firm confidentiality, $A >> 100,000$. In Panel C, regressions include month-year of application dummies, job position ID dummies, and office location dummies. Demographic controls are race and gender. The sample is applicants for engineering and computer programmer positions from June 2008-May 2011. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 15: Schooling, Cognitive Ability, and Non-cognitive Ability: Comparing Referred vs. Non-referred Applicants, as well as Referred vs. Non-referred Workers

Panel A: Call-centers	(1)	(2)	(3)	(4)	(5)	(6)
Dep var:	Schooling in years		Intelligence (normalized)		Big 5 index (normalized)	
Applicants or workers?	App	Workers	App	Workers	App	Workers
Referral	-0.100***	-0.068***	-0.024***	-0.029***	-0.017***	-0.013***
	(0.012)	(0.023)	(0.004)	(0.008)	(0.002)	(0.003)
Observations	47,360	9,956	302,022	62,110	341,788	73,214
R-squared	0.088	0.095	0.137	0.196	0.087	0.092

Panel B: Trucking	(1)	(2)	(3)
Dep var:	Schooling	IQ	Big 5 index
(sample is workers)	in years	(normalized)	(normalized)
Referral	-0.231	-0.115	-0.023
	(0.150)	(0.101)	(0.061)
Observations	628	598	628
R-squared	0.093	0.120	0.076

Panel C: High-tech	(1)	(2)	(3)
Dep var:	Schooling	SAT total	Big 5 index
(sample is workers)	in years	score (mean=1401)	(normalized)
Referral	0.006	11.96	0.015
	(0.027)	(9.45)	(0.026)
Observations	10,890	899	1,853
R-squared	0.188	0.210	0.040

Notes: This table compares schooling, cognitive ability, and non-cognitive ability among referred vs. non-referred applicants, as well as referred vs. non-referred workers. An observation is an applicant in the odd columns of Panel A, whereas an observation is a worker in all the other regressions. Robust standard errors in parentheses. For an explanation of how the different variables are measured and defined, see Appendix B. For the call-centers, the applicant regressions (odd-numbered columns) include month-year of application dummies, location dummies, and controls for race, age, and gender. The worker regressions (even-numbered columns) include the same controls, except they include month-year of hire dummies instead of month-year of application dummies, and also include client dummies. The call-center schooling analysis is based on one firm, whereas the analyses of cognitive and non-cognitive ability are based on seven firms. For trucking, the regressions include month-year of hire dummies, work type controls, state dummies, and controls for race, age, gender, and marital status. The drivers here are from the same training school and were hired in late 2005 or 2006. For high-tech, the regressions include month-year of hire dummies, job category dummies, job rank dummies, office location dummies, and controls for race, age, and gender. The SAT and Big 5 Index data are from a voluntary 2006 survey done by the high-tech firm's HR department. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 16: Referrals and Productivity

Panel A: Call-centers	(1)	(2)	(3)	(4)	(5)
Dep var:	Adherence share	Average handle time	Sales conversion	Quality assurance	Customer satisfaction
Referral	-0.027**	0.001	-0.014	0.016	0.003
	(0.013)	(0.010)	(0.008)	(0.017)	(0.003)
Observations	152,683	749,616	134,386	31,908	603,860
Clusters	3,136	12,496	3,192	2,864	11,859
R-squared	0.142	0.563	0.725	0.175	0.034

Panel B: Truckers (accident coefs are multiplied by 100)	(1) Miles	(2) Accident	(3) Preventable accident	(4) Non-preventable accident (placebo)
Referral	-0.001	-0.136***	-0.121***	-0.008
	(0.009)	(0.032)	(0.021)	(0.018)
% reduced accident risk	NA	6%	11%	1%
Observations	0.83M	M	M	M
Clusters	0.85N	N	N	N
R-squared	0.082	0.0032	0.0039	0.0008

Panel C: High-Tech	(1)	(2)	(3)	(4)
Dep var:	Subjective performance	Objective performance	Patents	Citation-weighted patents
Referral	0.035***	0.004	0.236***	0.272***
	(0.012)	(0.007)	(0.077)	(0.091)
Observations	104,255	289,689	333,492	333,492
Clusters	16,546	11,123	17,190	17,190
R-squared	0.093	0.170		

Notes: Standard errors clustered by worker in parentheses. In Panel A, all columns are OLS regressions. Productivity is one of 5 normalized measures. An observation is a worker-day. The controls are month-year of hire dummies, month-year dummies, a fifth-order polynomial in tenure, location dummies, client dummies, and the number of times that each outcome was measured to compute the dependent variable. In Panel B, all columns are OLS regressions. In column 1, productivity is measured in normalized miles driven per week (trimming zero mile weeks, as well as the lowest and highest 1% of the non-zero miles observations), whereas in columns 2-4, productivity is a dummy for having an accident in a given week. An observation is a worker-week. All regressions include month-year of hire dummies, month-year dummies, a fifth-order polynomial in tenure, driver training contracts, work type controls, training school dummies, state dummies, the annual state unemployment rate, and controls for gender, race, marital status, and age. The exact sample size is withheld to protect firm confidentiality, $M >> 100,000$, $N >> 10,000$. In Panel C, columns 1-2 are OLS regressions (with normalized subjective and objective performance) and columns 3-4 are negative binomial models. An observation is a worker-quarter in column 1 and a worker-month in columns 2-4. All regressions include a fifth-order polynomial in tenure, job category dummies, job rank dummies, office location dummies, and controls for race, age, gender, and education. In addition, column 1 includes quarter-year of hire dummies and quarter-year dummies; column 2 includes month-year of hire dummies and month-year dummies; and columns 3-4 include month-year of hire dummies.* significant at 10%; ** significant at 5%; *** significant at 1%

Table 17: Referrals and Quitting

Industry:	Call-centers		Trucking		High-tech	
	(1)	(2)	(3)	(4)	(5)	(6)
Referral	-0.107***	-0.108***	-0.110***	-0.103***	-0.262***	-0.253***
	(0.014)	(0.014)	(0.020)	(0.020)	(0.061)	(0.062)
Current state unemployment rate				-0.074***		
				(0.016)		
Additional Controls	No	Yes	No	Yes	No	Yes
Observations	4,446,155	4,446,155	0.94M	0.94M	301,272	301,272

Notes: This table examines whether a worker's referral status predicts quitting. All specifications are Cox proportional hazard models with standard errors clustered by worker in parentheses. For call-centers, an observation is a worker-day. Both columns 1 and 2 include month-year of hire dummies, location dummies, and client dummies. We restrict to workers who are with the company for 200 days or less. The additional controls in column 2 are job test score controls. For trucking, an observation is a worker-week. Both columns 3-4 include month-year of hire dummies, month-year dummies, driver training contracts, work type controls, training school dummies, and state dummies. The additional controls in column 4 are gender, race, marital status, and age. The exact sample size is withheld to protect firm confidentiality, $M >> 100,000$, $N >> 10,000$. For high-tech, an observation is a worker-month. Both columns 5-6 include month-year of hire dummies, job category dummies, job rank dummies, and office location dummies. The additional controls in column 6 are race, age, gender, and education. For all three industries, time since hire is fully controlled for (see Appendix A.8 for details). * significant at 10%; ** significant at 5%; *** significant at 1%

Table 18: Referrals and Wages (OLS, Dep Var = Log(Salary))

Industry:	Call-centers		Trucking		High-tech	
	(1)	(2)	(3)	(4)	(5)	(6)
Referral	0.0002	-0.0004	0.0002	0.0041	0.017**	0.017**
	(0.0021)	(0.0021)	(0.0041)	(0.0040)	(0.0065)	(0.0065)
Additional Controls	No	Yes	No	Yes	No	Yes
Observations	634,153	634,153	$0.65M$	$0.65M$	245,270	245,270
R-squared	0.275	0.284	0.064	0.067	0.518	0.524

Notes: This table examines whether referred workers earn higher salaries. Standard errors clustered by worker in parentheses. For call-centers, an observation is a worker-day. Both columns 1 and 2 include month-year of hire dummies, month-year dummies, a fifth-order polynomial in tenure, location dummies, and client dummies. The additional controls in column 2 are job test score controls. There are 11,174 workers. The data are from two of the call-center firms. For trucking, an observation is a worker-week. Both columns 3-4 include month-year of hire dummies, month-year dummies, a fifth-order polynomial in tenure, driver training contracts, work type controls, training school dummies, state dummies, and the annual state unemployment rate. The additional controls in column 4 are gender, race, marital status, and age. There are $0.74N$ workers. The exact sample size is withheld to protect firm confidentiality, $M >> 100,000$, $N >> 10,000$. For high-tech, an observation is a worker-month. Both columns 5-6 include month-year of hire dummies, month-year dummies, a fifth-order polynomial in tenure, job category dummies, job rank dummies, and office location dummies. The additional controls in column 6 are controls for race, age, gender, and education. There are 10,655 workers. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 19: Profits per Worker Among Referred and Non-referred Workers

Industry	Profits per Worker Call-centers	Profits per Worker Trucking
Referred (overall)	\$1,453	\$3,547
Non-referred (overall)	\$1,201	\$2,549
Decomposition:		
Share profit difference from lower turnover	46.6%	64.8%
Share profit difference from higher productivity	0%	1.7%
Share profit difference from lower recruiting costs	53.4%	33.4%
Comparisons Based on the Referrer:		
Referred (matched sample)		\$3,810
Referred, referring worker w/ above median productivity		\$6,490
Referred, referring worker w/ below median productivity		\$1,526

Notes: We present profits per worker for call-center workers and truckers. As described in Section 6, the profits for referred workers includes the cost of paying referral bonuses. The decomposition of profit differences into lower turnover, higher productivity, and lower recruiting costs is as described in Section 6. The "matched sample" refers to drivers for which we know who referred them (that is, it is based on administrative data from the trucking firm's employee referral program). We calculate profits when the driver referring them has above median productivity in miles and when the driver referring them has below median productivity in miles. For both industries, we assume an annual discount factor of 0.95.

Table 20: High-Productivity Workers Are More Likely to Ever Make a Referral

Panel A: Trucking	(1)	(2)
Miles per week (normalized)	0.0047***	0.0052***
	(0.0013)	(0.0013)
Demographic controls	No	Yes
R-squared	0.031	0.033
Mean dep var	0.042	0.042
Panel B: High-tech	(1)	(2)
Subjective performance rating	0.027***	0.027***
(normalized)	(0.004)	(0.004)
Patents per year	0.033*	0.030*
	(0.018)	(0.016)
Interview score		0.012***
		(0.004)
Incumbent worker was referred		0.043***
		(0.006)
Demographic controls	No	Yes
R-squared	0.151	0.157
Mean dep var	0.180	0.180

Notes: This table presents OLS regressions of whether an employee makes a referral on the employee's average productivity. An observation is an incumbent worker. Robust standard errors in parentheses. In Panel A, all regressions include month-year of hire dummies, work type controls, state dummies, and tenure at the job. Making a referral is defined according to administrative employee referral program data. Demographic controls are gender, race, marital status, and age. The sample size is $0.63N$ workers. The exact sample size is withheld to protect firm confidentiality, $N >> 10,000$. In Panel B, all regressions include month-year of hire dummies, job category dummies, job rank dummies, office location dummies, and tenure at the job. Demographic controls are race, age, gender, and education. The sample size is 15,810 workers. * significant at 10%; ** significant at 5%; *** significant at 1%

Chapter 3: Corporate Prediction Markets: Evidence from Google, Ford, and Firm X

This chapter is co-authored with Eric Zitzewitz

1 Introduction

The success of public prediction markets such as the Iowa Electronic Markets has led to considerable interest in running prediction markets inside organizations. Interest is motivated in part by the hope that prediction markets might help aggregate information that is trapped in hierarchies for political reasons, such as perceptions that messengers are punished for sharing bad news (e.g., Prendergast, 1993). A popular book arguing the benefits to organizations from harnessing The Wisdom of Crowds (Suroweicki, 2004) was a notable source of enthusiasm. Markets in organizations face issues distinct from public prediction markets, however. If markets are run on topics of strategic importance, there is often a need to limit participation for confidentiality reasons. Limited participation makes markets thinner. In thinner markets, biases in participants' trading may have more influence on prices. Employees may optimistically bias their trading in order to influence management's view of their projects' performance or prospects. In addition to strategic biases, members of an organization may not be sufficiently dispassionate when making predictions. Employees may select employers based partly on optimism about their future, and belonging to an organization may likewise engender a favorable view of its prospects. Employees may suffer from other biases, such as probability misperceptions or loss aversion. Whereas in public prediction markets arbitrageurs may enter to eliminate any resulting inefficiencies, in corporate prediction markets, this entry may be less feasible.

This paper examines the efficiency of corporate prediction markets by studying markets at three major companies: Google, Ford Motor Company, and Firm X.[74] These firms' markets were chosen because they are among the largest corporate markets we are aware of and they span the many diverse ways that other companies have employed prediction markets. Our sample includes all of the major types of corporate prediction markets we are aware of, including markets that forecast demand, product quality, deadlines being

[74]Firm X is a large, privately held, profitable basic materials and energy conglomerate headquartered in the Midwestern United States, but with global operations.

Table 21: Homophily in Truckdriver Referrals

Panel A: Miles	(1)	(2)	(3)
Avg miles per week of referring driver	0.350***	0.339***	0.356***
	(0.037)	(0.039)	(0.039)
Mean dep var	1,652	1,652	1,652
Demog controls for referred driver	No	Yes	Yes
Demog controls for referrer	No	No	Yes
R-squared	0.187	0.191	0.194
Panel B: Accidents	(1)	(2)	(3)
Referring driver ever had an accident	0.00383*	0.00373*	0.00371*
	(0.00209)	(0.00211)	(0.00209)
Mean dep var	0.0218	0.0218	0.0218
Demog controls for referred driver	No	Yes	Yes
Demog controls for referrer	No	No	Yes
R-squared	0.010	0.010	0.011

Notes: This table present OLS regressions of the productivity of referred workers on the productivity of referrers. The sample is restricted to matched referred truckers hired in Oct. 2007-Dec. 2009. Further, the sample is restricted to workers whose referrer was hired between 2003 and 2009. In Panel A, the dependent variable is worker productivity in miles per week (the estimated coefficients are similar if we exclude zero-mile weeks from the sample). An observation is a worker-week. The sample size is $0.017M$ worker-weeks. All regressions include month-year of hire dummies, month-year dummies, a fifth-order polynomial in tenure, driver training contracts, work type controls, training school dummies, state dummies, and the annual state unemployment rate for the *referred* driver. They also include work type controls, state dummies, and tenure at date of referral for the *referring* worker. Demographic controls refer to controls for gender, race, marital status, and age for both the referred and referring worker. In Panel B, we present linear probability models where the dependent variable is whether the worker has an accident in a given week. An observation is a worker-week. The sample size is $0.018M$ worker-weeks. Controls are the same as in Panel A. Exact sample size is withheld to protect firm confidentiality, $M >> 100,000$. * significant at 10%; ** significant at 5%; *** significant at 1%

met, and external events. It includes both markets into which the entire company was invited to trade and markets available only to hand-picked employees or specific divisions. It also includes diversity in the strength of incentives and in market mechanisms and design. Table 1 summarizes these characteristics and shows examples of other major corporations that we are aware of having used markets similar to those in our sample.

Despite large differences in market design, operation, participation, and incentives, we find that prediction market prices at our three companies are well-calibrated to probabilities and improve upon alternative forecasting methods. Ford employs experts to forecast weekly vehicle sales, and we show that contemporaneous prediction market forecasts outperform the expert forecast, achieving a 25% lower mean squared error (p-value 0.104). Google and Firm X did not have formal expert forecasts of the variables being predicted by its markets, but for markets forecasting continuous variables, expert opinion was used in the construction of the securities. Google and Firm X created securities tracking the probability of the outcome falling into one of 3 or more bins, and an expert was asked to create bin boundaries that equalized ex ante probabilities. Firm X also ran binary markets on whether a variable would be above or below an "over/under" median forecast. At both Google and Firm X market-based forecasts outperform those used in designing the securities, using market prices from the first 24 hours of trading so that we are again comparing forecasts of roughly similar vintage.

The strong relative predictive performance of the Google and Ford markets is achieved despite several pricing inefficiencies. Googles markets exhibit an optimism bias. Both Google and Ford's markets exhibit a bias away from a naive prior (1/N, where N is the number of bins, for Google and prior sales for Ford). We find that these inefficiencies disappear by the end of the sample, however. Improvement over time is driven by two mechanisms: first, more experienced traders trade against the identified inefficiencies and earn higher returns, suggesting that traders become better calibrated with experience. Second, traders (of a given experience level) with higher past returns earn higher future returns, trade against identified inefficiencies, and trade more in the future. These results together suggest that traders differ in their skill levels, that they learn about their ability over time, and that self-selection causes the average skill level in the market to rise over time.

Our Google data, which include information on traders job and product assignments, allow us to examine the role played by insiders in corporate markets. If we define an insider narrowly, as a team member for a project that is the subject of a market, or as a friend of a team member (as reported on a social network survey), we find that insiders account for 10 percent of trades, that insiders are more likely to be on the optimistic side of a market, and that insiders' trades are not systematically profitable or unprofitable. If we

instead define insiders more broadly, as those traders we would expect to be most central to social and professional networks at Google (software engineers located at the Mountain View headquarters with longer tenure), we find that these traders are less optimistic and more profitable than other traders. So while a small number of insiders may trade optimistically in markets on their own projects, perhaps reflecting either overconfidence or ulterior motivations, they are offset by a larger group of traders who also have relevant expertise and fewer professional reasons to be biased.

Taken together, these results suggest that despite limited participation, individual traders biases, and the potential for ulterior trading motives, corporate prediction markets perform reasonably well, and appear to do so for reasons anticipated by theory. Equilibrium market prices reflect an aggregation of the information and any subjective biases of their participants (Grossman, 1976; Grossman and Stiglitz, 1980; Ottaviani and Sorensen, 2014). Traders with an outside interest in manipulating prices may attempt to do so (Allen and Gale, 1992; Aggarwal and Wu, 2006; Goldstein and Guembel, 2008), but, as emphasized by Hanson and Oprea (2009), the potential for manipulation creates incentives for other traders to become informed. Similar logic applies to traders with subjective biases – their presence creates incentives for participation by informed traders. Our results of initial inefficiency disappearing, with more experienced and skilled traders trading against the inefficiencies, are consistent with this set of predictions.

Our paper contributes to an increasingly extensive empirical literature on prediction markets and a much smaller literature describing experimental markets run at companies. Forsythe, et. al. (1992) and Berg, et. al. (2008) analyze the empirical results from the Iowa Electronic Market on political outcomes, finding that markets outperform polls as predictors of future election results. Wolfers and Zitzewitz (2004) and Snowberg, Wolfers, and Zitzewitz (2005, 2012) examine a broader set of markets, again concluding that prediction markets at least weakly outperform alternative forecasts. A series of papers have used prices from public prediction markets to estimate the effects of policies and political outcomes (e.g., Rigobon and Sack, 2005; Snowberg, Wolfers and Zitzewitz, 2007a and 2007b; Wolfers and Zitzewitz, 2009).

While most of the smaller literature on corporate prediction markets is empirical, Ottoviani and Sorenson (2007) present a theoretical framework for prediction markets inside organizations. The empirical literature begins with Ortner (1998), which reports on markets run at Siemens about project deadlines.Chen and Plott (2002) and Gillen, Plott, and Shum (2013) report on sales forecasting markets run inside Hewlett-Packard and Intel, respectively. Hankins and Lee (2011) describe three experimental prediction markets run at Nokia, including one predicting smart phone sales. Most of these experiments are much smaller than the markets we study. The largest is the sales forecasting experiment at Intel,

which is about 60 percent as large as the sales forecasting portion of the markets run at Ford.[75]

Our study differs from these prior and concurrent studies in several ways. First, the larger scale of the markets we analyze allows us to test for market inefficiencies with great statistical power, as well as to characterize differences in efficiency over time and across types of markets. Second, the microdata available on Google participants allow us to identify the characteristics of employees who trade with and against inefficiencies. Third, the markets we analyze are non-experimental in the sense that they were initiated by the companies themselves.[76] They are thus more field than field experiment. While a downside to field data is that some research opportunities may have been missed, an advantage is that the markets we study are more likely to be representative of prediction markets as companies will implement them in the future.

Prior research informs our analyses of the specific inefficiencies we examine. Building on Ali's (1977) analysis of horseracing parimutuel markets, Manski (2006) shows that two common features of prediction markets – budget constraints and the skewed payoff structure of binary securities – can combine to cause a longshot bias in which prices of low-priced securities will be upwardly biased relative to median beliefs. Gjerstad (2005), Wolfers and Zitzewitz (2005), and Ottaviani and Sorensen (2014) generalize this result to a broader set of risk preferences and information environments, showing that the sign of any bias is ambiguous.

The optimistic bias we document could either arise from genuine optimism, an uncorrected-for bias in information (e.g., the "inside view" of Kahneman and Lovallo, 1993) or a conscious effort to manipulate prices. As Hanson and Oprea (2009) argue, the extent to which a (consciously or unconsciously) biased trader will affect prices depends on the ability of other traders to become informed and enter the market. In past episodes of apparent price manipulation in public prediction markets, other traders entered and traded against the apparent manipulation, reducing its impact on prices.[77] The price impact of manipulators in experimental markets is examined by Hanson, Oprea, and Porter (2006) and Jian and Sami (2012), with the former concluding that manipulation does not affect the accuracy of prices and the latter concluding that effects depend on the correlation of signals given to participants. In the field, the robustness of a corporate prediction market may depend on the ability and willingness of unbiased traders to enter the market and

[75]The Intel sales forecasting markets cover 46 product*period combinations, while the sales forecasting component of Fords markets cover 78 product*period combinations (6 models times 13 weeks).

[76]The markets at Google were created by a group that included an author on this paper (Cowgill), but several years prior to his beginning his career as an economist.

[77]See Wolfers and Zitzewitz (2004 and 2006b), Rhode and Strumpf (2004 and 2006), Hansen, Schmidt, and Strobel (2004), and Newman (2012) for discussions.

become informed, which may be constrained by limited participation.

To the extent that the optimistic bias we document is behavioral, our results also speak to the growing literature about overconfidence and excess optimism in organizations. Recent work shows that worker overconfidence has significant economic consequences for workers and firms. A theoretical literature explores how optimism may improve motivation of employees (Benabou and Tirole, 2002 and 2003; Compte and Postlewaite, 2004) or lead to risk-taking that generates positive externalities (Bernardo and Welch, 2001; Goel and Thakor, 2008).

Other work has discussed how employee optimism and equity compensation interact. Optimistic employees may overvalue equity compensation, and thus be cheaper to compensate. As Bergman and Jenter (2007) point out, however, the simplest version of this explanation of equity compensation ignores the fact that employees of public companies can buy equity with their cash compensation. Shorting employer equity is difficult for most employees, so when equity is included in compensation, it in practice likely provides a lower bound on employees' stock exposure. Oyer and Schaefer (2005) argue that firms may use a mixture of equity and cash compensation because it causes employees who are optimistic about firm prospects to self-select into employment, which could be beneficial if optimistic employees work harder, or if they take risks that are beneficial to their employers.

Empirical work finds that employee optimism or overconfidence is correlated with risk-taking, but suggests that the benefits of optimism-induced risk taking may be mixed. Hirshleifer, Low, and Teoh (2012) find that firms with overconfident CEOs invest more in research and development and attain more and more highly cited patents. Larkin and Leider (2012) find that overconfident employees select more convex incentives contracts, and Hoffman and Burks (2013) finds that overconfident truckers select more training. In both cases, employee overconfidence lowers costs for firms. At the same time, Malmendier and Tate (2008) find that overconfident CEOs undertake mergers that are associated with lower stock performance for their employers.

Corporate prediction markets provide tools for both measuring and potentially correcting employee optimism. The optimistic bias in Google's markets, and the fact that appears to arise from new employees who become better calibrated with experience, is interesting in light of the aforementioned work. Firm X told us that a primary motivation for running markets was a desire to help senior managers become better calibrated forecasters. It is possible that in their context of economic forecasting and strategic planning, correct calibration is paramount, while in other contexts, correcting employee optimism may or may not be in an employer's interests.

The remainder of the paper is organized as follows. The next section provides background on the markets at Google, Ford, and Firm X. The following section presents our empirical analysis of the efficiency and inefficiencies of these markets. A discussion concludes.

2 Background on the Corporate Prediction Markets

The three companies whose prediction markets we examine, Google, Ford, and Firm X, are in different industries, have distinct corporate cultures, and took different approaches in their prediction market implementations. We will describe them in turn, and then discuss commonalities and differences.

2.1 Background on the Companies and Their Markets

Google is a software company, headquartered in Mountain View, CA, with a highly educated workforce and a high level of internal transparency. Its prediction markets began as a "20% time project" initiated by a group of employees that included a co-author of this paper (Cowgill) prior to his PhD. Google opened its prediction markets to all employees. The focus of Google's markets were whether specific quarterly "Objectives and Key Results" (OKRs) would be achieved. OKRs are goals of high importance to the company (e.g., the number of users, a third-party quality rating, or the on-time completion of key products). The attainment of OKRs was widely discussed within the company, as described by Levy (2011):

> OKRs became an essential component of Google culture. Four times per year, everything stopped at Google for division-wide meetings to assess OKR progress. [...]
>
> It was essential that OKRs be measurable. An employee couldn't say, "I will make Gmail a success" but, "I will launch Gmail in September and have a million users by November." "It's not a key result unless it has a number," says [senior executive] Marissa Mayer.

Google's markets were run with twin goals: 1) aggregating information for management about the success of an important project and 2) further communicating management's interest in the success of the project. Prediction market prices were featured on the company intranet home page, and thus were of high visibility to employees. One particular

anecdote illustrates how the markets impacted executive behavior. At a company-wide meeting, a senior executive made the following comment:

> [...] I'd like to talk about one of our key objectives for the last six quarters. During this entire time, one of our quarterly objectives has been to hire a new senior-level executive in charge of an important new objective to work on [redacted].
>
> We have failed to do this for the past six quarters. Judging from the [internal prediction markets], you saw this coming. The betting on this goal was extremely harsh. I am shocked and outraged by the lack of brown-nosing at this company [laughter].
>
> We've decided to look into the problem and figure it out, and I think we have gotten to the bottom of it. We've made some adjustments in the plans for the new team, and made some hard decisions about exactly what type of candidates we're looking for. We're expecting to finally get it done in the upcoming quarter – which would take this objective off the list once and for all.

The objective in question was indeed completed that quarter.

While the prediction market project aspired to cover every company-wide OKR, information on some projects needed to be too compartmentalized for them to be appropriate for a prediction market with mass participation. Thus a cost of wide participation was that some topics were necessarily off limits. Despite this, over 60 percent of quarterly OKRs were covered by markets.

The markets on OKRs spanned the topics typically covered in other corporate prediction markets, including demand forecasting, project completion, and product quality (Table 1). Demand forecasting markets typically involved an outcome captured by a continuous variable (e.g., "How many Gmail users will there be by the end of Q2?"). An expert was asked to partition the continuum of possible outcomes into five equally likely ranges. In contrast, project completion and product quality OKRs were more likely to have binary outcomes (e.g., would a project be completed by the announced deadline), and these markets had two outcome securities. In addition to markets on OKRs, Google also ran markets on other business-related external events (e.g., will Apple launch a computer based on Intel's Power PC chip) and on fun topics that were designed to increase participation in the other markets.[78]

[78] Further detail on Google's prediction markets is available in the original version of this paper (Cowgill, Wolfers, and Zitzewitz, 2009) and in a Harvard Business School teaching case (Coles, Lakhani, and McAfee, 2007).

Ford Motor Company is a global automotive manufacturer based in Dearborn, Michigan, with operations and distribution on six continents and a financial services arm called Ford Motor Credit Company. Ford chose to focus its prediction markets on two topics of especially high importance: forecasting weekly sales volumes and predicting which car features would be popular with customers (as proxied in the interim by traditional market research, such as focus groups or surveys). Ford limited participation to employees with relevant expertise (in the Marketing and Product Development Divisions).

Sales forecasting is an important activity at an automaker, as it is essential for planning procurement and production so as to minimize parts and vehicle inventories. Ford has a long history of employing experts to forecast sales and other macroeconomic variables. Sales forecasting is also a common application for prediction markets: some of the Google OKRs involved future use of its products, and sales forecasts were the subject of markets at Hewlett-Packard (Chen and Plott, 2002) and Intel (Gillen, Plott, and Shum, 2013). Like H-P and Intel, Ford has an expert make official sales forecasts, with which we can compare the contemporaneous forecast of the market for accuracy. Unlike in the Google markets, in the Ford sales forecasting markets, a single security was traded with a payoff that was a linear function of the weekly sales for a particular model.

The features markets run by Ford were markets that sought to predict the success of a decision prospectively, which are sometimes called decision markets (Hanson, 2002). In a decision market, securities pay off based on an outcome variable, assuming the decision is undertaken. If the decision tracked by a security is not undertaken, then trades are cancelled. As a result, securities prices should reflect the expected value of the outcome variable, conditional on the decision being undertaken. Rather than defining the outcome as a feature's long-term success in the marketplace, Ford chose feedback from market research as a more immediate outcome measure. Its markets asked whether a series of potential car features (e.g., an in-car vacuum) would reach a threshold level of interest in market research, if that research were conducted.

Traditional market research is expensive to run, sample sizes are necessarily small, and so sampling errors can be meaningful. In contrast, opinions of employees may be cheaper to obtain, but employees are potentially biased, which is the reason non-employees are consulted in the first place. By asking employees to predict the results of the traditional market research, Ford sought to increase sample sizes while mitigating any biases in employee opinion. In a 2011 press release, Ford mentioned that it decided against including a Ford-branded bike carrier and an in-car vacuum in future models based on trading in its Features prediction market.[79] Ford also found the qualitative comments market par-

[79] See http://www.hpcwire.com/2011/02/22/ford_motor_company_turns_to_cloud-based_prediction_market_software/ (last accessed 6/30/2014).

ticipants made via the prediction market software to be of independent value. Ford cited employee education and engagement as additional benefits of running prediction markets.[80]

Unfortunately for research purposes, shortly after launching the features markets, Ford decided that results of its market research were too sensitive to be shared with its market participants, given the potential for imitation by competitors. As a result, it began settling markets based on the final trade price, rather than the market research outcomes. This turned the markets into "beauty contest" markets, in which security payoffs depend only on future market prices (Keynes, 1936). While an analysis of the predictive power of these markets would have been interesting, unfortunately this decision also meant the relevant market research outcomes were not recorded in our data, and subsequent attempts to obtain them were unsuccessful. As a result, we have reluctantly omitted them from the analysis.[81]

Firm X is a large, privately held, and profitable diversified basic materials and energy conglomerate headquartered in the Midwestern United States, but with global operations. It refines crude oil, transports oil and petroleum products, and manufactures products including chemicals, building materials, paper products, and synthetic fibers like spandex. Many of its businesses are very sensitive to the macro-economy and/or to commodity prices, both of which were quite volatile during the period their markets ran (March 2008 to present). Firm X decided to focus its prediction markets on macroeconomic and commodity prices that were relevant to its business. Some of these variables were already priced by existing futures markets (e.g., the future level of the Dow Industrials index or the West Texas Intermediate crude oil price) and some are the subject of macroeconomic forecasting (e.g., the unemployment rate and general price inflation), but many others were not (e.g., the Spandex price in China, the Kansas City Fed's Financial Stress Index). In addition, markets were run on policy and political outcomes of interest to Firm X, such as bailouts, health care reform, and the midterm and Presidential elections.

[80] Montgomery et. al. (2013) discusses these additional benefits in more detail.

[81] In an earlier version of this paper, we analyzed a single round of features markets that were run before this change was made. Those markets were poorly calibrated. Markets trading at high prices were roughly efficient, but those trading at low and intermediate prices displayed a very large optimism bias. Features with securities that traded below their initial price never achieved the threshold level of customer interest, and therefore were always expired at zero, and yet the market appeared to not anticipate this. Subsequent discussions with Ford revealed that these markets included features that were not shown to customers, and that these markets may have been unwound rather than expired at zero. Given the uncertainty about returns in these markets, we decided to omit on analysis of these markets from the revised paper, but include a graph documenting the poor calibration of the Features markets in the online appendix.

Firm X's markets were started by a Senior Manager in its strategic planning department, and participation was limited to a hand-selected group of employees with relevant expertise. While the number of participants in the Firm X markets was much smaller than at Google or Ford, 57 out of 58 invitees participated, and the average participant placed 220 trades (compared with 48 at Google and 10 at Ford).

Firm X's market creator had an additional motivation beyond obtaining forecasts. "People are overconfident in their predictions," he says. "They either say 'X will happen' or 'X won't happen.' They fail to think probabilistically, or confront their mistakes when they happen. The market therefore changes the way participants think, and I believe this not only improves our forecasts but has a positive spillover on everything else our team does." This stated goal is particularly interesting in light of our results, which suggest that markets are initially optimistic and overconfident (e.g., they display a bias away from a naive prior), but that these biases decline over time and that more experienced traders trade against them.[82]

Just under 60 percent of Firm X's markets predicted a continuous variable. About one-fifth of these markets divided the continuum of possible future outcomes into 3-10 bins as in Google's markets, while almost all of the other 80 percent specified a single "over/under" threshold. A very small number of markets (18 out of 1345) used the linear payoffs used by Fords Sales markets. For the remaining 40 percent of markets that predicted a discrete event (e.g., would President Obama be re-elected), there was a single security, which paid off if the specified event occurred.

2.2 Commonalities and Differences

Table 1 summarizes the types of markets run by the three companies, and provides examples of a few other companies we are aware of that have run related markets. All six types of markets we are aware of being run at other firms were run at our three firms. Google ran markets of all varieties, while Ford focused on sales forecasting and decision markets, and Firm X focused on external events. A few other firms have run many types of prediction markets (e.g., Eli Lilly, Best Buy), while others have run more focused experiments with one particular type of market.[83]

[82] Ironically, it was the markets at Google and Ford that displayed evidence of inefficiencies that disappeared over time; our analysis suggests that the Firm X markets were well-calibrated from the beginning.

[83] We base these statements on public comments made at conferences by firms, as well as on interviews. In the latter case, we do not identify specific firms (e.g., the reference to "other pharma") unless we have received permission to, and we omit some examples we are aware of for brevity. It is of course possible that firms have run markets we are unaware of.

Table 2 contrasts the scale and some key features of our three markets. One important difference was the structure of the securities in the markets. As discussed above, Google used multiple bins for continuous outcomes (e.g., demand) and two bins for discrete outcomes (e.g., deadlines). In contrast, Ford used securities with linear payoffs for the continuous outcomes in its Sales markets and single binary securities for the discrete outcomes in its Features markets. With a very small number of exceptions, Firm X used single binary securities for discrete outcomes and either bins or a single binary security combined with an "over/under" threshold for continuous outcomes. The choice between two bins and single binary securities for discrete outcomes can potentially affect market efficiency if some participants exhibit "short aversion" (i.e., prefer to take positions by buying rather than selling). With bins, choices of boundaries can affect efficiency if participants take cues from them, as the literature on partition dependence suggests some do (Fox and Clemen, 2005; Sonnenman, et. al., 2011). We will test whether pricing suggests bias towards buying, as well as whether there is a bias towards pricing each of N bins at 1/N.

Two other important differences were the market making mechanism and the incentives provided to participants, which we discuss in turn.

2.3 Market-making Mechanism

Google used an approach similar to the Iowa Electronic Markets (see, e.g., Forsythe, et. al., 1992), in which the range of possible future outcomes is divided into a set of mutually exclusive and completely exhaustive bins, and securities are offered for each. For continuous variable outcomes, such as future demand for a product, five bins were typically used, with the boundaries chosen by an expert to roughly equalize ex ante probability. For OKRs with discrete outcomes, like whether a deadline or quality target will be met, there are generally two outcomes, and no reason to expect the ex ante probability to be 0.5 (indeed, Google's official advice on forming OKRs is that they should be targets that will be met 65% of the time).

As on the Iowa Markets, participants can exchange a unit of artificial currency for a complete set of securities or vice versa. In markets with more than two outcomes, this approach does make shorting a security less convenient than taking a long position, since one must either first exchange currency for a complete set of securities and then sell the security, or else buy the securities linked to all other outcomes. On the other hand, any inconvenience cost of shorting should affect all securities in a market at least approximately equally, and biases to prices should be limited by the fact that other participants can simultaneously sell all outcomes if their bid prices sum to more than one. Google did

not have an automated market maker, but traders were observed placing such arbitrage trades (selling all possible outcomes when their bid prices summed to greater than one or, more rarely, buying when their ask prices summed to less than one).

Ford and Firm X used prediction market software developed by Inkling Markets.[84] Inkling's software uses an automated market maker that follows the logarithmic market scoring rule described in Hanson (2003). The market maker allows trading of infinitesimal amounts at zero transaction costs, and moves its price up or down in response to net buying or selling. The automated market maker ensures that traders can always place trades, which helps avoid frustration and is particularly important when participation is limited. In cases where securities are linked to a mutually exclusive and exhaustive set of outcomes, the automated market maker ensures that their prices always sum to one. The presence of the automated market maker also makes shorting or taking long positions equally convenient.

An issue with an automated market maker is that it must be set at an initial price, and market prices can therefore be biased towards this initial price, especially if participation is limited. Furthermore, if the initial price differs from a reasonable prior, then easy returns can be earned by being the first to trade. If relative performance (e.g., "bragging rights") is a source of motivation for trades, having performance depend too heavily on simply being the first to trade against an obviously incorrect price can be counterproductive. As a result, Inkling users take some care in setting initial prices, or in setting bin boundaries so that initial prices of 1/N are appropriate.

Thus the use of the Inkling mechanism could potentially reinforce potential biases toward pricing at 1/N discussed above. We will test whether prices at Ford and Firm X are biased towards their initial starting values, particularly early in markets' life, when compared with the markets at Google.

2.3.1 Incentives

Modest incentives for successful trading were provided at all three firms. Monetary incentives were largest at Google, although even these were quite modest. Google endowed traders with equal amounts of an artificial currency at the beginning of each quarter, and at the end of each quarter, this currency was converted linearly into raffle tickets for traders who placed at least one trade. The prize budget was $10,000 each quarter, or about 25-100 per active trader. The raffle approach creates the possibility that a poorly performing trader may win a prize through chance, but has the advantage of making incentives for traders

[84] http://www.inklingmarkets.com/

linear in artificial currency. Awarding a prize to the trader with the most currency would create convex incentives, which could make low-priced binary securities excessively attractive, potentially distorting prices.

Ford also used a lottery that created incentives that were linear in the currency used by the marketplace. For legal and regulatory reasons, it was not able to offer prizes to participants based outside the U.S., but we are told that these were a small share of participants in the markets we analyze.[85] Ford's incentives in North America were smaller than Google's, consisting of several $100 gift certificates.

Firm X did not offer monetary incentives for its traders, but publicized the most successful traders. The high participation rate of eligible Firm X traders suggests that the prediction markets were emphasized by management, and thus reputational incentives to perform should have been meaningful. If more attention was paid to the best performers than to the worst, the reputational incentives could have been convex in performance, encouraging risk taking. In particular, traders may have preferred the positively skewed payoffs of low-priced binary securities, potentially causing these securities to be mispriced.

Google also published league tables of the best performing traders, but any convexity may have been muted by the linear monetary incentives that were also provided. We therefore might expect low-priced binary securities to be more overpriced at Firm X than at Google. With smaller linear incentives for most of its participants, Ford might be expected to be an intermediate case between Google and Firm X.

3 Results

This section presents statistical tests in four subsections. The first subsection provides simple tests of the calibration of the three firms' markets. We test whether securities are priced at the expectation of their payoffs, conditional on price alone. The second subsection examines whether forecasts from prediction markets improve on contemporaneous expert forecasts. The third subsection expands our analysis of price efficiency to include tests for an optimism bias and for how pricing biases evolve over time. The final subsection examines how trader skill and experience are related to trading profits and to whether one trades with or against the aforementioned biases. This subsection also uses data on job and project assignments at Google to examine how insiders trade in markets.

[85] We have unfortunately been unable to obtain a precise percentage.

3.1 Calibration

In this subsection, we test whether the markets at Google, Ford, and Firm X make efficient forecasts, in the sense that they do not make forecasting errors that are predictable at the time of the forecast. This is equivalent to asking whether the markets yield predictable returns. In particular, if a market is asked to forecast Y (which could be a binary variable indicating whether an event occurred, or a continuous variable indicating, e.g., the sales of a car model), then an efficient forecast at time t will be E(Y — Ht), where Ht is the set of information known publicly at time t. If prediction market prices are efficient forecasts, then the price at time t is equal to this expectation, Pt = E(Y — Ht), and expected future returns are zero, E(Y Pt — Ht) = 0.

We focus our tests on variables that are known at time t and that our above review of the theory literature suggests may be correlated with mispricings. In this subsection, we begin by asking whether future prediction market returns are correlated with the current price level or the difference between the current price and a naive prior (either the market makers initial price or 1/N, where N is the number of mutually exclusive outcomes).

Figures 1 and 2 graph the future value of securities, conditional on current price for binary securities at Google and Firm X, respectively.[86] The prices and future values of binary securities range from 0 to 1, and trades are divided into 20 bins (0-0.05, 0.05-0.1, etc.) based on their trade price. The average trade price and ultimate payoffs for each bin are graphed on the x and y-axes, respectively. A 95% confidence interval for the average payoff is also graphed, along with a 45-degree line for comparison. The standard errors used to construct the confidence interval are heteroskedasticity-robust and allow for clustering within market.[87] Observations are weighted by time-to-next trade, which weights trades according to the amount of time that they persist as the last trade, and thus according to the likelihood they would be taken to be the current market forecast by a user consulting

[86] All securities in the Google markets are binary and none of the contracts in the Ford sales markets are (they had payoffs linear in vehicle sales). Almost all Firm X markets are binary the exceptions were a small number of markets with linear payoffs in commodity prices. These markets accounted for just under 1 percent of markets and trades, and they are excluded from Figure 2.

[87] Allowing for clustering at the market level allows for arbitrary correlations within the returns-to-expiry for trades within the same market: in this case for the fact that returns within securities will be positively correlated and returns across securities within markets will be negatively correlated. For Google and Firm X, we also cluster on calendar month as a second dimension in the regression tables presented below, using the code provided by Petersen (2009). This yields standard errors that are very similar to those that cluster only on market. The Ford prediction markets were short-lived enough that we do not have a sufficient number of calendar months for clustering to be valid, so we instead use one-dimensional clustering on markets (which, in the Sales markets, is also equivalent to clustering on time periods). With only six models in the Ford markets, clustering on model as well would not yield asymptotically valid standard errors.

the market at a random time.[88]

Google and Firm Xs markets appear approximately well-calibrated. Both markets exhibit an apparent underpricing of securities with prices below 0.2, and an overpricing for securities above that price level, but this is slight, especially for Firm X. For Google, the price level below which we observe overpricing differs in two and five-outcome markets. Figures 3A and 3B plot percentage point returns to expiry (i.e., the difference between payoff and price) against price for Google's two and five-outcome markets, respectively.[89] In both sets of markets, securities are underpriced when priced below 1/N and overpriced when priced above this level, implying a bias in prices away from 1/N.

Figure 4 examines the calibration of Fords sales markets. Given that these are linear markets and that they track sales for different models with differing overall sales levels, we scale prices and payoffs using a models past sales. In order to ensure that we do not condition our analysis on information that market participants would not have observed, we use 3-week lagged sales. The x-axis plots the log difference between the sales forecast by a trade and lagged sales, and the y-axis plots the average difference between actual log weekly sales and lagged sales. The graph suggests that in contrast to the Features markets, the Sales markets are generally well-calibrated, albeit perhaps with a mild optimistic bias.

Table 3 presents regressions that test the calibration of the three firms markets. For each market, we begin with regressions of payoff on price, where the unit of observation is a trade. If prices are efficient forecasts, then E(Yt — Pt) = Pt, and a regression of Yt on Pt should yield a slope of one and a constant of zero. The second regression reported for each market is a regression of percentage point returns to expiry (Yt Pt) on Pt. In these regressions, efficient forecasting would be consistent with a slope of zero and a constant of zero. For obvious reasons, the slope in the first regression is simply one plus the slope in the second regression.

The results imply that we cannot reject the null hypothesis of efficient forecasting for the Firm X and Ford Sales markets. For the Google markets, we can reject this null, but we still conclude that prices are informative, as they are strongly positively correlated with outcomes. For Google, the relationship is slightly less than one-for-one, which implies that high-priced contracts are overpriced and low-priced contracts are underpriced, consistent

[88]Note that weighting in this manner does not produce a look-ahead bias from a forecasting perspective. Equal weighting trades produces very similar, albeit slightly noisier, results.

[89]Following other work on binary prediction markets (see, e.g., Tetlock, 2008), we use percentage point returns to expiry (i.e., payoff minus price) rather than scaling returns by their price [i.e. (payoff price)/price]. We do so for two reasons: 1) we are primarily interested in returns as a measure of forecasting performance, rather than financial profit opportunities, and therefore there is no reason to be more interested in a given sized percentage point profit opportunity when the price is low, 2) scaling by price causes the returns of very low priced securities to dominate the results, and makes the outcome variable more heteroskedastic.

with Figure 1.

Table 3 also reports regressions for Google and Firm X that test whether returns (i.e., forecast errors) are better predicted by price or by the difference between price and 1/N. Whereas the Firm X markets exhibit no predictability with respect to either variable, returns in the Google markets are better predicted by (price 1/N) than by price, consistent with Figures 3A and 3B. We report separate regressions for 2 and 5-outcome markets, which collectively account for 92 percent of trades in Googles markets and 65 percent in Firm Xs markets. These regressions suggest predictability in both types of Googles markets, again consistent with the Figures, but neither subset of Firm X's markets. For Ford, we substitute the most recent sales figure reported prior the market commencing as our naive prior, and likewise test whether returns a better priced by (price - prior) sales than by price. We do not find statistically significant evidence that either variable predicts returns for Ford.

Taken together, the results from the Figures and Table 3 suggest that all markets have prices that are positively correlated with outcomes, and the Firm X, Ford Sales, and Google markets are reasonably well-calibrated. While the Firm X and Ford Sales markets exhibit no evidence of return predictability, the Google markets display a bias in pricing away from a naive prior of 1/N. This bias is the opposite of the longshot bias predicted by the Ali (1977) and Manski (2006) models and is also inconsistent with participants taking cues from security boundaries as in the partition dependence literature. It is instead consistent with investors collectively under reacting to the information used in designing the boundaries or overreacting to other information, such as new information or their own prior beliefs (as in Ottaviani and Sorensen, 2014).

3.2 Markets versus Experts

Given that firms run prediction markets at least partly to obtain predictions, a natural next question is whether the predictions from markets outperform alternatives, including forecasts by expert forecasters or managers. We compare markets predictions with three types of alternative forecasts. The first is a formal forecast from a team of expert forecasters. Ford forecasts weekly auto sales for different models, and for the six models covered by prediction markets, we can compare the experts' forecast with the prediction market forecast from immediately before the forecast was issued.[90]

[90]The expert forecasts were issued 11 days before the week in question began. The six forecasted models were the Escape, F-150, Focus, Fusion, Super Duty, and Lincoln (all models). The official sales forecasts are closely held at Ford and were not available to the vast majority of predict market participants.

A second type of forecast we compare with are percentile forecasts derived from bin boundaries used in constructing the prediction market securities. As mentioned above, in order to avoid minimize pricing biases from either partition dependence effects or the initialization of market maker prices at 1/N, both Google and Firm X sought expert help in choosing bin boundaries to equalize ex ante probabilities. For example, at Google, the prediction market organizers would ask the Product Manager for the relevant product (e.g., the Gmail Product Manager for markets on new Gmail users) for assistance in creating the bins. These experts were encouraged to use whatever sources they desired to set these boundaries, and they often consulted historical data or made statistical forecasts. The bin boundaries they chose can be interpreted as specific-percentile forecasts, and it is straightforward to obtain an approximate median forecast from these boundaries.[91]

A third, related, type of forecast can be obtained from over/under markets that were run by Firm X on continuous variables. In these markets, a single security was traded that paid off if a macroeconomic variable exceeded a threshold, and as above that was chosen to create a 50 percent ex ante probability. The threshold can therefore be interpreted as a median forecast. About half of the binary markets in our sample used a prior-period value as the threshold (e.g., will housing starts be up from last month?). We analyze only over/under markets where this approach was not used, in order to focus on instances where an over/under value was actively selected.

We compare forecasts that are as close to contemporaneous as possible. For Ford, prediction markets were begun several days before the expert forecast was made, so we were able to compare the expert forecast with the prediction market forecast immediately prior to the expert's forecast. For Google and Firm X, the expert forecast was used to design the securities, and so it was necessarily made a few days before the prediction market was opened. In order to limit the timing difference between the expert and prediction markets forecasts, we use prediction markets forecasts from only the first day that a market was open. The prediction market traders may have had access to a few days of information that was not yet available to the expert at Google and Firm X, but this should not have been the case for Ford.

Table 4 presents the results of these comparisons. In each column we report the results of horserace regressions (Fair and Shiller, 1989) of the security payoffs on the prediction market and expert forecasts. We also report the ratio of the prediction market and expert mean squared errors, and the p-value from a f-test for the equivalence of the two variances. In all four cases, the prediction market forecast has a lower mean squared error

[91]For example, when there are an even number of bins, the boundary between the two middle bins is a median forecast. When there are an odd number of bins, the midpoint of the middle bin is an approximate median forecast.

and receives a higher weight in the horserace regression.

The expert forecasts we study obviously differ in their formality. Ford has a long history of producing forecasts of weekly auto sales, which are clearly of high importance to planning procurement and production so as to minimize part and vehicle inventories. While the individuals setting the bin boundaries at Google and Firm X were chosen to be the most knowledgeable at the company, it is possible that less effort was put into their forecasts than was exerted at Ford. Nevertheless, it is interesting to note that the mean squared error improvement achieved by the prediction market at Ford is among the largest.[92]

3.3 Prediction market pricing biases

This subsection expands our earlier analysis of prediction market efficiency. In particular, we test whether forecasting errors can be predicted by a broader set of variables than price alone.

In Tables 5 and 6, we test for an optimism bias by adding a variable that captures whether a security is linked to an outcome that we judge would be good for the company. In the Ford Sales markets, all securities are structured so that buying involves an expression of optimism (i.e., predicting high sales) and thus it is impossible to distinguish between optimism and a preference for taking long rather than short positions. In Google and Firm Xs markets, however, securities were available that were linked to both positive and negative outcomes, and so we can separate these two effects. In these markets, we code the most optimistic outcome as +1, the least optimism as -1, and place intermediate outcomes at uniform intervals along this scale (e.g., in 5-outcome markets, the outcomes are given optimism -1, -0.5, 0, 0.5, and 1; in 2-outcome markets, they are given scores of -1 and 1). We limit the sample in Tables 5 and 6 to markets for which we can identify the outcome that would be good for the firm without making a difficult judgment call.

In Table 5, we find negative future returns for securities tied to optimistic outcomes in Googles markets. The evidence of the bias away from 1/N persists when controlling for optimism. There is also evidence of small biases towards purchasing rather than selling securities (reflected in the negative average returns to expiry) and against purchasing securities tied to the most extreme outcomes (reflected in the positive returns for these securities). We do not see evidence of any of these biases in Firm Xs markets, however. Likewise, the Ford Sales markets do not exhibit evidence of a combined optimism and short aversion bias (based on the near-zero constant in Table 3, Panel C, Column 4).

[92]The p-value for the test for the statistical significance of the improvement is largest at Ford, at 0.104, but this is related to the much smaller sample size at Ford.

In Table 6, we test for optimism separately for markets on different subjects. As in Table 5, we limit the sample to markets for which identifying the outcomes that are better for the firm can be done without difficult judgments calls. For both Google and Firm X, all Fun markets are excluded. For Google, high demand for products, timely completion of projects, and high product quality are all regarded as good for Google. Markets on external news were assigned optimism scores by a member of the company (who was not Cowgill) in cases where the assignments were regarded as uncontroversial (if there was any doubt about which outcome was better for Google, we did not assign optimism scores for that market)).

Firm X is a largely U.S.-based basic materials and energy producer. We code macroeconomic outcomes associated with a strong economy (e.g., high GDP growth, low unemployment, high employment, high industrial production, and high equity prices) as good for the firm. Most markets Firm X ran on commodity prices were on commodities it produced or assisted in the production of, so for these markets we code high commodity prices as good for Firm X. We exclude markets if we are uncertain about whether Firm X was a net buyer or seller of the commodity. Given the macroeconomic situation during the time period studied (2008-13), we code increases in inflation as good for Firm X. During the 2011 European banking and sovereign debt crisis, Firm X ran markets on future interest spreads and write downs for investors, and we regard high spreads and write downs as negative for the global macro economy and thus for Firm X. For markets on exchange rates between the US dollar and another currency, we code a weak dollar as good for Firm X, unless Firm X also produced in the country in question, in which case we omit that market from the sample. For markets on policy and politics, we code pro-business outcomes as good for Firm X, such as electoral victories by U.S. Republicans or UK Tories, or the passage of policies backed primarily by these parties.[93] Where applicable, our optimism codings are consistent with public statements by the firms executives. We suspect that most readers will regard all of these judgments as uncontroversial, however, the impact of reversing or omitting any of them can be ascertained from the disaggregated results in Table 6.

In Table 6, we find that the optimistic bias is largest for markets on project completion. There are several reasons to expect the bias to be largest in these markets. First, these markets are on outcomes that are most under Google employees' control, and thus perhaps the most influenced by overconfidence about one's own or one's colleagues' ability. Second, strategic concerns for biased trading by insiders may be larger for these markets, given that outcomes are more under employees' control. Third, information about project

[93]This is consistent with the fact that stock market prices for basic materials and energy firms increase on average when Republicans win close elections (see, e.g., Snowberg, Wolfers, and Zitzewitz, 2007a and 2007b and Zitzewitz, 2014).

completion is presumably less dispersed throughout the organization than information about demand or external news, and discouraging entry by arbitrageurs and making the potentially biased views of project insiders more influential.

Unlike optimism, the degree of bias away from 1/N does not vary statistically significantly across the categories of Google's markets.[94] In contrast to Google, Firm Xs markets exhibit almost no evidence of bias. This is not simply due to imprecision of the estimates, as the Firm X sample is larger in terms of markets and securities, and coefficients of the magnitude found at Google can be rejected for the (Price 1/N) and optimism variables.

As discussed above, the optimism in Google's markets could arise for either strategic or behavioral reasons. To help distinguish among the two, we conduct tests for company-wide mood swings in the optimism of prediction market pricing. In Cowgill and Zitzewitz (2013), we find daily frequency correlations between the company stock price and job satisfaction, physical output, hours worked, hiring decisions, and the evaluation of candidates and ideas. There is no persistence in these correlations (i.e., the stock price change from last week is not correlated with the outcome variables) which is inconsistent with standard explanations, such as an increase in employee wealth affecting labor supply decisions, or good news for a company affecting future labor demand and thus hiring. Instead we conclude that companywide mood swings are the likely explanation.

Table 7 presents tests for mood-swing effects on the size of the optimism bias at Google. The regressions repeat the specification in Table 5, Panel A, Column 4, with the optimism variable interacted with Google stock returns on days t+1, t, t-1, and t-2. In a variety of different specifications, we find that a 2% increase in Googles stock price (roughly a one standard deviation change) is associated with prediction market prices for securities tracking optimism outcomes being priced 3-4 percentage points higher, relative to their pricing on an average day. As in Cowgill and Zitzewitz (2013), these effects are quite temporary, as there is no association between the prediction market prices and day t-2 returns, as we would expect if the aforementioned relationship was driven by good news leading to both higher stock and prediction market prices.

We conclude this subsection by examining how pricing biases evolve over time: over the life of an individual market and over the life of the prediction market experiment as a whole. As discussed above, the Firm X and Ford prediction markets use an automated market maker that is initialized at a prior, and we might expect prices to be biased to-

[94]The degree of optimism is statistically significantly different in the completion and external news categories (p ¡ 0.001 in both cases), but biases away from the prior are not statistically significantly different from one another (p = 0.870). P-values are calculated using versions of the regression in Table 6, Panel A, Column 1 that allow for add an interaction between the bias variable (i.e., price - 1/N or optimism) and an indicator variable for the market category.

wards that initialization value, at least early in the life of the market. To investigate this possibility, we number the trades in each security sequentially and then split the sample according to this trade number (Table 8).[95] We find no evidence that prices are biased towards the price, even very earlier in a market's life.[96] The large bias away from the prior in Ford Sales markets after trade number 50 turns out to be driven by a single, very inaccurate, market for one model in the first week; if that market is excluded, the coefficient on Price Prior is consistent with other subsamples.

The Google markets did not use an automated market maker, and thus they have less reason to be biased towards the prior value early in their life. Indeed, the results in Table 8 imply that they are actually biased away from the prior early in their life and that this bias abates with more trading history. As discussed above, the bias away from the 1/N prior suggests that traders are either overweighting their own prior beliefs or information that arrives after the market begins. The fact that the bias away from the market prior declines over the life of the market is more consistent with the former possibility. In contrast, the optimistic bias in Google's markets is small early in a markets life, and grows over time. This is consistent with market participants overreacting to new positive information and underreacting to new negative information.[97]

Finally, Table 9 presents tests of how the aforementioned (Price Prior) and optimism biases evolved over our sample. Regressions from Tables 3 and 4 are modified by the inclusion of a time trend (which is scaled to equal 0 at the beginning of the sample and 1 at the end) and interactions of the time trend with the bias variables. The results suggest that biases away from the prior in the Google and Ford markets are large at the beginning of the sample and essentially disappear by the end of the sample. The same appears to be true of the optimism bias in Googles markets. Firm X's markets again appear efficient, albeit with weak evidence ($p = 0.09$) of a small optimism bias at the beginning of the sample that disappears by the sample's end.

[95] We take this approach to splitting the trading history of markets because the trade number is a variable that will be known at the time of the trade, while the whether a trade is in a given decile of a particular market's life would not be known.

[96] In an earlier version of the paper, we cut the "Trades 1-10" sample even finer, finding no evidence of biases toward 1/N, even in the prices of the first two trades in each market.

[97] Unfortunately, we lack a direct measure of new information arrival for most of Google's markets. To further investigate over and under reaction, we ran tests for price momentum or reversals, but found that results were not robust to small changes in time horizons.

3.4 Individual trader characteristics and market efficiency

This subsection analyzes how traders' characteristics are which traders contribute to the biases discussed above, which traders trade against these biases, and which traders earn positive returns. For all three firms we have trader identifiers, and so we can construct variables that describe a traders past history. For Google we also have data on traders job and project assignments, and so we also construct variables that capture a traders relationship with the subject of the market being traded.

In order to understand which traders contribute to and trade against pricing biases, we need analyze the relationship between the nature of a position being taken (e.g., its optimism) and the characteristics of the trader. We begin by analyzing all three companies, and thus focus on traders past experience and past success. Prior to each trade, we calculate for each trader the number of prior trades that each trader has participated in and their average past return to expiry on all trades in contracts that have settled by that time. In order to be included in the sample, a trader must have at least one past trade in a contract that has settled.

In the Google data, participants trade against each other, and thus every trade has a buyer and a seller. For Google, we structure the data so that each trade appears in the data twice (i.e., as a buy by one trader and as a sell by another). The characteristics of the security traded are first multiplied by the direction of that side of the trade (+1 if a buy, -1 if a sell) and then regressed on trade fixed effects and the trader characteristics for that trade*side. This yields coefficients that are identical to what we would obtain if we regress the securitys characteristics on the difference in the characteristics of the buyer and seller, but has the advantage of facilitating the adjustment of standard errors for clustering within traders as well as within markets.[98] The coefficients in the regression tell us whether the traders with greater experience or better past returns is systematically on the purchasing side, on the optimistic side, on the side that buys securities priced above 1/N or sells those priced below, and on the side that ultimately earns positive returns.

In the Ford and Firm X data, where participants trade against an automated market maker, we multiply the security characteristics by the direction of the trade (i.e., +1 if the participant is buying, -1 if selling) and then regress these on trader characteristics. Since we have one observation per trade rather than two, trade fixed effects are not included.[99] In these regressions, the coefficients tells us whether traders with greater experience or

[98] Note that clustering by market also adjusts standard errors for the inclusion of two observations per trade, as clustering allows for any correlation of errors within cluster groups.

[99] We include time period fixed effects (for weeks for Ford and months for Firm X) to control for changes in trading behavior over time, although doing so has limited impact on the results.

better past returns are more likely to buy than sell, are more likely to trade in an optimistic direction, are more likely to buy when prices are above 1/N and sell when they are below, and are more likely to buy securities that ultimately have positive returns to expiry.

Table 10 presents the results of these tests. In Panel A, we find that Google traders with high past returns trade in a pessimistic direction, are more likely to sell than buy, and trade against securities that are priced above 1/N. All three correlations are consistent with what the previous section found to be profitable, and consistent with this, we find that traders with high past returns earn high future returns. We also find that more experienced traders are more likely to sell and to trade against securities that are priced above 1/N, again in both cases consistent with what would be profitable. Thus we can conclude that less experienced traders and traders with less past success trade in a direction that would contribute to the biases discussed above.

In the Ford markets, we also find that traders with more past experience and more past success are more likely to sell than buy (which means they are also trading pessimistically), and both types of traders are more likely to sell when price is above its initial value (Panel B). The results presented in Sections 2.1 and 2.3 suggest that trading in this direction should be profitable, and indeed we find a positive and significant relationship between future returns and both past performance and past experience.

Given that the Firm X markets did not display pricing biases, there is less reason to expect proxies for trader experience or skill to be correlated with trading in a particular direction. Indeed, in Panel C, we see much less evidence of such correlations. We do see a positive correlation between past and future returns, consistent with traders displaying persistent skill.

In a previous version of the paper, we analyzed how continued participation by a particular trader was related to past performance and activity. At all three firms, continued participation more likely for traders with higher past returns and those who were more active in the prior period (see Table A1 in the Online Appendix). The reduction of pricing biases over time at Google and Ford are consistent with the fact that the more skillful and experienced traders trade against these biases, that traders gain experience over time, and that the most engaged and skillful traders are more likely to continue to participate.

Finally, we analyze the relationship between traders' job assignments and their prediction market trading, using data that are only available to us for Google. Table 11 presents regressions with the same structure as in Table 10, Panel A. We find that optimistic trades are made disproportionately by traders who are staffed on the project in question and by friends of those insiders (as indicated by either party on a social network survey). Insiders are also more likely to buy securities and to buy when securities are trading above

1/N. Consistent with this, they earn lower returns. Programmers and employees based in Mountain View and New York (Googles second largest office at the time of the study), who we might to be more knowledgeable, tend to trade against biases and earn higher returns. The results are consistent with those with the most knowledge of a markets subject trading in an unprofitable (and potentially strategically biased) way, but with other knowledgeable employees trading in the opposite direction, pushing prices back to their efficient level.

It is also interesting that newly hired employees trade more optimistically. It is worth noting that during this time period, the vast majority of new Google hires were hired directly from degree programs, and thus were inexperienced both in working at Google and in working in general. Therefore it is possible that their optimism reflected an initial miscalibration about the extent to which demand forecasts and deadlines are stretch targets rather than unbiased forecasts. Consistent with this, we find in unreported results that the correlation between hire date and optimism is strongest for markets on demand forecasts and on whether deadlines will be met.

4 Discussion

While much of our analysis above deals with inefficiencies, our results about corporate prediction markets are largely encouraging. First, we find that forecasts from predictions markets outperform other forecasts available to management, including, in the case of Ford, sales forecasts that are taken extremely seriously.[100] Second, we find that prediction markets get better with age. In both the Google and Ford Sales markets, initial pricing biases disappeared as our sample progressed. This is consistent with the fact that we find more experienced traders trading against pricing biases and earning high returns, and with the fact that traders who appear unskilled stop participating. It is also consistent with our best-calibrated prediction markets being the markets at Firm X. The Firm-X markets ran for almost 5 years and the average participant made over 200 trades.

Regarding the inefficiencies, some results match well with the prior literature, while others are more puzzling. Our finding of an optimistic bias in some markets is consistent with prior work on the role of optimism in organizations. At Google, the optimistic bias is

[100] We cannot distinguish whether it is the prediction market mechanism per se that leads to the better predictive performance, or simply the involvement of more people. It is possible that an averaging of forecasts from multiple experts, or a Delphi method approach to aggregating information from several forecasters, would have also outperformed a single expert (or in Ford's case, a forecasting group). See Graefe and Armstrong (2011) for a laboratory experiment that compares the predictive performance of other group forecasting methods, such as Delphi.

strongest for markets on project completion. Insiders and their friends trade optimistically at Google, potentially for strategic reasons, but also potentially due to overconfidence in ones own and teammates ability. The fact that the optimistic bias exhibits "mood swings" (i.e., that is it correlated with daily stock returns) is more consistent with optimism having at least a partly behavioral source. The fact that newly hired employees are the most optimistic is consistent with employees arriving at Google initially miscalibrated and then learning. The fact the optimistic bias diminishes over our 2005-7 sample period is also consistent with initial miscalibration and learning. Taken together, the evidence suggests that strategic biases, overconfidence, behavior biases, and inexperience (i.e., beginning a career with systematically erroneous priors) all play a role in the optimistic bias.

The bias in pricing away from nave priors in Google and Fords markets is less consistent with prior literature. Most of the extant literature, such as the Ali (1977) and Manski (2006) models, the partition dependence literature, and the work on probability misperceptions (Kahneman and Tversky, 1979), led us to expect a bias in the other direction. We also expected the Inkling market-making mechanism to impart a bias towards the prior, at least early in the life of a market, and likewise the potential convexity of reputational incentives should have made low priced securities more attractive, creating a bias in the opposite direction. The fact that the bias away from the prior was strongest at Google (which had the most linear incentives and did not use an automated market maker) was consistent with these expectations, but the overall sign of the bias was not. The pricing bias we did find (at Google and in Fords Sales markets) is consistent with an overreaction to their own priors or to new information or with participants' underappreciating the effort that was put into security design (i.e., insufficient partition dependence). While we still find the direction of the bias puzzling, it did diminish over time, consistent with participants becoming better calibrated. By the end of the sample, there was no evidence of pricing inefficiencies in any of the Google, Ford, and Firm X markets. We are limited to analyzing the markets of firms who shared data with us, and the decision to share data may have be related to the success of the prediction markets. Nevertheless, we regard the evidence on the efficiency of corporate prediction markets as largely encouraging.

Producing efficient forecasts that improved upon the available alternatives was only one of the goals the companies had for their markets. Googles management sought to communicate the importance of its OKRs. The anecdote described above, where a senior manager admitted to having been embarrassed by prediction market trading into a redoubling of efforts, provides at least one example of this working.[101] We are aware of contrasting

[101] We originally hoped to produce more systematic evidence on this point by randomizing which OKRs were covered by markets, in order to test whether the existence of a market had a causal effect on a projects outcomes. Unfortunately, power calculations revealed that given the number of OKRs at Google for which it was feasible to run markets, the causal effect would have to be implausibly large to be detectable. If this

anecdotes from other companies though. For example, we are aware of four cases at different companies (outside those in our sample) in which internal prediction markets were shut down or limited at the request of senior management after they forecast problems with projects. One of these projects became a high-profile debacle that we believe most readers would be aware of (but which unfortunately we cannot name).

Among our three sample firms, only Firm X's markets are still being run, despite relatively strong predictive performance at all three firms. The shutdown of Ford's sales markets is especially puzzling, as a 25 percent reduction the mean squared error of a sales forecast is presumably of significant value to an automaker. Our contacts at Ford tell us that budgets for experiments like prediction markets were limited by the still recessionary economic environment in 2010. Confidentiality concerns may have limited the usefulness of the Features markets, and given that these markets accounted for the majority of trades, they may have overshadowed the more successful markets on Sales. It is also possible that accurate sales forecasting is slightly less important during periods of overcapacity, like 2010.[102] Nevertheless, we still regard this decision as puzzling.

In Google's case, its prediction markets were begun as a "20 percent project" (Google allows it's engineering to spend up to 20 percent of their time working on a new project of their choice) in 2005. All members of the project team left full-time employment at Google around 2008-10, and so continuing the project would have either required recruiting new 20-percent engineers or assigning engineers to work on the project as their "80-percent" assignment. Engineers tend to prefer working on 20-percent projects of their own creation, and the bar for promotion of a 20-percent project is high.[103] One possible view of the non-continuation of Google's prediction markets is that opportunity cost of its engineers' time is high, and the "20 percent time" system intentionally sacrifices moderately successful projects to maximize the number of major successes.

An alternative view is that Google lacked the high-value application for prediction markets that Ford arguably had in sales forecasting. Forecasting demand for Google's products (such as Gmail) is probably less important than forecasting car sales, as share of marginal costs are presumably a much lower share of total costs in Google's case, and acquiring processing and storage capacity in response to anticipated demand is a much

was true at Google, which is among the largest corporate prediction markets run to date, it is likely to be an issue in many other settings.

[102]Ford sales in 2010 had rebounded about 20 percent from the low in 2009, but were still well below pre-2008 levels. See, e.g., Ford Motor Company, 2011 Annual Report, p. 180 (available at http://corporate.ford.com/doc/2011%20Ford%20Motor%20Company%20AR_LR.pdf, last accessed June 27, 2014).

[103]Past 20-percent projects that have become "80 pecent time" products include Gmail, Google News, Google Talk, suggested completions of Google queries, and AdSense (a second advertising system for Gmail and blogs that now accounts for 20 percent of its revenue). See, e.g., Tate (2013).

more reversible decision than building a specific model of car, at least assuming that processing and storage capacity has many alternative applications. If running markets on project completion increased the likelihood that projects would be completed on time, this would presumably of much greater value to Google. Unfortunately, as mentioned above, Google's markets lacked the sample size needed to run an experiment designed to test for these effects. Furthermore, it is not obvious a priori that the effects would be positive. As discussed above, running markets on OKRs draws attention to them, presumably intensifying reputational incentives to achieve them. At the same time, if a project's participants hold initially optimistic views of the likelihood that it will be completed on time, debiasing these views may not be in the company's interest.

Decisions about the adoption and continuation of corporate prediction markets are typically made by agents for organizations populated by other agents. Decisions about the adoption of corporate prediction markets may therefore depend on factors other than their utility in aggregating information. First, if agents in organizations earn rents from asymmetric information, adopting technologies that increase transparency may not be in their interests. A prediction market may provide an ex ante measure of a key projects expected quality, where otherwise only ex post measures (e.g., market acceptance) would have been available. Agents, including CEOs, may prefer noisier measures of performance, especially if performance is expected to be disappointing.

Alternatively, senior managers may have legitimate concerns about organizational side effects. Aggregating information about the success or failure of a key initiative helps informs management, but also informs other members of the organization. Informing these other members may have side effects, such as potentially adverse effects on effort levels, the leakage of information to competitors, or the facilitation of insider trading. The third concern limited the OKRs eligible for markets at Google, and the second concern constrained the design of Ford's markets on Features. Informing an organization about a coming failure may be more damaging, in terms of morale, effort reduction, and employee turnover, than informing about coming successes is beneficial. The may be particularly true if employees have optimistic biases that benefit employers and that information sharing will reduce in expectation.

Our initial motivation for our analysis was that there were several plausible reasons to expect that prediction markets would not work well in corporate settings. Compared with public prediction markets, corporate markets are thinner, involve traders with potential biases, and have less potential for entry by arbitrageurs who reduce pricing biases. Despite this, the corporate prediction markets we study performed quite well. This, however, leaves us with two new open questions. First, why are corporate prediction markets not more popular, including at firms that have already experimented with them? And

second, does this lack of popularity itself reflect agency problems? Would firms' owners benefit from insisting on their adoption?

TABLE 1
Summary of corporate prediction markets at Google, Ford, Firm X and selected other companies

Market Topic	Example	Google	Ford	Firm X	Other companies running similar markets
Company performance					
Demand forecasting	Ford F-150 sales next week	X	X		Arcelor Mittal, Best Buy, Chrysler, Eli Lilly, HP, Intel, Nokia
Project completion	Will chat be launched within Gmail by end of quarter?	X			Best Buy, Electronic Arts, Eli Lilly, Microsoft, Nokia, Siemens
Product quality	Google Talk sound quality rating	X			Electronic Arts, Eli Lilly, other pharma
External events	Spandex Price in China	X		X	Eli Lilly, other pharma
Decision markets	If feature X is offered, what will demand be?	X	X		Best Buy, GE, Motorola, pharma, Qualcomm, Rite Solutions, Starwood
Fun	Will the Interns win at the company picnic?	X		X	Electronic Arts

Incentive Type	Example	Google	Ford	Firm X	Other companies running similar markets
Monetary Prizes	$1 cash awards, credit in company store	X			Best Buy, Microsoft, Misys
Non-Monetary Prizes	T-shirts, plaques	X	X		Microsoft, Misys, other pharma
Reputational Incentives Only	Leaderboard		X	X	Boing, J&J, Microsoft

Market Mechanism	Example	Google	Ford	Firm X	Other companies running similar markets
Decentralized					
Continuous Double Auction		X			Hewlett Packard, Nokia, Siemens
Centralized					
Market Maker (following Hanson (2003) approach)			X	X	Chevron, CNBC, Electronic Arts, GE, General Mills,Lockheed Martin, Microsoft, Missile Defense Agency, Misys, MITRE Corp, Motorola, NASA, Nucor, Overstock.com, PayPal, Proctor and Gamble, Qualcomm, SanDisk, T-Mobile
Other market maker					Boeing, Electronic Arts, Genentech, Hallmark, J&J, Overstock, Sony, WD-40

Approach to Beauty Contest Markets	Example	Google	Ford	Firm X	Other companies running similar markets
Avoided	What will be the price of oil in 2020? Trades resolved according to price of oil in 2020	X		X	Best Buy, Electronic Arts, Google, Microsoft
Included At Least Some	What will be the price of oil in 2020? Trades resolved according to the consensus in the prediction market in July 2012.		X		GE, Motorola, Rite-Solutions

Information about prediction markets run by firms outside of our sample come from public comments by firms and interviews. In some cases, the firm asked not to be identified, or provided only partial information. We omit some examples we are aware of for brevity. It is of course possible that firms have run markets we are unaware of. Note that some companies are listed twice within a section in cases where they changed approaches.

TABLE 2
Summary statistics

	Google	Ford	Firm X
Industry	Software/Internet	Automobile	Basic materials
Ownership	Public (Ticker: GOOG)	Public (Ticker: F)	Private
Sample begins	April 2005	May 2010	March 2008
Sample ends	September 2007	December 2010	January 2013
Markets (questions)	270	101	1,345
Securities (answers)	1,116	17	4,278
Trades	70,706	3,262	12,655
Unique traders	1,465	294	57
Market mechanism	IEM-style CDA	LMSR	LMSR
Software	Internally developed	Inkling	Inkling
Style of market			
One continuous outcome (e.g., how many F-150s sold?)		100%	1.3%
One binary outcome (e.g., Project X done by Sep 30?)			59%
Two outcomes (e.g., Yes and No securities)	29%		0.7%
3+ outcomes (e.g., bins)	71%		39%
Topic of market			
Demand forecasting	20%	100%	
Project completion	15%		
Product quality	10%		
External news	19%		96%
Decision	2%		
Fun	33%		4%
Share for which optimism can be signed	58%	100%	71%

Notes: IEM-style CDA = continuous double auction with separate securities for each outcome (Forsythe, et. al., 1992)
LMSR = Logarithmic Market Scoring Rule (Hanson, 2003)

TABLE 3
Calibration tests

Panel A. Google

	(1)	(2)	(3)	(4)	(5)	(6)
Markets included	All	All	All	All	2-outcome	5-outcome
Dependent variable	Payoff	Payoff - Price	Payoff - Price	Payoff - Price	Payoff - Price	Payoff - Price
Price	0.812***	-0.188**	0.006			
	(0.083)	(0.083)	(0.024)			
(Price - 1/N)			-0.238**	-0.232**	-0.357*	-0.189**
			(0.094)	(0.102)	(0.217)	(0.075)
Constant	0.050	0.050	-0.007	-0.006	-0.005	-0.010*
	(0.031)	(0.031)	(0.011)	(0.005)	(0.005)	(0.005)
Trades	70,706	70,706	70,706	70,706	22,452	42,416
Securities	1,032	1,032	1,032	1,032	157	767
Markets	270	270	270	270	79	155
Calendar months	30	30	30	30	30	30
R-squared	0.255	0.018	0.023	0.023	0.043	0.017

Panel B. Firm X

	(1)	(2)	(3)	(4)	(5)	(6)
Markets included	All	All	All	All	2-outcome	5-outcome
Dependent variable	Payoff	Payoff - Price	Payoff - Price	Payoff - Price	Payoff - Price	Payoff - Price
Price	0.969***	-0.031	-0.069			
	(0.069)	(0.069)	(0.117)			
(Price - 1/N)			0.080	0.021	0.062	-0.018
			(0.113)	(0.046)	(0.054)	(0.097)
Constant	0.028	0.028	0.040	0.017	-0.029	0.032***
	(0.025)	(0.025)	(0.039)	(0.011)	(0.019)	(0.009)
Trades	12,655	12,655	12,655	12,655	5,702	2,570
Securities	2,801	2,801	2,801	2,801	825	782
Markets	1,345	1,345	1,345	1,345	818	195
Calendar months	59	59	59	59	59	49
R-squared	0.286	0.0004	0.0013	0.0001	0.0010	0.0001

Panel C. Ford Sales

	(1)	(2)	(3)	(4)
Dependent variable	Payoff	Payoff - Price	Payoff - Price	Payoff - Price
Price	1.046***	0.046	0.057	
	(0.032)	(0.032)	(0.035)	
(Price - Prior Sales)			-0.238	-0.222
			(0.144)	(0.153)
Constant	-0.026***	-0.026***	-0.025***	-0.009
	(0.004)	(0.004)	(0.006)	(0.007)
Trades	3,262	3,262	3,262	3,262
Securities	101	101	101	101
Markets	17	17	17	17
R-squared	0.922	0.022	0.126	0.092

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Each column in each panel represents a regression; each observation in these regressions is a trade. The dependent variable is either the ultimate payoff of a security or the difference between this payoff and the trade price as indicated. The independent variable is either the trade price or the difference between the trade price and a proxy for a naive prior. For the Google and Firm X markets, in which each market consists of securities linked to N mutually exclusive outcomes, we use 1/N as the naive prior probability for each outcome. For the Ford markets, in which each security has a payoff that is a linear function of the sales of a particular group of models in a given week or month, the naive prior is that most recent actual sales figure reported as of the beginning of the market in question. Standard errors are heteroskedasticity robust and allow for clustering on market (for all three firms) and calendar month (for Google and Firm X).

TABLE 4
Markets vs. Experts

Company	Ford	Google	Firm X	
Prediction market type	One continuous outcome	3-5 bins	3-10 bins	One binary outcome
Expert forecast source	Expert forecaster	Derived from Bins	Derived from Bins	Contract over/under
Market topic	Auto sales	Demand	Macro numbers	Macro numbers
Timing of prediction market forecast	Just before expert	First day of PM	First day of PM	First day of PM
Prediction market forecast	0.67	0.82	1.01	1.16
	(0.10)	(0.14)	(0.19)	(0.19)
Expert forecast	0.38	0.09	-0.11	-0.27
	(0.08)	(0.58)	(0.57)	(0.17)
Observations	78	197	1330	748
Unique markets	6	191	185	296
Time periods	13	30	45	58
MSE (prediction market)/MSE(expert)	0.742	0.727	0.924	0.908
P-value of difference with 1	0.104	0.00004	0.002	0.002

This table presents horserace regressions of the outcome being forecast on forecasts from prediction markets and experts. As described in the text, the prediction market and expert forecasts are as contemporaneous as possible. Standard errors are heteroskedasticity robust and allow for clustering on market and, for Google and Firm X, calendar month. In the bottom of the panel, the ratio of the mean squared errors of the two forecasts is reported. For Ford, the expert forecast is a formal expert forecast, whereas for Google and Firm X, the expert forecasts are derived from the prediction market security construction as described in the text.

TABLE 5
Tests for Pricing Biases

Panel A. Google

	(1)	(2)	(3)	(4)
(Price - 1/N)		-0.232***	-0.226**	-0.210**
		(0.089)	(0.090)	(0.088)
Optimism			-0.103**	-0.106***
(+1 if best outcome, -1 if worst)			(0.041)	(0.040)
Extreme outcome				0.130**
Abs(Optimism)				(0.055)
Constant	-0.017***	-0.010**	-0.006	-0.014***
(Captures short aversion)	(0.004)	(0.004)	(0.004)	(0.004)
Trades	37,910	37,910	37,910	37,910
Securities	612	612	612	612
Markets	157	157	157	157
R-squared	0.000	0.025	0.067	0.079

Panel B. Firm X

	(1)	(2)	(3)	(4)
(Price - 1/N)		0.026	0.017	0.020
		(0.050)	(0.050)	(0.051)
Optimism			0.021	0.021
(+1 if best outcome, -1 if worst)			(0.021)	(0.021)
Extreme outcome				0.033
Abs(Optimism)				(0.054)
Constant	-0.003	-0.003	-0.010	-0.010
(Captures short aversion)	(0.013)	(0.013)	(0.014)	(0.014)
Trades	8,910	8,910	8,910	8,910
Securities	1,704	1,704	1,704	1,704
Markets	945	945	945	945
R-squared	0.000	0.000	0.002	0.002

Each observation is a trade; the dependent variable is the percentage point return to expiry (i.e., expiry value - price). 1/N represents a naive prior, with N equal to the number of outcomes for the market (N = 2 for binary markets). Outcomes are ordered based on what would be beneficial for company profits -- the best outcome is scaled +1 and the worst is scaled -1. The extreme outcome measure is the absolute value of an outcome's optimism, less the mean of this value across a market's securities. In the sample size data, a security refers to a unique security with a specific payoff and a market refers to a group of securities with related payoffs (e.g., a group of securities tracking mutually exclusive outcomes). Standard errors are heteroskedasticity robust and allow for clustering on market and calendar month.

TABLE 6
Biases by Subsample

Panel A: Google

	(1)	(2)	(3)	(4)	(5)
	All	Demand Forecasting	Project Completion	Product Quality	External News
Good outcome		High demand	On time	High quality	See text
Price - 1/N	-0.226**	-0.203	-0.247	-0.186	-0.489**
	(0.090)	(0.133)	(0.175)	(0.145)	(0.207)
Optimism	-0.103**	-0.039	-0.239***	-0.085	0.109**
	(0.041)	(0.046)	(0.068)	(0.083)	(0.052)
Constant	-0.006	-0.012*	-0.008	0.006	-0.003
	(0.004)	(0.007)	(0.006)	(0.012)	(0.009)
Trades	37,910	12,387	11590	5897	6,898
Markets	157	51	38	22	42
R-squared	0.067	0.024	0.211	0.207	0.104

Panel B: Firm X

	(1)	(2)	(3)	(4)	(5)
	All	Politics	Policy	Stocks	Growth
Good outcome		GOP wins	GOP policies	High values	Rapid growth
Price - 1/N	0.017	-0.058	-0.081	-0.257	-0.013
	(0.050)	(0.221)	(0.086)	(0.158)	(0.125)
Optimism	0.021	0.163*	0.034	0.001	0.049
	(0.021)	(0.090)	(0.120)	(0.086)	(0.032)
Constant	-0.010	-0.026	-0.004	0.028	-0.019
	(0.014)	(0.055)	(0.075)	(0.060)	(0.029)
Trades	8,910	449	382	1,309	2,205
Markets	945	35	12	53	425
R-squared	0.002	0.119	0.005	0.018	0.003

Panel B (continued): Firm X

	(6)	(7)	(8)	(9)	(10)	(11)
	Jobs	Commodities	Exchange Rates	Eurozone	Energy	Inflation
Good outcome	More jobs	Higher prices	Weak dollar	No crisis	Higher prices	Faster inflation
Price - 1/N	-0.453	0.699***	0.129	0.452***	0.412***	0.034
	(0.278)	(0.123)	(0.090)	(0.090)	(0.116)	(0.122)
Optimism	0.175***	-0.019	0.012	-0.050	0.020	-0.095**
	(0.044)	(0.043)	(0.066)	(0.057)	(0.051)	(0.041)
Constant	0.043**	0.019	-0.059	0.002	-0.062**	-0.006
	(0.017)	(0.038)	(0.062)	(0.017)	(0.024)	(0.018)
Trades	657	290	462	166	492	1,429
Markets	39	35	46	20	41	93
R-squared	0.088	0.143	0.005	0.120	0.055	0.032

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Regressions identical to those in Table 5, Column 3 are presented for subsets of the Google and Firm X markets. Only markets for which optimism can be signed are included, and thus all "Fun" markets are excluded. See text for more details on the rationale applied in signing the optimism of different categories of outcome. Standard errors are heteroskedasticity robust and allow for clustering on market and calendar month.

TABLE 7
Optimism and Stock Returns

	(1)	(2)	(3)	(4)	(5)	(6)
Optimism*Google log stock return (t+1)	-0.869	-0.228	-0.303	-1.006	-0.646	-0.830
	(0.720)	(0.659)	(0.671)	(0.675)	(0.603)	(0.565)
Optimism*Google log stock return (t)	-1.158	-0.185	-0.243	-0.015	0.196	0.255
	(0.796)	(0.455)	(0.434)	(0.610)	(0.613)	(0.488)
Optimism*Google log stock return (t-1)	-2.022***	-1.318**	-1.296**	-2.618***	-2.112***	-1.414**
	(0.744)	(0.569)	(0.561)	(0.767)	(0.658)	(0.628)
Optimism*Google log stock return (t-2)	-0.695	0.037	0.063	-0.103	-0.043	-0.042
	(0.436)	(0.302)	(0.287)	(0.316)	(0.354)	(0.316)
Topics included	All	All	All	Completion	Completion	Completion
Google stock returns (t+1, t, t-1, t-2)	Y	Y	Y	Y	Y	Y
Interactions of Google stock returns (t+1 to t-2) with calendar quarter fixed effects		Y	Y	Y	Y	Y
Interactions of Google stock returns (t+1, t, t-1, t-2) with extremeness and price-1/N			Y	Y	Y	Y
S&P and Nasdaq returns (t+1, t, t-1, t-2) and interactions with optimism					Y	Y
Day of week fixed effects and interactions with optimism						Y
Observations	37,910	37,910	37,910	11,590	11,590	11,590
R-squared	0.095	0.155	0.159	0.489	0.505	0.522

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The regressions in this table extend the regression in Table 5, Panel A, Column 4 by adding Google stock returns from surrounding periods and their interaction with the optimism variable. Columns 1-3 include all trades included in Table 4, Column 5 (i.e., all markets for which optimism can be signed), while columns 4-6 include only markets on the timing of project completion (i.e., those included in Table 6, Column 3). Standard errors are heteroskedasticity-robust and allow for clustering within markets and calendar months.

TABLE 8
Pricing Biases over the life of markets

Panel A. Google

	Trades 1-10	Trades 11-50	Trades 50+
(Price - Naïve Prior)	-0.475***	-0.340***	-0.126
	(0.085)	(0.069)	(0.127)
Optimism	-0.013	-0.081**	-0.140**
(+1 if best outcome, -1 if worst)	(0.033)	(0.033)	(0.055)
Constant	-0.007*	-0.010*	-0.001
(Captures short aversion)	(0.004)	(0.005)	(0.007)
Trades	5,251	13,737	18,922
Markets	157	144	81
R-squared	0.069	0.069	0.098

Panel B. Firm X

	Trades 1-10	Trades 11-25	Trades 26+
(Price - Naïve Prior)	0.003	0.055	-0.008
	(0.059)	(0.088)	(0.167)
Optimism	0.020	0.014	0.094
(+1 if best outcome, -1 if worst)	(0.019)	(0.042)	(0.102)
Constant	-0.006	-0.029	-0.038
(Captures short aversion)	(0.012)	(0.035)	(0.117)
Trades	7650	1,129	131
Markets	945	187	12
R-squared	0.001	0.004	0.066

Panel C. Ford Sales

	Trades 1-10	Trades 11-50	Trades 51+
(Price - Naïve Prior)	-0.122	-0.178	-0.811***
	(0.121)	(0.152)	(0.166)
Constant	-0.006	-0.011	-0.004
(Captures optimism and short aversion)	(0.008)	(0.008)	(0.005)
Trades	957	1747	558
Markets	101	86	20
R-squared	0.034	0.059	0.710

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Regressions identical to those in Table 3, Column 4 are presented, except that trades in each security are numbered sequentially and the sample is split according to trade number. Standard errors are heteroskedasticity robust and allow for clustering on market (for all three firms) and calendar month (for Google and Firm X).

TABLE 9
Reduction in Biases Over Time

	Google	Ford	Firm X
(Price - Prior)	-0.379***	-0.290**	0.063
	(0.134)	(0.109)	(0.120)
(Price - Prior)*Date	0.355	-1.251***	-0.114
	(0.287)	(0.227)	(0.212)
Optimism	-0.210***		0.088*
	(0.061)		(0.052)
Optimism*Date	0.274**		-0.129
	(0.115)		(0.082)
Constant	-0.010	-0.012	0.074
	(0.012)	(0.011)	(0.053)
Constant*Date (Min 0, Max 1)	-0.003	0.049*	-0.050
	(0.005)	(0.027)	(0.037)
Trades	37,910	3,262	8,910
R-squared	0.090	0.26	0.006

Regressions identical to those in Table 3, Column 4 (for Ford) and Table 5, Column 3 (for Google and Firm X) are presented, with the variables interacted with a linear time trend, which is scaled to equal 0 at the beginning of the sample and 1 at the end. Standard errors are heteroskedasticity robust and allow for clustering on market (for all three firms) and calendar month (for Google and Firm X).

TABLE 10
Biases, Experience and cumulative returns

Panel A: Google

	(1) Optimism	(2) Price - Prior	(3) Buy	(4) Returns
Cumulative Returns	-0.520**	-0.036	-0.587**	0.178**
	(0.215)	(0.065)	(0.247)	(0.071)
Experience	-0.019**	-0.032***	-0.122***	0.015***
	(0.009)	(0.003)	(0.019)	(0.003)
Observations	75,820	141,412	141,412	141,412
R-squared	0.005	0.049	0.055	0.006

Panel B: Ford Sales

	(1) Buy/Optimism	(2) Price - Prior	(3) Returns
Cumulative Returns	-0.149	-0.017**	0.023***
	(0.126)	(0.007)	(0.006)
Experience	-0.131***	-0.006***	0.006***
	(0.025)	(0.002)	(0.002)
Observations	2,810	2,810	2,810
R-squared	0.023	0.01	0.019

Panel C: Firm X

	(1) Optimism	(2) Price - Prior	(3) Buy	(4) Returns
Cumulative Returns	-5.804	0.927	10.658	6.984***
	(5.075)	(1.607)	(9.324)	(2.490)
Experience	0.002	0.021***	-0.068**	-0.011***
	(0.011)	(0.005)	(0.033)	(0.004)
Observations	8,696	12,318	12,318	12,318
R-squared	0.001	0.018	0.010	0.005

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table presents regressions testing whether traders with more past experience or higher past returns trade in a direction that is correlated with certain security characteristics or with future returns. In Google's markets, each trade has two participants (a buyer and a seller), and thus each trade appears in the dataset twice. For Ford and Firm X, participants trade with an automated market maker, and so each trade appears in the data once. For each observation, the dependent variable is a security characteristc multipled by the side (+1 if a buy, -1 if a sell). The dependent variable "Buy" is this side variable; "Returns" is returns to expiry multiplied by side. Standard errors are heteroskedasticity-robust and allow for clustering within participants and markets. Regressions include fixed effects for trades for Google and time periods for Ford and Firm X (weeks and months, respectively).

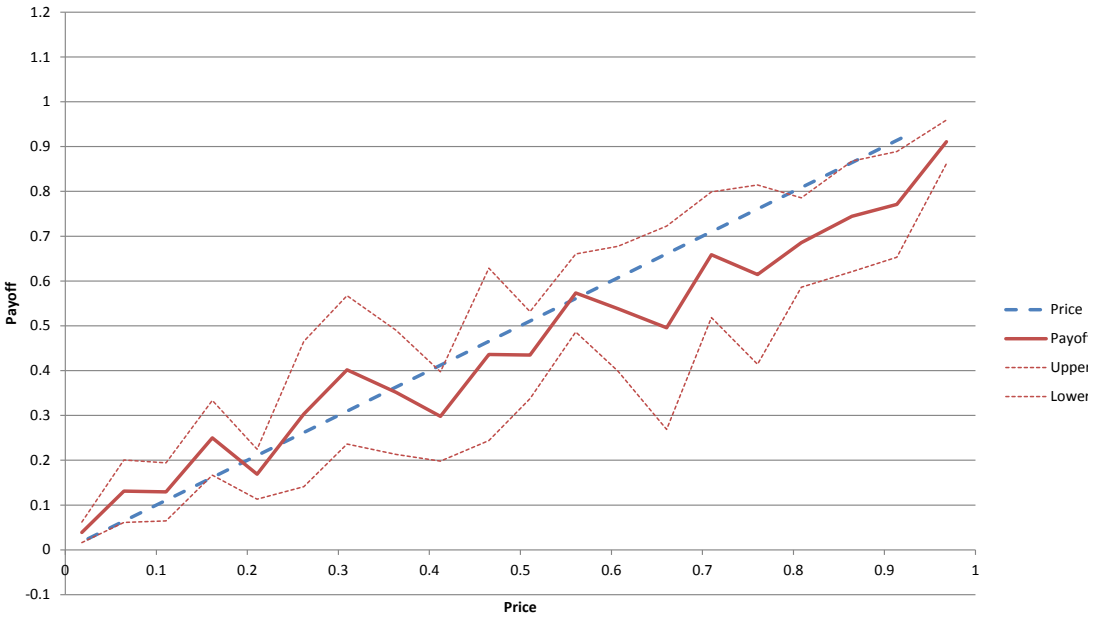
TABLE 11
Trader Characteristics, Biases and Returns at Google

	(1)	(2)	(4)	(5)
	Optimism	Price - Prior	Buy	Returns
Market Insider	0.127	0.066*	0.401***	-0.012
	(0.099)	(0.035)	(0.153)	(0.046)
Friend of Insider	0.146**	0.033*	-0.200	-0.026
	(0.069)	(0.019)	(0.179)	(0.019)
Coder / Engineer	-0.008	-0.156***	-0.467***	0.102***
	(0.084)	(0.028)	(0.133)	(0.031)
Hire Date (in Years)	0.073**	-0.014	-0.152**	0.007
	(0.031)	(0.010)	(0.062)	(0.012)
NYC-Based	-0.152	-0.117***	-0.089	0.064*
	(0.115)	(0.033)	(0.159)	(0.037)
Mountain-View Based	-0.177*	-0.039	-0.130	0.037
	(0.094)	(0.026)	(0.105)	(0.025)
Observations	75,820	141,412	141,412	141,412
R-squared	0.009	0.046	0.065	0.005

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

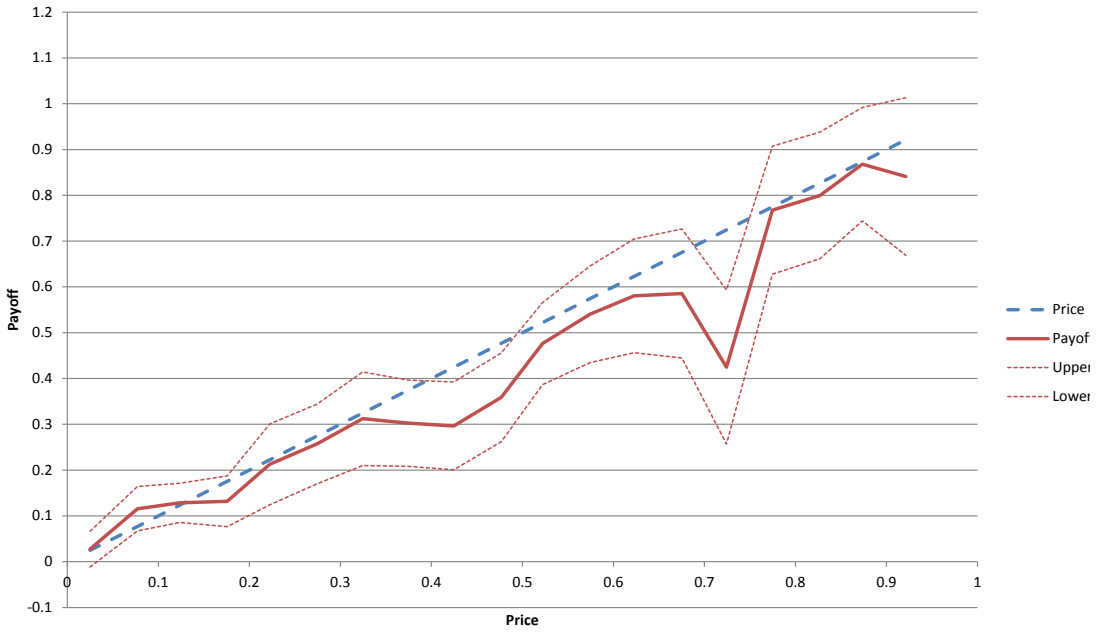
This Table presents regressions analogous to those in Table 10, Panel A, except that traded characteristics are included rather than experience variables. A market insider is a participant on the project covered by the market. Friends of insiders are as indicated by either party on a social networking survey. Standard errors are heteroskedasticity-robust and adjust for clustering within participants and markets.

Figure 1. Prices and payoffs in Google's prediction markets



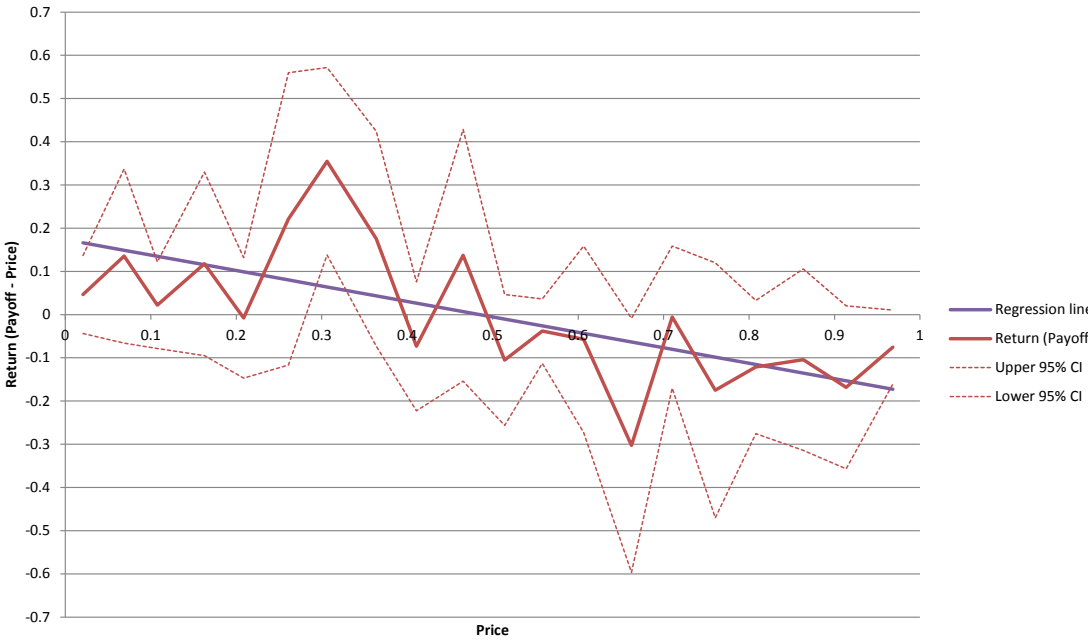
Trades in Google's prediction markets (N = 70,706) are sorted into 20 bins based on their price (0-0.05, 0.05-0.01, etc.). The graph plots the average price and ultima
payoff for each bin. 95% confidence intervals are reported for payoffs, based on standard errors that allow for clustering in payoffs for securities with related outcor

Figure 2. Prices and payoffs in Firm X's binary markets



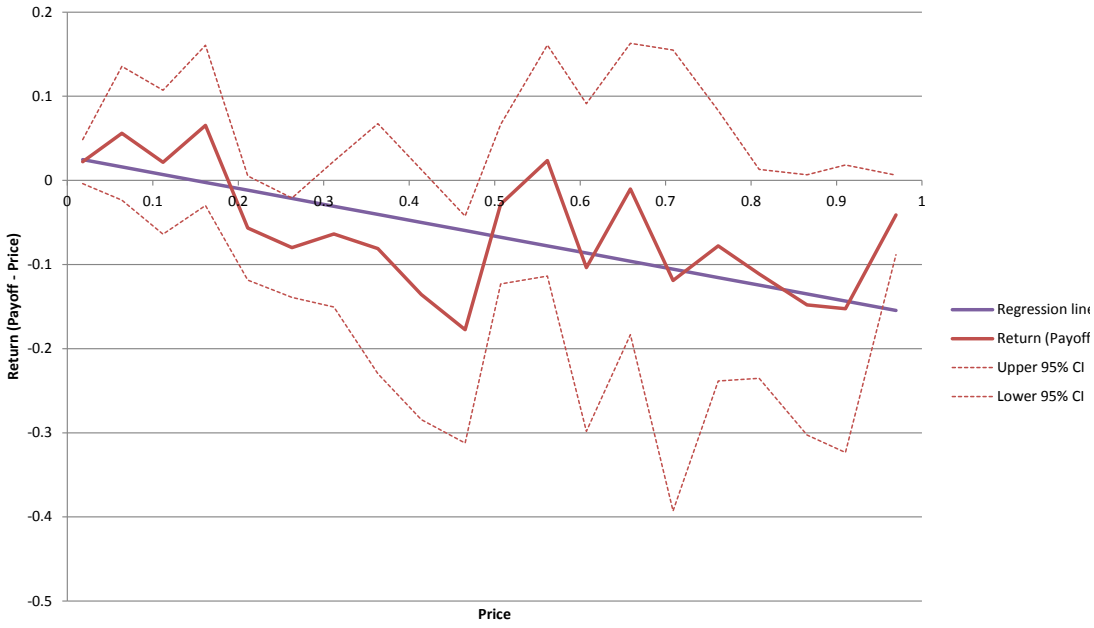
Trades in Firm X's binary prediction markets (N = 9,237) are sorted into 20 bins based on their price (0-0.05, 0.05-0.1, etc.). The graph plots the average price and ultima payoff for each bin. 95% confidence intervals are reported for payoffs, based on standard errors that allow for clustering in payoffs for securities with related outcomes.

Figure 3A. Prices and returns in Google's 2-outcome markets



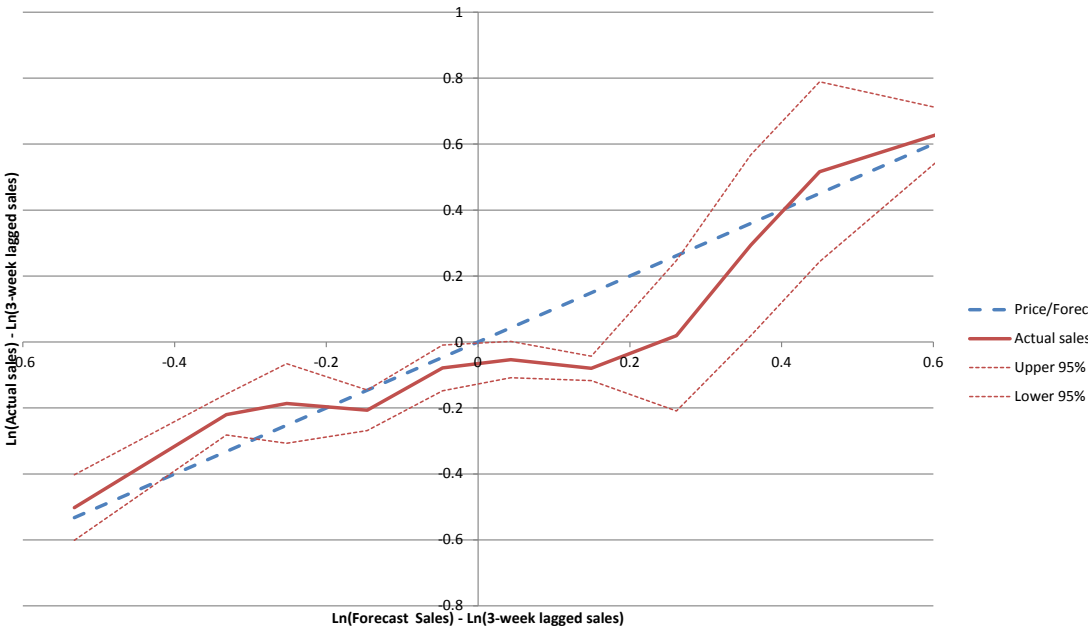
Trades in Google's two-outcome prediction markets (N = 22,452) are sorted into 20 bins based on their price (0-0.05, 0.05-0.1, etc.). The graph plots the average price and ultimate payoff for each bin. 95% confidence intervals are reported for payoffs, based on standard errors that allow for clustering in payoffs for securities with relat outcomes.

Figure 3B. Prices and returns in Google's 5-outcome markets



Trades in Google's five-outcome prediction markets (N = 42,416) are sorted into 20 bins based on their price (0-0.05, 0.05-0.1, etc.). The graph plots the average price and ultimate payoff for each bin. 95% confidence intervals are reported for payoffs, based on standard errors that allow for clustering in payoffs for securities with rela outcomes.

Figure 4. Forecast and actual sales in Ford's sales markets



Trades in Ford's sales prediction markets (N = 3,262) are sorted into bins based on the log difference between the sales predicted by their price and three-week-prior s
for the given model. The graph plots the average price and ultimate payoff for each bin. 95% confidence intervals are reported for payoffs, based on standard errors
allow for clustering in payoffs for securities for the same time period.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), pp. 251–333.

Abramowicz, Michael and Todd Henderson, "Prediction markets for corporate governance," *bepress Legal Series*, 2006, p. 1571.

Acemoglu, Daron, "Training and Innovation in an Imperfect Labour Market," *Review of Economic Studies*, 1997, *64* (3), 445–464.

_ **and Alexander Wolitzky**, "The Economics of Labor Coercion," 2009. Mimeo, MIT.

_ **and Jorn-Steffen Pischke**, "Why Do Firms Train? Theory And Evidence," *Quarterly Journal of Economics*, 1998, *113* (1), 78–118.

_ **and** _ , "Beyond Becker: Training in Imperfect Labour Markets," *Economic Journal*, 1999a, *109* (453), F112–42.

_ **and** _ , "The Structure of Wages and Investment in General Training," *Journal of Political Economy*, 1999b, *107* (3), 539–572.

_ **and Martin Kaae Jensen**, "Robust comparative statics in large static games," in "Decision and Control (CDC), 2010 49th IEEE Conference on" IEEE 2010, pp. 3133–3139.

Acland, Dan and Matthew Levy, "Habit Formation, Naivete, and Projection Bias in Gym Attendance," 2011. Mimeo.

Adams, William, Liran Einav, and Jonathan Levin, "Liquidity Constraints and Imperfect Information in Subprime Lending," *American Economic Review*, 2009, *99* (1), 49–84.

Adler, Paul S. and Seok-Woo Kwon, "Social Capital: Prospects for a New Concept," *The Academy of Management Review*, 2002, *27* (1), pp. 17–40.

Aghion, Philippe and Jean Tirole, "Formal and Real Authority in Organizations," *Journal of Political Economy*, 1997, *105* (1), 1–29.

_ **, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt**, "Competition and Innovation: an Inverted-U Relationship," *The Quarterly Journal of Economics*, 2005, *120* (2), 701–728.

Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis, "Recent Developments in the Economics of Time Use," *Annu. Rev. Econ.*, 2012, *4* (1), 373–397.

Allison, Paul D and Nicholas A Christakis, "Logit models for sets of ranked items," *Sociological methodology*, 1994, *24* (1994), 199–228.

Altonji, Joseph and Rebecca Blank, "Race and Gender in the Labor Market," *Handbook of Labor Economics*, 1999.

Altonji, Joseph G. and Charles R. Pierret, "Employer Learning And Statistical Discrimination," *Quarterly Journal of Economics*, 2001, *116* (1), 313–350.

Amabile, Teresa M, *Creativity in context: Update to" the social psychology of creativity."*, Westview press, 1996.

Anderson, Jon, Stephen Burks, Jeffrey Carpenter, Lorenz Gtte, Karsten Maurer, Daniele Nosenzo, Ruth Potter, Kim Rocha, and Aldo Rustichini, "Self-selection and variations in the laboratory measurement of other-regarding preferences across subject pools: evidence from one college student and two adult samples," *Experimental Economics*, 2012, *Forthcoming*, 1–20.

Anderson, Lisa and Charles Holt, "Information Cascades in the Laboratory," *American Economic Review*, 1997, *87* (5), 847–862.

Anenberg, Elliot, "Uncertainty, Learning and the Value of Information in the Residential Real Estate Market," 2011. Mimeo.

Arbatskaya, Maria and Hugo M Mialon, "Dynamic Multi-Activity Contests," *Scandinavian Journal of Economics*, 2010.

Arcidiacono, Peter, "Ability Sorting and the Returns to College Major," *Journal of Econometrics*, 2004, *121* (1-2), 343–375.

_ , "Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?," *Econometrica*, 2005, *73* (5), 1477–1524.

Arkes, Hal R., Caryn Christenson, Cheryl Lai, and Catherine Blumer, "Two Methods of Reducing Overconfidence," *Organizational Behavior and Human Decision Processes*, 1987, *39*, 133–144.

Armstrong, Christopher S., David F. Larcker, and Che-Lin Su, "Endogenous Selection and Moral Hazard in Compensation Contracts," *Operations Research*, 2010, *58* (4-Part-2), 1090–1106.

Arrow, Kenneth J, Robert Forsythe, Michael Gorham, Robert Hahn, Robin Hanson, John O Ledyard, Saul Levmore, Robert Litan, Paul Milgrom, Forrest D Nelson et al., "The promise of prediction markets," *Science*, 2008, *320* (5878), 877–878.

Asch, Beth and John Warner, "A Theory of Military Compensation and Personnel Policy," *RAND, Monograph Report*, 1994.

Asch, Beth J and John T Warner, "A Theory of Compensation and Personnel Policy in Hierarchical Organizations with Application to the United States Military," *Journal of Labor Economics*, 2001, *19* (3), 523–62.

Ashraf, Nava, Erica Field, and Jean Lee, "Household Bargaining and Excess Fertility: An Experimental Study in Zambia," 2009. Mimeo.

Autor, David H., "Why Do Temporary Help Firms Provide Free General Skills Training?," *Quarterly Journal of Economics*, 2001, *116* (4), 1409–1448.

_ , "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing," *Journal of Labor Economics*, 2003, *21* (1), 1–42.

_ **and David Scarborough**, "Does Job Testing Harm Minority Workers? Evidence from Retail Establishments," *Quarterly Journal of Economics*, 2 2008, *123* (1), 219–277.

Azoulay, Pierre, Joshua S Graff Zivin, and Gustavo Manso, "Incentives and creativity: evidence from the academic life sciences," *The RAND Journal of Economics*, 2011, *42* (3), 527–554.

Babcock, Linda, George Loewenstein, and Samuel Issacharoff, "Creating Convergence: Debiasing Biased Litigants," *Law and Social Inquiry*, 1997, *22* (4), 913–925.

Bach, Norbert, Oliver Gürtler, and Joachim Prinz, "Incentive Effects in Tournaments with Heterogeneous Competitorsan Analysis of the Olympic Rowing Regatta in Sydney 2000," *management revue*, 2009, pp. 239–253.

Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin, "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 09 2007, *75* (5), 1331–1370.

_ **, Han Hong, John Krainer, and Dennis Nekipelov**, "Estimating Static Models of Strategic Interactions," Mimeo 2006.

_ **, Victor Chernozhukov, Han Hong, and Denis Nekipelov**, "Nonparametric and Semiparametric Analysis of a Dynamic Discrete Game," 2009. Mimeo.

Baker, George, Michael Gibbs, and Bengt Holmstrom, "The Internal Economics of the Firm: Evidence from Personnel Data," *Quarterly Journal of Economics*, 1994, *109* (4), pp. 881–919.

_ , _ , **and** _ , "The Wage Policy of a Firm," *Quarterly Journal of Economics*, 1994, *109* (4), pp. 921–955.

Baker, George P, "Incentive contracts and performance measurement," *Journal of Political economy*, 1992, pp. 598–614.

Baker, George P., "Incentive Contracts and Performance Measurement," *The Journal of Political Economy*, 1992, *100* (3), pp. 598–614.

_ , "Distortion and Risk in Optimal Incentive Contracts," *Journal of Human Resources*, 2002, *37* (4), 728–751.

_ **and Thomas N. Hubbard**, "Contractibility And Asset Ownership: On-Board Computers and Governance In U. S. Trucking," *Quarterly Journal of Economics*, 2004, *119* (4), 1443–1479.

Balafoutas, Loukas, Florian Lindner, and Matthias Sutter, "Sabotage in tournaments: evidence from a natural experiment," *Kyklos*, 2012, *65* (4), 425–441.

Balmaceda, Felipe, "Firm-Sponsored General Training," *Journal of Labor Economics*, 2005, *23* (1), 115–134.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Social Connections and Incentives in the Workplace: Evidence From Personnel Data," *Econometrica*, 2009, *77* (4), 1047–1094.

_ , _ , **and** _ , "Social Incentives in the Workplace," *Review of Economic Studies*, 2010, *77* (2), 417–458.

_ , _ , **and** _ , "Team incentives: Evidence from a firm level experiment," *Journal of the European Economic Association*, 2013, *11* (5), 1079–1114.

Barankay, Iwan, "Rank incentives: Evidence from a randomized workplace experiment," *Discussion papers*, 2012.

Barber, Brad M. and Terrance Odean, "Boys Will Be Boys: Gender, Overconfidence, And Common Stock Investment," *Quarterly Journal of Economics*, 2001, *116* (1), 261–292.

Baron, James and David Kreps, *Strategic Human Resources: Frameworks for General Managers*, New York, NY: Wiley, 1999.

Barro, Robert and Xavier Sala i Martin, *Economic Growth*, 2nd ed., Cambridge, MA: MIT Press, 2004.

Barron, John M., Mark C. Berger, and Dan A. Black, "Do Workers Pay for On-The-Job Training?," *Journal of Human Resources*, 1999, *34* (2), 235–252.

Bartel, Ann P, "Training, Wage Growth, and Job Performance: Evidence from a Company Database," *Journal of Labor Economics*, 1995, *13* (3), 401–25.

Baye, Michael R and Heidrun C Hoppe, "The strategic equivalence of rent-seeking, innovation, and patent-race games," *Games and Economic Behavior*, 2003, *44* (2), 217–226.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa, "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," *Journal of Political Economy*, December 2008, *116* (6), 1150–1196.

Bayus, Barry L, "Crowdsourcing New Product Ideas over Time: An Analysis of the Dell IdeaStorm Community," *Management Science*, 2013, *59* (1), 226–244.

Beaman, Lori, "Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S.," *Review of Economic Studies*, 2012, *79* (1).

_ **and Jeremy Magruder**, "Who Gets the Job Referral? Evidence from a Social Networks Experiment," *American Economic Review*, 2012, *102* (7), 3574–93.

_ **, Niall Keleher, and Jeremy Magruder**, "Do Job Networks Disadvantage Women? Evidence From a Recruitment Experiment in Malawi," working paper, 2013.

Becker, Brian E and Mark A Huselid, "The incentive effects of tournament compensation systems.," *Administrative Science Quarterly*, 1992, *37* (2).

Becker, Gary, *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, Chicago, IL: University of Chicago Press, 1964.

Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 1968, *76*, 169.

Bellemare, Charles and Bruce Shearer, "On The Relevance And Composition Of Gifts Within The Firm: Evidence From Field Experiments," *International Economic Review*, 2011, *52* (3), 855–882.

Belman, Dale L. and Kristen A. Monaco, "The Effects of Deregulation, De-Unionization, Technology, and Human Capital on the Work and Work Lives of Truck Drivers," *Industrial and Labor Relations Review*, 2001, *54* (2A), pp. 502–524.

_ **,** _ **, and Taggert Brooks**, *Sailors of the Conrete Sea: A Portrait of Truck Drivers' Work and Lives*, Lansing, MI: Michigan State University Press, 2005.

Belzer, Michael, *Sweatshops on Wheels: Winners and Losers in Trucking Deregulation*, New York, NY: Oxford University Press, 2000.

Belzil, Christian, J. Hansen, and Nicolai Kristensen, "Estimating Complementarity Between Education and Training," 2009. Mimeo.

Bénabou, Roland and Jean Tirole, "Self-confidence and personal motivation," *The Quarterly Journal of Economics*, 2002, *117* (3), 871–915.

Benabou, Roland and Jean Tirole, "Intrinsic and extrinsic motivation," *The Review of Economic Studies*, 2003, *70* (3), 489–520.

Benoit, Jean-Pierre and Juan Dubra, "Apparent Overconfidence," *Econometrica*, 2011, *79* (5), 1591–1625.

Berg, Joyce, Dickhaut John, and McCabe Kevin, "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, July 1995, *10* (1), 122–142.

Berg, Joyce E and Thomas A Rietz, "Prediction markets as decision support systems," *Information Systems Frontiers*, 2003, *5* (1), 79–93.

Berg, Joyce, Robert Forsythe, Forrest Nelson, and Thomas Rietz, "Results from a dozen years of election futures markets research," *Handbook of Experimental Economics Results*, 2008, *1*, 742–751.

Berger, Johannes and Petra Nieken, "Heterogeneous Contestants and the Intensity of Tournaments An Empirical Investigation," *Journal of Sports Economics*, 2014, p. 1527002514538639.

Bergman, Nittai K. and Dirk Jenter, "Employee Sentiment and Stock Option Compensation," *Journal of Financial Economics*, 2007, *84* (3), 667–712.

Bernardo, Antonio E and Ivo Welch, "On the evolution of overconfidence and entrepreneurs," *Journal of Economics & Management Strategy*, 2001, *10* (3), 301–330.

Bernheim, B. Douglas and Antonio Rangel, "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics," *Quarterly Journal of Economics*, 2009, *124* (1), 51–104.

Besley, Timothy and Maitreesh Ghatak, "Status incentives," *The American Economic Review*, 2008, pp. 206–211.

Bingham, Kelly, "Paying for Truck Driving School," 2007.

Blinder, Alan S. and Alan B. Krueger, "Labor Turnover in the USA and Japan: A Tale of Two Countries," *Pacific Economic Review*, 1996, *1* (1), 27–57.

Bloom, Nicholas and John Van Reenen, "Patents, Real Options and Firm Performance," *Economic Journal*, 2002, *112* (478), C97–C116.

_ **and** _ , "Human Resource Management and Productivity," *Handbook of Labor Economics*, 2011, pp. 1697–1767.

BLS, "Bureau of Labor Statistics Occupational Outlook Handbook, 2010-11: Truck Drivers and Driver/Sales Workers," 2010.

Blume, Brian D, Timothy T Baldwin, Robert S Rubin, and William Bommer, "ALL FORCED RANKING SYSTEMS ARE NOT CREATED EQUAL: A POLICY CAPTURING STUDY.," in "Academy of Management Proceedings," Vol. 2006 Academy of Management 2006, pp. H1–H6.

Bodily, Samuel E and Robert F Bruner, *Enron, 1986-2001*, Darden School, UVA, 2002.

Bojilov, Raicho, "Incentives to Work or Incentives to Quit?," 2011. Mimeo.

Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas ter Weel, "The Economics and Psychology of Personality Traits," *Journal of Human Resources*, 2008, *43* (4), 972–1059.

Boudreau, Kevin J, Nicola Lacetera, and Karim R Lakhani, "Incentives and problem uncertainty in innovation contests: An empirical analysis," *Management Science*, 2011, *57* (5), 843–863.

Boyer, Kenneth D. and Stephen V. Burks, "Stuck in the Slow Lane: Undoing Traffic Composition Biases in the Measurement of Trucking Productivity," *Southern Economic Journal*, 2009, *75* (4), 1220–1237.

Branch, Gregory F., Eric A. Hanushek, and Steven G. Rivkin, "Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals," Working Paper 17803, National Bureau of Economic Research February 2012.

Breaugh, James A., "Relationships between Recruiting Sources and Employee Performance, Absenteeism, and Work Attitudes," *The Academy of Management Journal*, 1981, *24* (1), pp. 142–147.

_ , "Employee recruitment: Current knowledge and important areas for future research," *Human Resource Management Review*, 2008, *18* (3), 103 – 118.

_ **and Rebecca B. Mann**, "Recruiting source effects: A test of two alternative explanations," *Journal of Occupational Psychology*, 1984, *57* (4), 261–267.

Brown, Clair and Greg Linden, "Chips and Change," 2009.

Brown, Jennifer, "Quitters never win: The (adverse) incentive effects of competing with superstars," *Journal of Political Economy*, 2011, *119* (5), 982–1013.

Brown, Meta, Elizabeth Setren, and Giorgio Topa, "Do Informal Referrals Lead to Better Matches? Evidence from a Firm's Employee Referral System," New York Federal Reserve Bank working paper, 2013.

Brunello, Giorgio and Alfredo Medio, "An explanation of international differences in education and workplace training," *European Economic Review*, 2001, *45* (2), 307–322.

Brunnermeier, Markus K. and Jonathan A. Parker, "Optimal Expectations," *American Economic Review*, 2005, *95* (4), 1092–1118.

Bryant, Adam, "In head-hunting, big data may not be such a big deal," *The New York Times*, 2013.

Buddin, Richard J., "Success of First-Term Soldiers: The Effects of Recruiting Practices and Recruit Characteristics," 2005. Santa Monica, CA: RAND Corporation.

_ **and Kanika Kapur**, "The Effect of Employer-Sponsored Education on Job Mobility: Evidence from the U.S. Navy," *Industrial Relations*, 2005, *44*, 341–363.

Buera, Francisco, "A dynamic model of entrepreneurship with borrowing constraints: theory and evidence," *Annals of Finance*, 2009, *5*, 443–464.

Bull, Clive, Andrew Schotter, and Keith Weigelt, "Tournaments and piece rates: An experimental study," *The Journal of Political Economy*, 1987, pp. 1–33.

Burdett, Kenneth and Kenneth L. Judd, "Equilibrium Price Dispersion," *Econometrica*, 1983, *51* (4), pp. 955–969.

Burks, Stephen, Bo Cowgill, Mitchell Hoffman, and Michael Housman, "The value of hiring through referrals," *Available at SSRN 2253738*, 2013.

_ **,** _ **,** _ **, and** _ , "The Value of Hiring through Referrals," 2013. Mimeo.

_ **, Jeffrey Carpenter, and Lorenz Goette**, "Performance pay and worker cooperation: Evidence from an artefactual field experiment," *Journal of Economic Behavior & Organization*, June 2009, *70* (3), 458–469.

_ , _ , _ , **and Aldo Rustichini**, "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment," *Proceedings of the National Academy of Science*, 2009, *106* (19), 7745–7750.

_ , _ , _ , **and** _ , "Overconfidence is a Social Signaling Bias," *Review of Economic Studies*, 2012, *Forthcoming*.

_ , _ , _ , **Kristen Monaco, Kay Porter, and Aldo Rustichini**, "Using Behavioral Economic Field Experiments at a Firm: The Context and Design of the Truckers and Turnover Project," in *The Analysis of Firms and Employees: Quantitative and Qualitative Approaches*, S. Bender et al., ed. (Amsterdam: University of Chicago Press, 2008).

_ , **Michael Belzer, Q. Kwan, S. Pratt, and S. Shackelford**, "Trucking 101: An Industry Primer," *Transportation Research Board*, 2011.

Caldwell, David F. and W. Austin Spivey, "The Relationship Between Recruiting Source and Employee Success: An Analysis by Race," *Personnel Psychology*, 1983, *36* (1), 67–72.

Calvo-Armengol, Antoni and Matthew O. Jackson, "The Effects of Social Networks on Employment and Inequality," *American Economic Review*, 2004, *94* (3), 426–454.

Camerer, Colin F and Robin M Hogarth, "The effects of financial incentives in experiments: A review and capital-labor-production framework," *Journal of risk and uncertainty*, 1999, *19* (1), 7–42.

Cameron, Colin and Pravin Trivedi, *Microeconometrics: Methods and Applications* (Cambridge: Cambridge University Press, 2005).

Cameron, Lisa A, "Raising the stakes in the ultimatum game: Experimental evidence from Indonesia," *Economic Inquiry*, 1999, *37* (1), 47–59.

Cameron, Stephen V. and James J. Heckman, "The Dynamics of Educational Attainment for Black, Hispanic, and White Males," *Journal of Political Economy*, 2001, *109* (3), 455–499.

Cappelli, Peter, "Why Do Employers Pay for College?," *Journal of Econometrics*, 2004, *121* (1-2), 213–241.

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez, "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction," *American Economic Review*, 2012, *102* (6), 2981–3003.

_ **and Dean R. Hyslop**, "Estimating the Effects of a Time-Limited Earnings Subsidy for Welfare-Leavers," *Econometrica*, 2005, *73* (6), 1723–1770.

_ **, Jochen Kleve, and Andrea Weber**, "Active Labor Market Policy Evaluations: A Meta-Analysis," 2009. Mimeo.

_ **, Joerg Heining, and Patrkick Kline**, "Workplace Heterogeneity and the Rise of German Wage Inequality," 2012. Mimeo.

CareerBuilder, "Referral Madness: How Employee Referral Programs Turn Good Employees Into Great Recruiters and Grow Your Bottom Line," CareerBuilder e-Book 2012.

Carlson, Nicholas, *Marissa Mayer and the Fight to Save Yahoo!*, Twelve, 2014.

Carmichael, H Lorne, "Incentives in academics: Why is there tenure?," *Journal of Political Economy*, 1988, *96* (3), 453–72.

Carmichael, Lorne, "Firm-Specific Human Capital and Promotion Ladders," *Bell Journal of Economics*, Spring 1983, *14* (1), 251–258.

Carneiro, Pedro, James J. Heckman, and Edward J. Vytlacil, "Estimating Marginal Returns to Education," *American Economic Review*, 2011, *101* (6), 2754–81.

Carpenter, Jeffrey, Eric Verhoogen, and Stephen Burks, "The effect of stakes in distribution experiments," *Economics Letters*, 2005, *86* (3), 393–398.

_ **, Peter Hans Matthews, and John Schirm**, "Tournaments and office politics: Evidence from a real effort experiment," *The American Economic Review*, 2010, *100* (1), 504–517.

Carrell, Scott E, Bruce I Sacerdote, and James E West, "From natural variation to optimal policy? The importance of endogenous peer group formation," *Econometrica*, 2013, *81* (3), 855–882.

Carter, Nathan, "Group Explorer," 2007.

Casas-Arce, Pablo and F Asís Martínez-Jerez, "Relative performance compensation, contests, and dynamic incentives," *Management Science*, 2009, *55* (8), 1306–1320.

Casella, Alessandra and Nobuyuki Hanaki, "Information Channels in Labor Markets: On the Resilience of Referral Hiring," *Journal of Economic Behavior & Organization*, 2008, *66* (3-4), 492–513.

Castilla, Emilio J., "Social Networks and Employee Performance in a Call Center," *American Journal of Sociology*, 2005, *110* (5), pp. 1243–1283.

Caves, Richard E and Michael E Porter, "From entry barriers to mobility barriers: Conjectural decisions and contrived deterrence to new competition*," *The Quarterly Journal of Economics*, 1977, pp. 241–261.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein, "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design," *Quarterly Journal of Economics*, 2010, *125* (1), 215–261.

Cesarini, David, Magnus Johannesson, Paul Lichtenstein, and Bjorn Wallace, "Heritability of Overconfidence," *Journal of the European Economic Association*, 2009, *7* (2-3), 617–627.

Chan, Tat, Barton Hamilton, and Christopher Makler, "Using Expectations Data to Infer Managerial Objectives and Choices," 2008. Mimeo.

Charness, Gary, Aldo Rustichini, and Jeroen van de Ven, "Overconfidence, Subjectivity, and Self-Deception," 2010. Mimeo.

Che, Yeon-Koo, "Design competition through multidimensional auctions," *The RAND Journal of Economics*, 1993, pp. 668–680.

Chen, Kay-Yut and Charles R Plott, "Information aggregation mechanisms: Concept, design and implementation for a sales forecasting problem," *California Institute of Technology Social Science Working Paper*, 2002, *1131.*

Chen, Yiling, Chao-Hsien Chu, Tracy Mullen, and David M Pennock, "Information markets vs. opinion pools: An empirical comparison," in "Proceedings of the 6th ACM conference on Electronic commerce" ACM 2005, pp. 58–67.

Chetty, Raj, "The Transformative Potential of Administrative Data for Microeconometric Research," 2012. Presentation, NBER Summer Institute.

_ **, John Friedman, and Jonah Rockoff**, "The Long-Term Impacts of Teachers: Teacher Value-added and Student Outcomes in Adulthood," 2011. Mimeo.

Chiappori, Pierre-Andre and Bernard Salanie, "Testing Contract Theory: A Survey of Some Recent Work," in L. Hansen M. Dewatripont and S. Turnovsky, eds., *Advances in Economics and Econometrics*, 2003.

Ching, Andrew, Tulin Erdem, and Michael Keane, "Learning Models: An Assessment of Progress, Challenges and New Developments," 2011. Mimeo.

Choi, Hyunyoung and Hal Varian, "Predicting the present with google trends," *Economic Record*, 2012, *88* (s1), 2–9.

Cingano, Federico and Alfonso Rosolia, "People I Know: Job Search and Social Networks," *Journal of Labor Economics*, 2012, *30* (2), 291 – 332.

Clark, Derek J and Kai A Konrad, "Contests with Multi-tasking*," *The Scandinavian Journal of Economics*, 2007, *109* (2), 303–319.

Clark, Jeremy and Lana Friesen, "Overconfidence in Forecasts of Own Performance: An Experimental Study," *Economic Journal*, 2009, *119* (534), 229–251.

Coase, Ronald H, "The nature of the firm," *economica*, 1937, *4* (16), 386–405.

Coles, Peter A, Karim R Lakhani, and Andrew McAfee, *Prediction markets at Google*, Harvard Business School, 2007.

Commercial Carrier Journal, "The CCJ Top 250," 2005.

Compte, Olivier and Andrew Postlewaite, "Confidence-enhanced performance," *American Economic Review*, 2004, pp. 1536–1557.

Conlin, Michael, Ted O'Donoghue, and Timothy J. Vogelsang, "Projection Bias in Catalog Orders," *American Economic Review*, 2007, *97* (4), 1217–1249.

Corchón, Luis and Matthias Dahm, "Foundations for contest success functions," *Economic Theory*, 2010, *43* (1), 81–98.

Coscelli, Andrea and Matthew Shum, "An Empirical Model of Learning and Patient Spillovers in New Drug Entry," *Journal of Econometrics*, 2004, *122* (2), 213 – 246.

Cowgill, B, J Wolfers, and E Zitzewitz, "Using Prediction Markets to Track Information Flows: Evidence from Google," *Dartmouth College*, 2009.

Cowgill, Bo and Eric Zitzewitz, "Mood Swings at Work: Stock Price Movements, Effort and Decision Making," 2008.

_ **and** _ , "Incentive Effects of Equity Compensation: Employee Level Evidence from Google," *Dartmouth Department of Economics working paper*, 2009.

_ **and** _ , "Corporate Prediction Markets: Evidence from Google, Ford, and Firm X1," 2013.

_ , **Justin Wolfers, and Eric Zitzewitz**, "Using prediction markets to track information flows: Evidence from Google," *Dartmouth College*, 2008.

Cox, David, "Regression models and life-tables," *Journal of the Royal Statistics Society, Series B*, 1972, *34*, 187–220.

Crawford, Gregory S. and Matthew Shum, "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 2005, *73* (4), 1137–1173.

Crawford, Vincent P and Juanjuan Meng, "New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income," *American Economic Review*, 2011, *101* (5), 1912–32.

Cringely, Robert, "Triumph of the Nerds–The Television Program Transcripts," *PBS Online*, 1996.

Dal Bo, Ernesto, Frederico Finan, and Martn A. Rossi, "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service," *Quarterly Journal of Economics*, 2013, *128* (3), 1169–1218.

Dalton, Dan R. and William D. Todor, "Turnover Turned over: An Expanded and Positive Perspective," *The Academy of Management Review*, 1979, *4* (2), pp. 225–235.

D'Amour, Alexander, David Doolin, Guan-Cheng Li, Ye Sun, Vetle Torvik, Amy Yu, and Lee Fleming, "Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975-2010)," 2013. Mimeo.

Datcher, Linda, "The Impact of Informal Networks on Quit Behavior," *Review of Economics and Statistics*, 1983, *65* (3), 491–95.

Daula, Thomas and Robert Moffitt, "Estimating Dynamic Models of Quit Behavior: The Case of Military Reenlistment," *Journal of Labor Economics*, 1995, *13* (3), 499–523.

Davis, Steven J., R. Jason Faberman, and John C. Haltiwanger, "The Establishment-Level Behavior of Vacancies and Hiring," *Quarterly Journal of Economics*, 2013, *128* (2), 581–622.

De Bondt, Werner and Richard Thaler, "Financial Decision-making in Markets and Firms: a Behavioral Perspective," in R. A. Jarrow, V. Maksimovic, and W. T. Ziemba, eds., *Handbooks in Operations Research and Management Science*, North Holland, 1995, pp. 385–410.

Decker, Phillip J. and Edwin T. Cornelius, "A note on recruiting sources and job survival rates," *Journal of Applied Psychology*, 1979, *64* (4), 463–464.

DeGroot, Morris, *Optimal Statistical Decisions*, New York, NY: Mcgraw-Hill College, 1970.

Delavande, Adeline, "Pill, Patch, Or Shot? Subjective Expectations And Birth Control Choice," *International Economic Review*, 2008, *49* (3), 999–1042.

DellaVigna, Stefano and M. Daniele Paserman, "Job Search and Impatience," *Journal of Labor Economics*, July 2005, *23* (3), 527–588.

_ **and Ulrike Malmendier**, "Contract Design and Self-control: Theory and Evidence," *The Quarterly Journal of Economics*, 2004, *119* (2), 353–402.

_ **and** _ , "Paying Not to Go to the Gym," *American Economic Review*, June 2006, *96* (3), 694–719.

_ **, John List, and Ulrike Malmendier**, "Testing for Altruism and Social Pressure in Charitable Giving," *Quarterly Journal of Eocnomics*, 2012, *Forthcoming*.

Demougin, Dominique and Aloysius Siow, "Careers in ongoing hierarchies," *The American Economic Review*, 1994, pp. 1261–1277.

Dey, Matthew S. and Christopher J. Flinn, "An Equilibrium Model of Health Insurance Provision and Wage Determination," *Econometrica*, 2005, *73* (2), 571–627.

Dickstein, Michael, "Efficient Provision of Experience Goods: Evidence from Antidepressant Choice," 2011. Mimeo.
Does the size of the legislature affect the size of government? Evidence from two natural experiments

***Does the size of the legislature affect the size of government? Evidence from two natural experiments**, Journal of Public Economics, 2012, 96 (34), 269 – 278.*

Dominitz, Jeff, *"Using Expectations Data To Study Subjective Income Expectations," Review of Economics and Statistics, 1998, 80 (3), 374–388.*

_ **and Charles F. Manski**, *"Using Expectations Data To Study Subjective Income Expectations," Journal of the American Statistical Association, 1997, 92 (439), 855–867.*

Drago, Robert and Gerald T Garvey, *"Incentives for helping on the job: Theory and evidence," Journal of Labor Economics, 1998, 16 (1), 1–25.*

Duflo, Esther, Rema Hanna, and Stephen Ryan, *"Incentives Work: Getting Teachers to Come to School," American Economic Review, 2012, 102 (4), 1241–78.*

Dupas, Pascaline, *"Relative Risks and the Market for Sex: Teenagers, Sugar Daddies and HIV in Kenya," 2008. Mimeo.*

Dustmann, Christian, Albrecht Glitz, and Uta Schoenberg, *"Referral-based Job Search Networks," working paper, 2012.*

_ **and Uta Schoenberg**, *"Training and Union Wages," Review of Economics and Statistics, 2009, 91 (2), 363–376.*

_ **and Uta Schonberg**, *"What Makes Firm-Based Vocational Training Schemes Successful? The Role of Commitment," American Economic Journal: Applied Economics, 2012, 4 (2), 36–61.*

Dye, Ronald A, *"The trouble with tournaments," Economic Inquiry, 1984, 22 (1), 147–149.*

Eckstein, Zvi and Kenneth I. Wolpin, *"Dynamic Labour Force Participation of Married Women and Endogenous Work Experience," Review of Economic Studies, 1989, 56 (3), pp. 375–390.*

Ederer, Florian and Gustavo Manso, *"Is Pay for Performance Detrimental to Innovation?," Management Science, 2013, 59 (7), 1496–1513.*

Edwards, Ward, *"Conservatism in Human Information Processing," in B. Kleinmuntz, ed., Formal Representation of Human Judgement, 1968.*

Efron, Bradley, *"Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve," Journal of the American Statistical Association, 1988, 83 (402), 414–425.*

Ehrenberg, Ronald G and Michael L Bognanno, *"Do tournaments have incentive effects?," 1988.*

_ **and** _ , *"DO TOURNAMENTS HAVE INCENTIVE EFFECTS?," Journal of Political Economy, 1990, 98, 1307–1324.*

_ **and** _ , *"The incentive effects of tournaments revisited: Evidence from the European PGA tour," Industrial and Labor Relations Review, 1990, pp. 74S–88S.*

Eichenwald, Kurt, *Conspiracy of fools: A true story, Random House LLC, 2005.*

_ , *"Microsoft's Lost Decade," Vanity Fair, 2012.*

Eil, David and Justin Rao, *"The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself," American Economic Journal: Microeconomics, 2011, Forthcoming.*

Einav, Liran, Mark Jenkins, and Jonathan Levin, *"Contract Pricing in Consumer Credit Markets," 2009. Mimeo.*

_ , _ , and _ , *"Contract Pricing in Consumer Credit Markets," 2009. Mimeo.*

Eisenberg, Melvin Aron, *"The Limits of Cognition and the Limits of Contract," Stanford Law Review, 1995, 47 (2), pp. 211–259.*

Elbaum, Bernard, *"Why Apprenticeship Persisted in Britain But Not in the United States," Journal of Economic History, 1989, 49 (2), 337–349.*

Eliaz, Kfir and Ran Spiegler, *"Contracting with Diversely Naive Agents," Review of Economic Studies, 2006, 73 (3), 689–714.*

Eliazar, Iddo I and Igor M Sokolov, *"Diversity of Poissonian populations," Physical Review E, 2010, 81 (1), 011122.*

Elliott, James R, *"Social Isolation and Labor Market Insulation," The Sociological Quarterly, 1999, 40 (2), 199–216.*

Ellison, Glenn, *Bounded Rationality in Industrial Organization, Cambridge University Press, 2006.*

_ and Alexander Wolitzky, *"A search cost model of obfuscation," The RAND Journal of Economics, 2012, 43 (3), 417–441.*

_ and Sara Fisher Ellison, *"Search, obfuscation, and price elasticities on the internet," Econometrica, 2009, 77 (2), 427–452.*

Erdem, Tulin, Michael Keane, Sabri Oncu, and Judi Strebel, *"Learning About Computers: An Analysis of Information Search and Technology Choice," Quantitative Marketing and Economics, 2005, 3 (3), 207–247.*

Ericson, Keith, *"Market Design when Firms Interact with Inertial Consumers: Evidence from Medicare Part D," 2010. Mimeo.*

_ , *"Forgetting We Forget: Overconfidence and Prospective Memory," Journal of the European Economic Association, 2011, Forthcoming.*

Fafchamps, Marcel and Alexander Moradi, *"Referral and Job Performance: Evidence from the Ghana Colonial Army," 2011. Mimeo.*

Fair, Ray C and Robert J Shiller, *"The Informational Content of Ex Ante Forecasts," The Review of Economics and Statistics, 1989, pp. 325–331.*

_ **and** _ , *"Comparing information in forecasts from econometric models," The American Economic Review, 1990, pp. 375–389.*

Fairburn, James A and James M Malcomson, *"Performance, promotion, and the Peter Principle," The Review of Economic Studies, 2001, 68 (1), 45–66.*

Falk, Armin and Andrea Ichino, *"Clean Evidence on Peer Effects," Journal of Labor Economics, 2006, 24 (1), 39–58.*

Fama, Eugene F, *"Agency Problems and the Theory of the Firm," Journal of Political Economy, 1980, 88 (2), 288–307.*

Fang, Hanming and Dan Silverman, *"Time-Inconsistency And Welfare Program Participation: Evidence From The NLSY," International Economic Review, 2009, 50 (4), 1043–1077.*

_ **and Yang Wang**, *"Estimating Dynamic Discrete Choice Models with Hyperbolic Discounting, with an Application to Mammography Decisions," 2010. Mimeo,.*

Farber, Henry S., *"The Analysis of Interfirm Worker Mobility," Journal of Labor Economics, 1994, 12 (4), pp. 554–593.*

_ **and Robert Gibbons**, *"Learning and Wage Dynamics," Quarterly Journal of Economics, 1996, 111 (4), 1007–1047.*

Fernandez, Roberto and Emilio Castilla, *"How Much Is That Network Worth? Social Capital Returns for Referring Prospective Hires," in Nan Lin Karen Cook and Ronald Burt, eds., Social Capital: Theory Research, 2006.*

Fernandez, Roberto M. and Isabel Fernandez-Mateo, *"Networks, Race, and Hiring," American Sociological Review, 2006, 71 (1), pp. 42–71.*

_ **and M. Lourdes Sosa**, *"Gendering the Job: Networks and Recruitment at a Call Center," American Journal of Sociology, 2005, 111 (3), pp. 859–904.*

_ **and Nancy Weinberg**, *"Sifting and Sorting: Personal Contacts and Hiring in a Retail Bank," American Sociological Review, 1997, 62 (6), pp. 883–902.*

_ , **Emilio J. Castilla, and Paul Moore**, *"Social Capital at Work: Networks and Employment at a Phone Center," American Journal of Sociology, 2000, 105 (5), pp. 1288–1356.*

Festinger, Leon, *A Theory of Cognitive Dissonance, Stanford, CA: Stanford University Press, 1957.*

Figlewski, Stephen, *"Subjective information and market efficiency in a betting market," The Journal of Political Economy, 1979, pp. 75–88.*

Finkelstein, Amy and Kathleen McGarry, *"Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market," American Economic Review, 2006, 96 (4), 938–958.*

Fischoff, Baruch, *"Debiasing," in Paul Slovic Daniel Kahneman and Amos Tversky, eds., Judgment under Uncertainty: Heuristics and Biases, University of Chicago Press, 1982, pp. 422–444.*

Flabbi, Luca and Andrea Ichino, *"Productivity, Seniority and Wages: New Evidence from Personnel Data," Labour Economics, 2001, 8 (3), 359–387.*

Fleisig, Dida, *"Adding Information May Increase Overconfidence In Accuracy Of Knowledge Retrieval," Psychological Reports, 2011, 108 (2), 379–392.*

Fogel, Robert W. and Stanley L. Engerman, *Time on the Cross: The Economics of American Negro Slavery, New York, NY: Little, Brown, 1974.*

Frank, Robert H, *Choosing the right pond: Human behavior and the quest for status., Oxford University Press, 1985.*

Frederick, Shane, George Loewenstein, and Ted O'Donoghue, *"Time Discounting and Time Preference: A Critical Review," Journal of Economic Literature, June 2002, 40 (2), 351–401.*

Frederiksen, Anders, Fabian Lange, and Ben Kriechel, *"Subjective Performance Evaluations and Employee Careers," 2014. Mimeo.*

Fullerton, Richard L and R Preston McAfee, *"Auctioning Entry into Tournaments," Journal of Political Economy, 1999, 107 (3), 573–605.*

Galenianos, Manolis, *"Hiring Through Referrals," working paper, 2012.*

_ , *"Learning About Match Quality and the Use of Referrals," Review of Economic Dynamics, 2013, 16 (4), 668–690.*

Gannon, Martin J., *"Sources of referral and employee turnover," Journal of Applied Psychology, 1971, 55 (3), 226–228.*

Garicano, Luis, *"Hierarchies and the Organization of Knowledge in Production," Journal of Political Economy, 2000, 108 (5), 874–904.*

_ **and Ignacio Palacios-Huerta**, *Sabotage in tournaments: Making the beautiful game a bit less beautiful, Centre for Economic Policy Research, 2005.*

Genicot, Garance, *"Bonded Labor and Serfdom: a Paradox of Voluntary Choice," Journal of Development Economics, 2002, 67 (1), 101 – 127.*

Gibbons, Robert and Lawrence Katz, *"Layoffs and Lemons," Journal of Labor Economics, 1991, 9 (4), 351–80.*

_ **and Michael Waldman**, *"A theory of wage and promotion dynamics inside firms," The Quarterly Journal of Economics, 1999, 114 (4), 1321–1358.*

_ **and** _ , *"Enriching a theory of wage and promotion dynamics inside firms," Technical Report, National Bureau of Economic Research 2003.*

_ **, Lawrence F. Katz, Thomas Lemieux, and Daniel Parent**, *"Comparative Advantage, Learning, and Sectoral Wage Determination," Journal of Labor Economics, 2005, 23 (4), 681–724.*

Gicheva, Dora, *"Worker Mobility, Employer-Provided Tuition Assistance, and the Choice of Graduate Management Program Type," 2009. Mimeo.*

Gillen, Benjamin J, Charles R Plott, and Matthew Shum, *"Information Aggregation Mechanisms in the Field: Sales Forecasting Inside Intel," 2012.*

Gilleskie, Donna B., *"A Dynamic Stochastic Model of Medical Care Use and Work Absence," Econometrica, 1998, 66 (1), pp. 1–45.*

Gjerstad, Steven and McClelland Hall, *"Risk aversion, beliefs, and prediction market equilibrium," Economic Science Laboratory, University of Arizona, 2005.*

Gladwell, Malcolm, *"The talent myth," The New Yorker, 2002, 22 (2002), 28–33.*

Gneezy, Uri and Aldo Rustichini, *"Pay enough or don't pay at all," The Quarterly Journal of Economics, 2000, 115 (3), 791–810.*

Goel, Anand M and Anjan V Thakor, *"Overconfidence, CEO selection, and corporate governance," The Journal of Finance, 2008, 63 (6), 2737–2784.*

Goel, Sharad, Daniel M Reeves, Duncan J Watts, and David M Pennock, *"Prediction without markets," in "Proceedings of the 11th ACM conference on Electronic commerce" ACM 2010, pp. 357–366.*

Goettler, Ronald and Karen Clay, *"Tariff Choice with Consumer Learning and Switching Costs," Journal of Marketing Research, 2011, 48 (4), 633–652.*

Goldfarb, Avi and Mo Xiao, *"Who Thinks about the Competition? Managerial Ability and Strategic Entry in US Local Telephone Markets," American Economic Review, 2011, 101 (7), 3130–61.*

Goodwin, Liz, *"Law grads sue school, say degree is 'indentured servitude'," Yahoo News, August 2011.*

Gotz, Glenn and John J. McCall, *"A Dynamic Retention Model for Air Force Officers," Technical Report, RAND Corporation 1984.*

Granovetter, Mark, *"The Strength of Weak Ties," American Journal of Sociology, 1973, 78 (6), 1360–1380.*

_ , *Getting a Job (Cambridge, MA: Harvard University Press, 1974).*

Green, Gary Paul, Leann M Tigges, and Daniel Diaz, *"Racial and Ethnic Differences in Job-search Strategies in Atlanta, Boston, and Los Angeles," Social Science Quarterly, 1999.*

Green, Jerry R and Nancy L Stokey, *"A Comparison of Tournaments and Contracts," Journal of Political Economy, 1983, 91 (3), 349–64.*

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, *"Identifying Agglomeration Spillovers: Evidence from Million Dollar Plants," NBER Working Papers 13833, National Bureau of Economic Research, Inc March 2008.*

Greenwald, Bruce C, *"Adverse Selection in the Labour Market," Review of Economic Studies, July 1986, 53 (3), 325–47.*

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano, *"Policy Responses to Fiscal Restraints: A Difference-in-Discontinuities Design," working paper, 2012.*

Grisley, William and Earl D. Kellogg, *"Farmers' Subjective Probabilities in Northern Thailand: An Elicitation Analysis," American Journal of Agricultural Economics, 1983, 65 (1), pp. 74–82.*

Grossman, Zachary and David Owens, *"An Unlucky Feeling: Overconfidence and Noisy Feedback," 2010. Mimeo.*

Grubb, Michael, *"Selling to Overconfident Consumers," American Economic Review, 2009, 99 (5), 1770–1807.*

_ and Matthew Osborne, *"Cellular Service Demand: Tariff Choice, Usage Uncertainty, Biased Beliefs, and Learning," 2011. Mimeo.*

Gruber, Jonathan and Brigitte C. Madrian, *"Health Insurance, Labor Supply, and Job Mobility: A Critical Review of the Literature," 2002. NBER Working Paper.*

Gubler, Timothy, Ian Larkin, and Lamar Pierce, *"The dirty laundry of employee award programs: Evidence from the field," Harvard Business School NOM Unit Working Paper, 2013, (13-069).*

Gürtler, Marc and Oliver Gürtler, *"The optimality of heterogeneous tournaments," Technical Report, Working Papers, Institut für Finanzwirtschaft, TU Braunschweig 2013.*

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg, *"Market Value and Patent Citations," RAND Journal of Economics, 2005, 36 (1), 16–38.*

Hamermesh, Daniel S, Harley Frazis, and Jay Stewart, *"Data watch: The American time use survey," Journal of Economic Perspectives, 2005, pp. 221–232.*

Han, Jing and Jian Han, *"Network-based recruiting and applicant attraction in China: insights from both organizational and individual perspectives," The International Journal of Human Resource Management, 2009, 20 (11), 2228–2249.*

Hankins, Richard A and Alison Lee, *"Crowd Sourcing and Prediction Markets," 2011.*

Hanson, Robin, *"Combinatorial information market design," Information Systems Frontiers, 2003, 5 (1), 107–119.*

Hanson, Robin D, *"Decision markets," Entrepreneurial Economics: Bright Ideas from the Dismal Science, 2002, p. 79.*

Haran, Uriel, Don A. Moore, and Carey K. Morewedge, *"A simple remedy for overprecision in judgment," Judgment and Decision Making, 2010, 5 (7), 467–476.*

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

Harris, Milton and Bengt Holmstrom, *"A Theory of Wage Dynamics," Review of Economic Studies, 1982, 49 (3), 315–333.*

Heath, Rachel, *"Why do Firms Hire using Referrals? Evidence from Bangladeshi Garment Factories," mimeo, 2013.*

Heckman, James J. and Burton Singer, *"A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," Econometrica, 1984, 52 (2), 271–320.*

_ , Robert LaLonde, and Jeffrey Smith, *"Economics and Econometrics of Active Labor Market Programs," Handbook of Labor Economics, 1999, 31, 1866–2097.*

Hellerstein, Judith K., Melissa McInerney, and David Neumark, *"Neighbors and Coworkers: The Importance of Residential Labor Market Networks," Journal of Labor Economics, 2011, 29 (4), 659 – 695.*

Henderson, Rebecca and Iain Cockburn, *"Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery," RAND Journal of Economics, 1996, 27 (1), 32–59.*

Hendren, Nathaniel, *"Private Information and Insurance Rejections," 2012. Mimeo.*

Hensvik, Lena and Oskar Nordstrom Skans, *"Social Networks, Employee Selection and Labor Market Outcomes: Toward an Empirical Analysis," IFAU Working Paper 2013:15, 2013.*

Herfindahl, Orris Clemens, *"Concentration in the steel industry." PhD dissertation, Columbia University. 1950.*

Hermalin, Benjamin E., *"Toward an Economic Theory of Leadership: Leading by Example," The American Economic Review, 1998, 88 (5), pp. 1188–1206.*

Hirschman, Albert, *"National Power and the Structure of International Trade," 1945.*

Hirschman, Albert O, *"The paternity of an index," The American Economic Review, 1964, pp. 761–762.*

_ , *National power and the structure of foreign trade, Vol. 105, Univ of California Press, 1980.*

Hoffman, Mitchell, *"Overconfidence at Work and the Evolution of Beliefs: Evidence from a Field Experiment," 2011. Mimeo, UC Berkeley.*

_ , *"Training Contracts, Worker Overconfidence, and the Provision of Firm-Sponsored General Training," Technical Report, Working paper, Univ. California, Berkeley 2011.*

_ , *"How is Information (Under-) Valued? Evidence from Framed Field Experiments," 2012. Mimeo.*

_ **and John Morgan**, *"Naughty or Nice? Social preferences in Online Industries," 2009. Mimeo, UC Berkeley.*

_ **and Stephen Burks**, *"Training Contracts, Worker Overconfidence, and the Provision of Firm-Sponsored General Training," 2012. Mimeo, Yale University.*

_ **and** _ , *"Training Contracts, Worker Overconfidence, and the Provision of Firm-Sponsored General Training," working paper, 2014.*

Holmström, Bengt, *"Managerial incentive problems: A dynamic perspective," The Review of Economic Studies, 1999, 66 (1), 169–182.*

Holmstrom, Bengt and Paul Milgrom, *"Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design," JL Econ. & Org., 1991, 7, 24.*

_ **and** _ , *"Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," Journal of Law, Economics and Organization, Special I 1991, 7 (0), 24–52.*

Holt, Charles A. and Angela M. Smith, *"An Update on Bayesian Updating," Journal of Economic Behavior and Organization, 2009, 69 (2), 125 – 134.*

Holt, Charles and Susan Laury, *"Risk Aversion and Incentive Effects," American Economic Review, 2002, 92 (5), 1644–1655.*

Holzer, Harry J., *"Job Search By Employed and Unemployed Youth," Industrial and Labor Relations Review, 1986, 40, 601.*

_ , *"Hiring Procedures in the Firm: Their Economic Determinants and Outcomes," NBER Working Paper No. 2185, 1987a.*

_ , *"Informal Job Search and Black Youth Unemployment," American Economic Review, 1987b, 77 (3), pp. 446–452.*

Hossain, Tanjim and Ryo Okui, *"The Binarized Scoring Rule of Belief Elicitation," 2010. Mimeo.*

Hotelling, Harold, *"Stability in Competition," The Economic Journal, 1929, 39 (153), 41–57.*

Hotz, V Joseph and Robert A Miller, *"Conditional Choice Probabilities and the Estimation of Dynamic Models," Review of Economic Studies, 1993, 60 (3), 497–529.*

Hubbard, Thomas N., *"Information, Decisions, and Productivity: On-Board Computers and Capacity Utilization in Trucking," American Economic Review, 2003, 93 (4), 1328–1353.*

Huck, Steffen, Imran Rasul, and Andrew Shephard, *"Comparing Charitable Fundraising Schemes: Evidence from a Field Experiment and a Structural Model," 2012. Mimeo.*

Hurd, Michael D. and Kathleen McGarry, *"The Predictive Validity of Subjective Probabilities of Survival," Economic Journal, 2002, 112 (482), 966–985.*

Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi, *"The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines," American Economic Review, 1997, 87 (3), 291–313.*

Imbens, Guido W. and Thomas Lemieux, *"Regression discontinuity designs: A guide to practice," Journal of Econometrics, 2008, 142 (2), 615–635.*

Ioannides, Yanis and Linda D. Loury, *"Job Information Networks, Neighborhood Effects, and Inequality," Journal of Economic Literature, 2004, 42 (4), 1056–1093.*

Ippolito, Richard A., *"Stayers as "Workers" and "Savers": Toward Reconciling the Pension-Quit Literature," Journal of Human Resources, 2002, 37 (2), 275–308.*

Jackson, Matthew, *Social and Economic Networks (Princeton, NJ: Princeton University Press, 2008).*

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson, *"Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," Quarterly Journal of Economics, 1993, 108 (3), 577–598.*

James, Jon, *"Ability Matching and Occupational Choice," 2011. Mimeo.*

Jenkins, Mark, *"Measuring the Costs of Ex Post Moral Hazard in Secured Lending," 2010. Mimeo.*

Jensen, Robert, *"The Perceived Returns to Education and the Demand for Schooling," Quarterly Journal of Economics, 2010, pp. 515–548.*

Jia, Hao, Stergios Skaperdas, and Samarth Vaidya, *"Contest functions: Theoretical foundations and issues in estimation," International Journal of Industrial Organization, 2013, 31 (3), 211–222.*

Jin, Ginger and Philip Leslie, *"The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," Quarterly Journal of Economics, 2004, 118 (2), 409–451.*

John, O. P., E. M. Donahue, and R. L. Kentle, *"The Big Five Inventory–Versions 4a and 54," 1991. Berkeley, CA: University of California,Berkeley, Institute of Personality and Social Research.*

John, OP, EM Donahue, and RL Kentle, *"The big five inventory: Versions 4a and 54, Institute of personality and social research," University of California, Berkeley, CA, 1991.*

Johnson, Erin M., *"Ability, Learning and the Career Path of Cardiac Specialists," 2009. Mimeo.*

Jolls, Christine and Cass R. Sunstein, *"Debiasing through Law," Journal of Legal Studies, 2006, 35, 199–242.*

Jovanovic, Boyan, *"Job Matching and the Theory of Turnover," Journal of Political Economy, 1979, 87 (5), 972–90.*

_ **and Yaw Nyarko**, *"The Transfer of Human Capital," Journal of Economic Dynamics and Control, 1995, 19 (5-7), 1033–1064.*

_ **and** _ , *"Learning by Doing and the Choice of Technology," Econometrica, 1996, 64 (6), 1299–1310.*

Judd, Kenneth, *Numerical Methods in Economics, Cambridge, Massachusetts: MIT Press, 1998.*

Kahn, Lisa and Fabian Lange, *"Employer Learning, Productivity and the Earnings Distribution: Evidence from Performance Measures," 2013. Mimeo.*

Kahn, Lisa B and Fabian Lange, *"Employer learning, productivity and the earnings distribution: Evidence from performance measures," Technical Report, Discussion paper series//Forschungsinstitut zur Zukunft der Arbeit 2010.*

_ **and** _ , *"Employer Learning, Productivity and the Earnings Distribution: Evidence from Performance Measures," Review of Economic Studies (forthcoming), 2014.*

Kahneman, Daniel and Amos Tversky, *"Choices, values, and frames.," American psychologist, 1984, 39 (4), 341.*

Kaniel, Ron, Cade Massey, and David Robinson, *"The Importance of Being an Optimist: Evidence from Labor Markets," 2011. Mimeo.*

Kaplan, Greg, *"Moving Back Home: Insurance Against Labor Market Risk," 2010. Mimeo.*

Keane, Michael P. and Kenneth I. Wolpin, *"The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," Review of Economics and Statistics, 1994, 76 (4), 648–72.*

_ **and** _ , *"The Career Decisions of Young Men," Journal of Political Economy, 1997, 105 (3), 473–522.*

_ **and** _ , *"The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment," International Economic Review, 2001, 42 (4), pp. 1051–1103.*

Kennan, John, *"Bonding and the Enforcement of Labor Contracts," Economics Letters, 1979, 3 (1), 61–66.*

Keynes, John Maynard, *The general theory of Employment, Interest and Money, Harcourt Brace and Co., 1936.*

Khwaja, Ahmed, Frank Sloan, and Sukyung Chung, *"The Relationship Between Individual Expectations and Behaviors: Mortality Expectations and Smoking Decisions," Journal of Risk and Uncertainty, 2007, 35, 179–201.*

Kirnan, Jean Powell, John A. Farley, and Kurt F. Geisinger, *"The Relationship Between Recruiting Source, Applicant Quality, and Hire Performance: An Analysis by Sex, Ethnicity, and Age," Personnel Psychology, 1989, 42 (2), 293–308.*

Kling, Jeffrey R., Sendhil Mullainathan, Eldar Shafir, Lee C. Vermeulen, and Marian V. Wrobel, *"Comparison Friction: Experimental Evidence from Medicare Drug Plans," Quarterly Journal of Economics, 2012, 127 (1), 199–235.*

Knoeber, Charles R, *"A real game of chicken: contracts, tournaments, and the production of broilers," Journal of Law, Economics, & Organization, 1989, pp. 271–292.*

_ **and Walter N Thurman**, *"Testing the theory of tournaments: An empirical analysis of broiler production," Journal of Labor Economics, 1994, pp. 155–179.*

Kohn, A, *"Punished by rewards: the trouble with gold stars, incentive plans, A's, praise, and other bribes," 1993.*

Korenman, Sanders and Susan Turner, *"Employment Contacts and Minority-White Wage Differences," Industrial Relations, 1996, 35 (1), 106–122.*

Koszegi, Botond, *"Health anxiety and patient behavior," Journal of Health Economics, 2003, 22 (6), 1073 – 1084.*

Köszegi, Botond, *"Ego utility, overconfidence, and task choice," Journal of the European Economic Association, 2006, 4 (4), 673–707.*

Krackhardt, David and Lyman W. Porter, *"When Friends Leave: A Structural Analysis of the Relationship between Turnover and Stayers Attitudes," Administrative Science Quarterly, 1985, 30 (2), 24261.*

Kraemer, Carlo, Markus Noth, and Martin Weber, *"Information Aggregation with Costly Information and Random Ordering: Experimental Evidence," Journal of Economic Behavior and Organization, 2006, 529, 432–442.*

Kramarz, Francis and Oskar Nordstrom Skans, *"When Strong Ties are Strong: Networks and Youth Labour Market Entry," Review of Economic Studies, forthcoming, 2014.*

Krasnokutskaya, Elena, Kyungchul Song, and Xun Tang, *"The Role of Quality in Service Markets Organized as Multi-Attribute Auctions," Technical Report, Citeseer 2013.*

Kraus, Anthony W., *"Repayment Agreements for Employee Training Costs," Labor Law Journal, 1993, 44 (1), 49–55.*

_ , *"Employee Agreements for Repayment of Training Costs: The Emerging Case Law," Labor Law Journal, 2008, pp. 213–226.*

Kronman, Anthony T., *"Paternalism and the Law of Contracts," The Yale Law Journal, 1983, 92 (5), pp. 763–798.*

Kubler, Dorothea and Georg Weizsacker, *"Limited Depth of Reasoning and Failure of Cascade Formation in the Laboratory.," Review of Economic Studies, 2004, 71, 425–441.*

Kugler, Adriana D., *"Employee Referrals and Efficiency Wages," Labour Economics, 2003, 10 (5), 531–556.*

Kyle, Albert S, *"Informed speculation with imperfect competition," The Review of Economic Studies, 1989, 56 (3), 317–355.*

LaComb, Christina Ann, Janet Arlie Barnett, and Qimei Pan, *"The imagination market," Information Systems Frontiers, 2007, 9 (2), 245–256.*

Lai, Ronald, Alexander D'Amour, Amy Yu, Ye Sun, and Lee Fleming, *"Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975 - 2010)," 2010. http://hdl.handle.net/1902.1/15705 UNF:5:RqsI3LsQEYLHkkg5jG/jRg== V4 [Version].*

Laibson, David, *"Golden Eggs and Hyperbolic Discounting," Quarterly Journal of Economics, 1997, 112 (2), pp. 443–477.*

_ **, Andrea Repetto, and Jeremy Tobacman**, *"Estimating Discount Functions with Consumption Choices over the Lifecycle," 2007. Mimeo.*

Lange, Fabian, *"The Speed of Employer Learning," Journal of Labor Economics, 2007, 25, 1–35.*

Larkin, Ian and Stephen Leider, *"Incentive Schemes, Sorting, and Behavioral Biases of Employees: Experimental Evidence," American Economic Journal: Microeconomics, 2012, 4 (2), 184–214.*

_ **and** _ , *"Incentive Schemes, Sorting and Behavorial Biases of Employees: Experimental Evidence," American Economic Journal: Microeconomics, 2012, 4 (2), 184–214.*

Laschever, Ron, *"The Doughboys Network: Social Interactions and the Employment of World War I Veterans.," 2009. Mimeo.*

Lau, Annie Y.S. and Enrico W. Coiera, *"Can Cognitive Biases during Consumer Health Information Searches Be Reduced to Improve Decision Making?," Journal of the American Medical Informatics Association, 2009, 16 (1), 54 – 65.*

Lazear, Edward, *"Why Is There Mandatory Retirement?," Journal of Political Economy, 1979, 87 (6), 1261–1284.*

_ , *"Job Security Provisions and Employment," Quarterly Journal of Economics, 1990, 105 (3), 699–726.*

_ , *"Performance Pay and Productivity," American Economic Review, 2000, 90 (5), 1346–1361.*

_ , *"Firm-Specific Human Capital: A Skill-Weights Approach," Journal of Political Economy, 2009, 117 (5), 914–940.*

_ **, Kathryn Shaw, and Christopher Stanton**, *"The Value of Bosses," 2012. Mimeo.*

Lazear, Edward P, *"Pay equality and industrial politics," Journal of Political Economy, 1989, pp. 561–580.*

Lazear, Edward P., *"Leadership: A personnel economics approach," Labour Economics, 2012, 19 (1), 92–101.*

Lazear, Edward P and Paul Oyer, *"Personnel economics," Technical Report, National Bureau of Economic Research 2007.*

_ **and Sherwin Rosen**, *"Rank-Order Tournaments as Optimum Labor Contracts," The Journal of Political Economy, 1981, 89 (5), 841–864.*

Lee, David S. and Thomas Lemieux, *"Regression Discontinuity Designs in Economics," Journal of Economic Literature, 2010, 48 (2), 281–355.*

Lee, Robert, *"Social capital and business and management: Setting a research agenda," International Journal of Management Reviews, 2009, 11 (3), 247–273.*

Leigh, Andrew, Justin Wolfers, and Eric Zitzewitz, *"Is there a favorite-longshot bias in election markets," in "Preliminary version presented at the 2007 UC Riverside Conference on Prediction Markets, Riverside, CA" 2007.*

Lemieux, Thomas, W. Bentley MacLeod, and Daniel Parent, *"Performance Pay and Wage Inequality," Quarterly Journal of Economics, 2009, 124 (1), 1–49.*

Lerner, Josh, Parag A Pathak, and Jean Tirole, *"The dynamics of open-source contributors," The American Economic Review, 2006, pp. 114–118.*

Levitt, Steven D., *"An Economist Sells Bagels: A Case Study in Profit Maximization," Working Paper 12152, National Bureau of Economic Research April 2006.*

_ **and John A. List**, *"What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?," Journal of Economic Perspectives, 2007, 21 (2), 153–174.*

Levy, Steven, *In the plex: How Google thinks, works, and shapes our lives, Simon and Schuster, 2011.*

Lewis, Michael, *Liar's poker: Rising through the wreckage on Wall Street, WW Norton & Company, 1989.*

Lewis, Randall A and Justin M Rao, *"On the near impossibility of measuring advertising effectiveness," Technical Report, Working paper 2012.*

Li, Danielle, *"Information, Bias, and Efficiency in Expert Evaluation: Evidence from the NIH," Job market paper, 2012, pp. 1–57.*

List, John A., *"Young, Selfish and Male: Field evidence of social preferences," Economic Journal, 2004, 114 (492), 121–149.*

Lohr, Steve, *"Betting to Improve the Odds," New York Times, April, 2008, 9.*

Loury, Linda Datcher, *"Some Contacts Are More Equal than Others: Informal Networks, Job Tenure, and Wages," Journal of Labor Economics, 2006, 24 (2), 299–318.*

Lunn, Mary and Don McNeil, *"Applying Cox Regression to Competing Risks," Biometrics, 1995, 51 (2), 524–532.*

Lynch, Lisa M, *"The Economics of Youth Training in the United States," Economic Journal, 1993, 103 (420), 1292–302.*

Lynch, Lisa M. and Sandra E. Black, *"Beyond the Incidence of Employer-Provided Training," Industrial and Labor Relations Review, 1998, 52 (1), pp. 64–81.*

Macan, Therese, *"The employment interview: A review of current studies and directions for future research," Human Resource Management Review, 2009, 19 (3), 203–218.*

MacDonald, Glenn and Leslie M Marx, *"Adverse specialization," Journal of Political Economy, 2001, 109 (4), 864–899.*

Madrian, Brigitte C., *"Employment-Based Health Insurance and Job Mobility: Is There Evidence of Job-Lock?," Quarterly Journal of Economics, 1994, 109 (1), 27–54.*

Malmendier, Ulrike and Geoffrey Tate, *"CEO overconfidence and corporate investment," The Journal of Finance, 2005, 60 (6), 2661–2700.*

_ **and** _ , *"CEO Overconfidence and Corporate Investment," Journal of Finance, December 2005, 60 (6), 2661–2700.*

_ **and** _ , *"Who makes acquisitions? CEO overconfidence and the market's reaction," Journal of Financial Economics, 2008, 89 (1), 20–43.*

_ **and** _ , *"Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction," Journal of Financial Economics, 2008, 89 (1), 20 – 43.*

Manchester, Colleen, *"Investment in General Human Capital and Turnover Intention," American Economic Review Papers & Proceedings, 2010, 100 (2), 209–13.*

_ , *"How Does General Training Increase Retention? Examination Using Tuition Reimbursement Programs," Industrial and Labor Relations Review, 2012, 65 (4).*

Manski, Charles F., *"Measuring Expectations," Econometrica, 2004, 72 (5), 1329–1376.*

_ , *"Interpreting the predictions of prediction markets," Economics Letters, 2006, 91 (3), 425 – 429.*

Marinovic, Iván, Marco Ottaviani, and Peter Norman Sørenson, *"Modeling Idea Markets: Between Beauty Contests and Prediction Markets," Leighton Vaughn Williams. Routledge, 2011.*

Marmaros, David and Bruce Sacerdote, *"Peer and social networks in job search," European Economic Review, 2002, 46 (45), 870 – 879.*

Marsden, Peter V. and Elizabeth H. Gorman, *"Social Networks, Job Changes, and Recruitment," in Ivar Berg and Arne L. Kalleberg, eds., Sourcebook of Labor Markets, Plenum Studies in Work and Industry, Springer US, 2001, pp. 467–502.*

Marx, Matt, Deborah Strumsky, and Lee Fleming, *"Noncompetes and Inventor Mobility: Specialists, Stars, and the Michigan Experiment," 2009. Mimeo.*

Maskin, Eric and Jean Tirole, *"Markov Perfect Equilibrium: I. Observable Actions," Journal of Economic Theory, 2001, 100 (2), 191–219.*

Maslow, Abraham Harold, *"A theory of human motivation.," Psychological review, 1943, 50 (4), 370.*

Massey, Cade, Joseph P. Simmons, and David A. Armor, *"Hope Over Experience," Psychological Science, 2011, 22 (2), 274–281.*

Mattock, Michael and Jeremy Arkes, *"The Dynamic Retention Model for Air Force Officers: New Estimates and Policy Simulations of the Aviator Continuation Pay Program," RAND, Technical Report, 2007.*

Mayraz, Guy, *"Priors and Desires: a Model of Payoff Dependent Beliefs," 2011. Mimeo.*

_ , *"Wishful Thinking," 2011. Mimeo.*

McCall, Brian, *"Occupational Matching: A Test of Sorts," Journal of Political Economy, 1990, 98 (1), 45–69.*

McClelland, David C, John W Atkinson, Russell A Clark, and Edgar L Lowell, *"The achievement motive.," 1953.*

McCue, Kristin, *"Promotions and wage growth," Journal of Labor Economics, 1996, pp. 175–209.*

McCullers, John C, *"Issues in learning and motivation," The hidden costs of reward, 1978, pp. 5–18.*

McDonald, Paul, Matt Mohebbi, and Brett Slatkin, *"Comparing google consumer surveys to existing probability and non-probability based internet surveys," Google Whitepaper, Retrieved from http://www. google. com/insights/consumersurveys/static/consumer_surveys_whitepaper. pdf, 2012.*

McEvily, Bill and Marco Tortoriello, *"Measuring trust in organisational research: Review and recommendations," Journal of Trust Research, 2011, 1 (1), 23–63.*

McGraw, Kenneth O, *"The detrimental effects of reward on performance: A literature review and a prediction model," The hidden costs of reward, 1978, pp. 33–60.*

McGreevy, Patrick, *"LAPD Suing Former Officers," Los Angeles Times, March 2006.*

McKelvey, Richard and Talbot Page, *"Public and Private Information: An Experimental Study of Information Pooling," Econometrica, 1990, 58 (6), 1321–1339.*

McKinney, W. Paul, David L. Schiedermayer, Nicole Lurie, Deborah E. Simpson, Jesse L. Goodman, and Eugene C. Rich, *"Attitudes of Internal Medicine Faculty and Residents Toward Professional Interaction With Pharmaceutical Sales Representatives," Journal of the American Medical Association, 1990, 264 (13), 1693–1697.*

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook, *"Birds of a Feather: Homophily in Social Networks," Annual Review of Sociology, 2001, 27, 415–444.*

Medoff, James L. and Katharine G. Abraham, *"Experience, Performance, and Earnings," Quarterly Journal of Economics, 1980, 95 (4), 703–736.*

_ and _, *"Are Those Paid More Really More Productive? The Case of Experience," Journal of Human Resources, 1981, pp. 186–216.*

Meltzer, David and Jeanette W. Chung, *"Coordination, Switching Costs and the Division of Labor in General Medicine: An Economic Explanation for the Emergence of Hospitalists in the United States," 2011. Mimeo.*

Merlo, Antonio and Kenneth I. Wolpin, *"The Transition from School to Jail: Youth Crime and High School Completion Among Black Males," 2008. Mimeo.*

Michaels, Ed, Helen Handfield-Jones, and Beth Axelrod, *The war for talent, Harvard Business Press, 2001.*

Milgrom, Paul and John Roberts, *"An economic approach to influence activities in organizations," American Journal of sociology, 1988, pp. S154–S179.*

_ and Nancy Stokey, *"Information, trade and common knowledge," Journal of Economic Theory, 1982, 26 (1), 17–27.*

Milgrom, Paul R, *"Employment contracts, influence activities, and efficient organization design," The Journal of Political Economy, 1988, pp. 42–60.*

Misra, Sanjog and Harikesh Nair, *"A Structural Model of Sales-Force Compensation Dynamics: Estimation and Field Implementation," 2009. Mimeo, Stanford University.*

Mobius, Markus M., Muriel Niederle, Paul Niehaus, and Tanya Rosenblat, *"Managing Self-Confidence: Theory and Experimental Evidence," 2010. Mimeo.*

Moldovanu, Benny, Aner Sela, and Xianwen Shi, *"Contests for status," Journal of Political Economy, 2007, 115 (2), 338–363.*

Montgomery, James D, *"Social Networks and Labor-Market Outcomes: Toward an Economic Analysis," American Economic Review, December 1991, 81 (5), 1407–18.*

Moore, Don A. and Paul J. Healy, *"The Trouble With Overconfidence," Psychological Review, 2008, 115 (2), 502 – 517.*

Moore, Don and Samuel Swift, *"The Three Faces of Overconfidence in Organizations," in Rolf Van Dick and J Keith Murnighan, eds., Social Psychology and Organizations, 2010.*

Morgan, John, *"Financing public goods by means of lotteries," Review of Economic Studies, 2000, 67 (4), 761–784.*

_ **and Martin Sefton**, *"Funding public goods with lotteries: experimental evidence," The Review of Economic Studies, 2000, 67 (4), 785–810.*

_ **, Dana Sisak, and Felix Vardy**, *"On the merits of meritocracy," Technical Report, Tinbergen Institute Discussion Paper 2012.*

Mortensen, Dale and Tara Vishwanath, *"Personal contacts and earnings: It is who you know!," Labour Economics, 1995, 2 (1), 187–201.*

Mullainathan, Sendhil, Joshua Schwarzstein, and William J. Congdon, *"A Reduced-Form Approach to Behavioral Public Finance," Annual Review of Economics, 2012, 4, 511–540.*

Munier, Bertrand and Costin Zaharia, *"High stakes and acceptance behavior in ultimatum bargaining," Theory and Decision, 2002, 53 (3), 187–207.*

Nagypal, Eva, *"Learning by Doing vs. Learning About Match Quality: Can We Tell Them Apart," Review of Economic Studies, 2007, 74 (1), pp. 537–566.*

Naidu, Suresh and Noam Yuchtman, *"How Green Was My Valley? Coercive Contract Enforcement in 19th Century Industrial Britain," 2009. Mimeo, Harvard University.*

Nakamura, Emi, *"Layoffs and Lemons Over the Business Cycle," Economics Letters, 2008, 99 (1), 55 – 58.*

Nalebuff, Barry J and Joseph E Stiglitz, *"Prizes and incentives: towards a general theory of compensation and competition," The Bell Journal of Economics, 1983, pp. 21–43.*

Needleman, Jack, Peter Buerhaus, V. Shane Pankratz, Cynthia L. Leibson, Susanna R. Stevens, and Marcelline Harris, *"Nurse Staffing and Inpatient Hospital Mortality," New England Journal of Medicine, 2011, 364 (11), 1037–1045.*

Nelson, Robert G. and David A. Bessler, *"Subjective Probabilities and Scoring Rules: Experimental Evidence," American Journal of Agricultural Economics, 1989, 71 (2), pp. 363–369.*

Neumark, David and William Wascher, *"Minimum Wages and Training Revisited," Journal of Labor Economics, 2001, 19 (3), 563–95.*

Oettinger, Gerald S., *"The Effects of Sex Education on Teen Sexual Activity and Teen Pregnancy," Journal of Political Economy, 1999, 107 (3), 606–635.*

of Pediatrics, American Academy, *"Emergency Contraception: Policy Statement," Pediatrics, 2005, 116 (4), 1038–1047.*

Offerman, Theo, Joep Sonnemans, Gijs Van De Kuilen, and Peter P. Wakker, *"A Truth Serum for Non-Bayesians: Correcting Proper Scoring Rules for Risk Attitudes," Review of Economic Studies, 2009, 76 (4), 1461–1489.*

Oman, Nathan B., *"Specific Performance and the Thirteenth Amendment," Minnesota Law Review, 2009, 93 (6), 2020–2099.*

Orrison, Alannah, Andrew Schotter, and Keith Weigelt, *"Multiperson tournaments: An experimental examination," Management Science, 2004, 50 (2), 268–279.*

Ortner, Gerhard, *"Forecasting markets– An industrial application," 1998.*

Ottaviani, Marco, *"The Design of Idea Markets: An Economist's Perspective," The Journal of Prediction Markets, 2009, 3 (1), 41.*

_ **and Peter Norman Sørensen**, *"Aggregation of information and beliefs in prediction markets," London Business School, 2006.*

_ **and** _ , *"Outcome manipulation in corporate prediction markets," Journal of the European Economic Association, 2007, 5 (2-3), 554–563.*

Oyer, Paul and Scott Schaefer, *"Why do some firms give stock options to all employees?: An empirical examination of alternative theories," Journal of financial Economics, 2005, 76 (1), 99–133.*

_ **and** _ , *"Why Do Some Firms Give Stock Options to All Employees?: An Empirical Examination of Alternative Theories," Journal of Financial Economics, 2005, 76 (1), 99–133.*

_ **and** _ , *"Personnel Economics: Hiring and Incentives," Handbook of Labor Economics, 2011.*

Pakes, Ariel and David Pollard, *"Simulation and the Asymptotics of Optimization Estimators," Econometrica, 1989, 57 (5), 1027–1057.*

Pallais, Amanda and Emily Sands, *"Why the Referential Treatment? Evidence from Field Experiments on Referrals," working paper, 2013.*

Pantano, Juan and Yu Zheng, *"Using Subjective Expectations Data to Allow for Unobserved Heterogeneity in Hotz-Miller Estimation Strategies," 2010. Mimeo.*

Papageorgiou, Theodore, *"Learning Your Comparative Advantages," 2010. Mimeo, Pennsylvania State University.*

Parent, Daniel, *"Wages and Mobility: The Impact of Employer-Provided Training," Journal of Labor Economics, 1999, 17 (2), pp. 298–317.*

Paserman, M.Daniele, *"Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation," Economic Journal, 2008, 118 (531), 1418–1452.*

Pellizzari, Michele, *"Do Friends and Relatives Really Help in Getting a Good Job?," Industrial and Labor Relations Review, 2010, 63 (3), 494–510.*

Peltzman, Sam, *"The Effects of Automobile Safety Regulation," Journal of Political Economy, 1975, 83 (4), 677–725.*

Peterson, Jonathan, *"Employee Bonding and Turnover Efficiency," 2010. Mimeo, Cornell University.*

Phatak, Narahari, *"Information Inside the Firm - Evidence from a Prediction Market," 2012.*

_ , *"Shifting Incentives in Forecasting Contests," 2012.*

Pigou, Arthur C., *Wealth and Welfare, London: Macmillan, 1912.*

Pinkston, Joshua C., *"How Much Do Employers Learn from Referrals?," Industrial Relations, 2012, 51 (2), 317–341.*

Plott, Charles R and Shyam Sunder, *"Efficiency of experimental security markets with insider information: An application of rational-expectations models," The Journal of Political Economy, 1982, pp. 663–698.*

Powell, James, *"Least Absolute Deviations Estimation of the Censored Regression Model," Journal of Econometrics, 1984, 25 (3), 303–325.*

Prendergast, Canice, *"The provision of incentives in firms," Journal of economic literature, 1999, pp. 7–63.*

_ , *"The Provision of Incentives in Firms," Journal of Economic Literature, 1999, 37 (1), pp. 7–63.*

_ , *"Contracts and Conflicts in Organizations," 2010. Mimeo, University of Chicago.*

Professional Truck Driver Institute, *"Certification Standards & Requirements for Entry-Level Tractor-Trailer Driver Courses, http://www.ptdi.org/errata/CERTIFICATION_STANDARDS1.pdf," 2010.*

Quaglieri, Philip L., *"A note on variations in recruiting information obtained through different sources," Journal of Occupational Psychology, 1982, 55 (1), 53–55.*

Radner, Roy, *"Hierarchy: The economics of managing," Journal of economic literature, 1992, pp. 1382–1415.*

Radzevick, Joseph R. and Don A. Moore, *"Competing to Be Certain (But Wrong): Market Dynamics and Excessive Confidence in Judgment," Management Science, 2011, 57 (1), 93–106.*

Rafaeli, Anat, Ori Hadomi, and Tal Simons, *"Recruiting through advertising or employee referrals: Costs, yields, and the effects of geographic focus," European Journal of Work and Organizational Psychology, 2005, 14 (4), 355–366.*

Rammstedt, Beatrice and Oliver P. John, *"Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German," Journal of Research in Personality, 2007, 41 (1), 203 – 212.*

Rees, Albert, *"Information Networks in Labor Markets," American Economic Review, 1966, 56 (1/2), 559–566.*

Rob, Rafael and Peter Zemsky, *Cooperation, corporate culture and incentive intensity, INSEAD, 1997.*

Rockoff, Jonah E., *"The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data," The American Economic Review, 2004, 94 (2), pp. 247–252.*

Rodriguez, Daniel A., Felipe Targa, and Michael H. Belzer, *"Pay Incentives and Truck Driver Safety: A Case Study," Industrial and Labor Relations Review, 2006, 59 (2), 205–225.*

Roll, Richard, *"Orange juice and weather," The American Economic Review, 1984, 74 (5), 861–880.*

Rose, Nancy L, *"Labor Rent Sharing and Regulation: Evidence from the Trucking Industry," Journal of Political Economy, 1987, 95 (6), 1146–78.*

Roth, Alex, *"Layoffs Deepen Pool of Applicants for Jobs Driving Big Rigs," Wall Street Journal, 2009. 2/27/2009.*

Rothschild, David, *"Forecasting Elections Comparing Prediction Markets, Polls, and Their Biases," Public Opinion Quarterly, 2009, 73 (5), 895–916.*

Rothstein, Jesse, *"Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement," Quarterly Journal of Economics, 2010, 125 (1), 175–214.*

Roy, Andrew Donald, *"Some thoughts on the distribution of earnings," Oxford economic papers, 1951, 3 (2), 135–146.*

Rust, John, *"Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," Econometrica, 1987, 55 (5), 999–1033.*

_ , *"Maximum Likelihood Estimation of Discrete Control Processes," SIAM Journal on Control and Optimization, 1988, 26, 1006–1024.*

_ , *"Structural Estimation of Markov Decision Processes," Handbook of Econometrics, 1994, pp. 3081–3143.*

_ , *"Numerical Dynamic Programming in Economics," in H. M. Amman, D. A. Kendrick, and J. Rust, eds., Handbook of Computational Economics, Vol. 1 of* Handbook of Computational Economics, *Elsevier, 1996, chapter 14, pp. 619–729.*

_ **and Christopher Phelan**, *"How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets," Econometrica, 1997, 65 (4), pp. 781–831.*

Rustichini, Aldo, Colin DeYoung, Jon Anderson, and Stephen Burks, *"Toward the Integration of Personality Theory and Decision Theory in the Explanation of Economic Behavior," Mimeo, University of Minnesota 2012.*

Ruud, Paul, *Introduction to Classical Econometric Theory, New York, NY: Oxford University Press, 2000.*

Sacerdote, Bruce, *"Peer effects in education: How might they work, how big are they and how much do we know thus far?," Handbook of the Economics of Education, 2011, 3, 249–277.*

Saks, Alan M., *"A psychological process investigation for the effects of recruitment source and organization information on job survival," Journal of Organizational Behavior, 1994, 15 (3), 225–244.*

Sanders, Carl, *"Skill Uncertainty, Skill Accumulation, and Occupational Choice," 2011. Mimeo.*

Sandroni, Alvaro and Francesco Squintani, *"Overconfidence, Insurance, and Paternalism," American Economic Review, 2007, 97 (5), 1994–2004.*

Sanna, Lawrence J., Norbert Schwarz, and Shevaun L. Stocker, *"When Debiasing Backfires: Accessible Content and Accessibility Experiences in Debiasing Hindsight," Journal of Experimental Psychology: Learning, Memory, and Cognition, 2002, 28 (3), 497 – 502.*

Schlag, Karl H. and Joel van der Weeley, *"Eliciting Probabilities, Means, Medians, Variances and Covariances without assuming Risk Neutrality," 2009. Mimeo, Universitat Pompeu Fabra.*

Schmidt, E. and J. Rosenberg, *How Google Works, Grand Central Publishing, 2014.*

Schmutte, Ian M., *"Job Referral Networks and the Determination of Earnings in Local Labor Markets," Journal of Labor Economics, forthcoming, 2015.*

Schoenberg, Uta, *"Testing for Asymmetric Employer Learning," Journal of Labor Economics, 2007, 25, 651–691.*

Schwarzstein, Joshua, *"Selective Attention and Learning," 2012. Mimeo.*

Segal, David, *"Law Students Lose the Grant Game as Schools Win," New York Times, April 2011.*

Servan-Schreiber, Emile, Justin Wolfers, David M Pennock, and Brian Galebach, *"Prediction markets: Does money matter?," Electronic Markets, 2004, 14 (3), 243–251.*

Shaw, Kathryn and Edward Lazear, *"Tenure and Output," Labour Economics, 2008, 15 (4), 704–723.*

Sheremeta, Roman M, *"Behavioral Dimensions of Contests," Available at SSRN 2476020, 2014.*

Shinnar, Rachel S., Cheri A. Young, and Marta Meana, *"The Motivations for and Outcomes of Employee Referrals," Journal of Business and Psychology, 2004, 19, 271–283.*

Siegel, Ron, *"All-Pay Contests," Econometrica, 2009, 77 (1), 71–92.*

Silver, Nate, *"Which polls fared best (and worst) in the 2012 presidential race," The New York Times, 2012, 10.*

Simon, Curtis J. and John T. Warner, *"Matchmaker, Matchmaker: The Effect of Old Boy Networks on Job Match Quality, Earnings, and Tenure," Journal of Labor Economics, 1992, 10 (3), 306–30.*

Simpson, Edward H, *"Measurement of diversity.," Nature, 1949.*

Sleuwaegen, Leo and Wim Dehandschutter, *"The critical choice between the concentration ratio and the H-index in assessing industry performance," The Journal of Industrial Economics, 1986, pp. 193–208.*

Slonim, Robert and Alvin E Roth, *"Learning in high stakes ultimatum games: An experiment in the Slovak Republic," Econometrica, 1998, pp. 569–596.*

Small, Mario Luis, *"Racial Differences in Networks: Do Neighborhood Conditions Matter?," Social Science Quarterly, 2007, 88 (2), 320–343.*

Smith, Adam, *"The Wealth of Nations, London: W," Strahan and T. Cadell, 1776.*

Smith, Vernon L, *"An experimental study of competitive market behavior," The Journal of Political Economy, 1962, pp. 111–137.*

Snowberg, Erik and Justin Wolfers, *"Explaining the favorite-longshot bias: Is it risk-love or misperceptions?," Technical Report, National Bureau of Economic Research 2010.*

Soskice, David, *"Reconciling Markets and Institutions: The German Apprenticeship System," in "Training and the Private Sector" NBER Chapters, National Bureau of Economic Research, Inc, 1994, pp. 25–60.*

Soukhoroukova, Arina, Martin Spann, and Bernd Skiera, *"Sourcing, filtering, and evaluating new product ideas: an empirical exploration of the performance of idea markets," Journal of Product Innovation Management, 2012, 29 (1), 100–112.*

Spinnewjin, Johannes, *"Training and Search during Unemployment," 2009. Mimeo, LSE.*

_ , *"Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs," 2010. Mimeo, LSE.*

Stange, Kevin M., *"An Empirical Investigation of the Option Value of College Enrollment," American Economic Journal: Applied Economics, 2012, 4 (1), 49–84.*

Staw, Barry M., *"The Consequences of Turnover," Journal of Occupational Behavior, 1980, 1, 25373.*

Stevens, Margaret, *"A Theoretical Model of On-the-Job Training with Imperfect Competition," Oxford Economic Papers, 1994, 46 (4), 537–62.*

Stigler, George J, *"A theory of oligopoly," The Journal of Political Economy, 1964, pp. 44–61.*

Stiglitz, Joseph E, *"Incentives and Risk Sharing in Sharecropping," Review of Economic Studies, 1974, 41 (2), 219–55.*

Stinebrickner, Ralph and Todd Stinebrickner, *"Learning about academic ability and the college drop-out decision," 2011. Mimeo.*

Stinebrickner, Todd, *"Compensation Policies and Teacher Decisions," International Economic Review, 2001, 42 (3), 751–79.*

Stracke, Rudi and Uwe Sunde, *"Dynamic Incentive Effects of Heterogeneity in Multi-Stage Promotion Contests," 2014.*

Sunstein, Cass R, *Infotopia: How many minds produce knowledge, Oxford University Press, USA, 2006.*

Suzuki, Yoshinori, Michael R. Crum, and Gregory R. Pautsch, *"Predicting Truck Driver Turnover," Transportation Research Part E: Logistics and Transportation Review, 2009, 45 (4), 538 – 550.*

Swisher, Kara, *"'Because Marissa Said So' – Yahoos Bristle at Mayer's QPR Ranking System and 'Silent Layoffs'," AllThingsD, 2013.*

Taber, Mary E. and Wallace Hendricks, *"The effect of workplace gender and race demographic composition on hiring through employee referrals," Human Resource Development Quarterly, 2003, 14 (3), 303–319.*

Terpstra, David E. and Elizabeth J. Rozell, *"The Relationship of Staffing Practices to Organizational Level Measures of Performance," Personnel Psychology, 1993, 46 (1), 27–48.*

Tervio, Marko, *"Superstars and Mediocrities: Market Failure in the Discovery of Talent," Review of Economic Studies, 2009, 76 (2), 829–850.*
The Chicago Manual of Style

The Chicago Manual of Style, thirteenth ed., University of Chicago Press,

Thomas, C William, "The rise and fall of Enron," *JOURNAL OF ACCOUNTANCY-NEW YORK-*, 2002, *193* (4), 41–52.

Thornton, Rebecca, "The Demand For, and Impact of, Learning HIV Status," *American Economic Review*, 2008, *98* (5), 1829–1863.

Tirole, Jean, *The theory of industrial organization*, MIT press, 1988.

Todd, Petra E. and Kenneth I. Wolpin, "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, 2006, *96* (5), 1384–1417.

Topa, Giorgio, "Social Interactions, Local Spillovers and Unemployment," *Review of Economic Studies*, 2001, *68* (2), 261–295.

_ , "Labor Markets and Referrals," *Handbook of Social Economics*, 2011, *1*, 1193–1221.

Trachter, Nicholas, "Ladders in Post-Secondary Education: Academic 2-year Colleges as a Stepping Stone," Mimeo, University of Chicago 2009.

Train, Kenneth, *Discrete Choice Models with Simulation*, New York, NY: Cambridge University Press, 2003.

Transport Topics, "Transport Topics: Top 100 For-Hire Carriers," 2009.

Tziralis, Georgios and Ilias Tatsiopoulos, "Prediction markets: An extended literature review," *The journal of prediction markets*, 2007, *1* (1), 75–91.

Van den Steen, Eric, "Rational Overoptimism (and Other Biases)," *American Economic Review*, 2004, *94* (4), 1141–1151.

_ , "On the origin of shared beliefs (and corporate culture)," *RAND Journal of Economics*, 2010, *41* (4), 617–648.

van der Klaauw, Wilbert, "On the Use of Expectations Data in Estimating Structural Dynamic Choice Models," 2011. Mimeo.

_ **and Kenneth I. Wolpin**, "Social security and the retirement and savings behavior of low-income households," *Journal of Econometrics*, 2008, *145* (1-2), 21–42.

Van Hoye, Greet and Filip Lievens, "Tapping the grapevine: A closer look at word-of-mouth as a recruitment source," *Journal of Applied Psychology*, 2009, *94* (2), 341–352.

Vecchio, Robert P., "The Impact of Referral Sources on Employee Attitudes: Evidence from a National Sample," *Journal of Management*, 1995, *21* (5), 953–965.

Viscusi, W. Kip and Charles J. O'Connor, "Adaptive Responses to Chemical Labeling: Are Workers Bayesian Decision Makers?," *The American Economic Review*, 1984, *74* (5), pp. 942–956.

_ **and Michael J. Moore**, "Worker Learning and Compensating Differentials," *Industrial and Labor Relations Review*, 1991, *45* (1), pp. 80–96.

Wahba, Jackline and Yves Zenou, "Density, social networks and job search methods: Theory and application to Egypt," *Journal of Development Economics*, December 2005, *78* (2), 443–473.

Waldman, Michael, "Up-or-Out Contracts: A Signaling Perspective," *Journal of Labor Economics*, 1990, *8* (2), 230–250.

Walker, James M, Vernon L Smith, and James C Cox, "Inducing Risk-Neutral Preferences: An Examination in a Controlled Market Environment," *Journal of Risk and Uncertainty*, March 1990, *3* (1), 5–24.

Walton, D., "Examining the self-enhancement bias: professional truck drivers' perceptions of speed, safety, skill and consideration," *Transportation Research Part F: Traffic Psychology and Behaviour*, 1999, *2* (2), 91 – 113.

Wang, Yang, "Subjective Expectations: Test for Bias and Implications for Choices," Mimeo, Lafayette College 2010.

Warner, John T. and Saul Pleeter, "The Personal Discount Rate: Evidence from Military Downsizing Programs," *American Economic Review*, 2001, *91* (1), 33–53.

Welch, Jack, "The vitality curve," *Leadership Excellence*, 2005, *22* (9), 4.

_ **and John A Byrne**, "Straight from the Gut," *New York*, 2001.

Werbel, James D. and Jacqueline Landau, "The Effectiveness of Different Recruitment Sources: A Mediating Variable Analysis," *Journal of Applied Social Psychology*, 1996, *26* (15), 1337–1350.

Williamson, Oliver E, "Markets and hierarchies: some elementary considerations," *The American economic review*, 1973, pp. 316–325.

_ , "Markets and hierarchies," *New York*, 1975, pp. 26–30.

Williamson, Oliver E., "Calculativeness, Trust, and Economic Organization," *Journal of Law and Economics*, 1993, *36* (1), 453–86.

Wiswall, Matthew and Basit Zafar, "Determinants of College Major Choices: Identification from an Information Experiment," Mimeo, New York Federal Reserve Bank 2012.

Wolfers, Justin and Eric Zitzewitz, "Prediction markets," Technical Report, National Bureau of Economic Research 2004.

_ **and** _ , "Interpreting prediction market prices as probabilities," Technical Report, National Bureau of Economic Research 2006.

Wolthoff, Ronald, "Applications and Interviews: A Structural Analysis of Two-Sided Simultaneous Search," working paper, 2012.

Woodland, Linda M and Bill M Woodland, "Market Efficiency and the Favorite-Longshot Bias: The Baseball Betting Market," *The Journal of Finance*, 1994, *49* (1), 269–279.

Wooldridge, Jeffrey, *Econometric Analysis of Cross Section and Panel Data*, New York, NY: MIT Press, 2002.

Yakubovich, Valery and Daniela Lup, "Stages of the Recruitment Process and the Referrers Performance Effect," *Organization Science*, 2006, *17* (6), 710–723.

yi Wang, Shang, "Marriage Networks and Labor Market Outcomes in China," Mimeo, NYU 2012.

Yoganarasimhan, Hema, "Estimation of beauty contest auctions," *Available at SSRN 2250472*, 2013.

Zafar, Basit, "How Do College Students Form Expectations?," Mimeo, New York Federal Reserve Bank 2010.

Zenou, Yves, "Explaining the Black/White Employment Gap: The Role of Weak Ties," working paper, 2012.

Zottoli, Michael A and John P Wanous, "Recruitment Source Research: Current Status and Future Directions," *Human Resource Management Review*, 2000, *10* (4), 353 – 382.