Personnel Economics: Hiring and Incentives

Paul Oyer and Scott Schaefer*

April 20, 2010

Abstract

We survey the Personnel Economics literature, focusing on how firms establish, maintain, and end employment relationships and on how firms provide incentives to employees. This literature has been very successful in generating models and empirical work about incentive systems. Some of the unanswered questions in this area — for example, the empirical relevance of the risk/incentive tradeoff and the question of whether CEO pay arrangements reflect competitive markets and efficient contracting — are likely to be very difficult to answer due to measurement problems. The literature has been less successful at explaining how firms can find the right employees in the first place. Economists understand the broad economic forces — matching with costly search and bilateral asymmetric information — that firms face in trying to hire. But the main models in this area treat firms as simple black-box production functions. Less work has been done to understand how different firms approach the hiring problem, what determines the firm-level heterogeneity in hiring strategies, and whether these patterns conform to theory. We survey some literature in this area and suggest areas for further research.

*Prepared for The Handbook of Labor Economics, Vol. 4, Ashenfelter and Card, eds. Oyer: Stanford GSB and NBER, oyer.paul@gsb.stanford.edu. Schaefer: David Eccles School of Business and Institute for Public and International Affairs, University of Utah, scott.schaefer@utah.edu. This survey has benefitted from conversations with colleagues too numerous to mention. We thank the editors, Orley Ashenfelter and David Card, for detailed comments.
1 Introduction

Personnel Economics is the study of the employment relationship. It is unlike most other fields of labor economics for two reasons. First, Personnel Economics has grown up largely within leading business schools, not economics departments. This has given the field a more normative orientation than what is typically found in economics. Because many researchers in this field must take their insights into MBA classrooms and offer advice to future managers, Personnel Economists are typically interested in how firms can solve human resource management problems and how the solutions to HR problems are related to firms’ broader strategic contexts. Second, Personnel Economics is notable in that it is shared between the fields of Labor Economics and Organizational Economics. Because of this, Personnel Economists typically do not treat a firm as a mere “black box” production function. The field is instead interested in understanding and explaining the wide array of human resource management choices made by firms.

Personnel Economics has made great progress in the past few decades, especially in the area of incentives. Personnel Economists, often applying key insights from advances in information economics, have developed theoretical models that capture both the broad issues and many of the details facing firms as they set up incentive systems. Rigorous and clever empirical work has confirmed the relevance of these models and, in some cases, found some potential holes as well. We highlight the success of both empirical and theoretical studies of incentives relative to the literature on hiring. As the labor market continues to get more skilled and employer human resource strategies continue to get more sophisticated, the opportunities to create economic surplus through efficient matching of employees and firms have likely grown and probably will continue to grow. We argue that hiring models developed to date are too far removed from the strategic issues firms face and the empirical work is simply too limited. The relative weakness of the hiring literature is a function of several things, including idiosyncrasies in how firms approach the issues and data limitations. But we are hopeful that new data and new approaches will make research advances possible in this area and we suggest some avenues for future research.

More specifically, we believe that, in developing the literature on incentives in firms, economists have got it right (mostly). Agency-theoretic models — which explain the risk/incentive trade-
off, multitasking, gaming, subjective performance evaluation, career concerns, tournaments, and the like — are probably right (again, mostly). Empirical work has either confirmed the relevance of these theories or researchers have reached a point where limits on measurement preclude sharp tests of the theory. Further, this research provides a good sense for what factors explain across-firm and within-firm variation in the use of various incentive tools. Scholars who teach MBAs at leading (and some not-so-leading) business schools have used these ideas — specifically ideas about firm-level factors that influence the efficiency of various incentive mechanisms — to integrate the economics of incentives into broader discussions of organizational and product-market strategy.

While economists have a lot to say about how a firm can motivate an employee, we have far less to say about how the firm should go about finding the right employee in the first place. We do think there are models that help identify the main economic problem — matching in the presence of search costs and bilateral asymmetric information — in hiring. And there is empirical research consistent with the hypotheses that matching, search costs, and asymmetric information all affect firms’ hiring choices, as well as scattered research on specific strategies that firms might pursue to hire employees.1

But as business economists, our critique of this hiring literature is that for the most part the firm is treated as a black box. What’s lacking is (a) documentation of across-firm variation in hiring strategies, (b) linkage of this across-firm variation in strategy to firm-level characteristics, and (c) a tie from these facts back to theory. For example, as we shall discuss, Lazear (1998) offers conditions under which hiring risky workers can be a profit-maximizing strategy for firms. But there are notably few studies that examine across-firm variation in propensity to hire risky workers, and then whether the observed variation fits with Lazear’s theory. As another example, Montgomery (1991) suggests that firms can mitigate asymmetric information problems by accessing workers’ social networks in making hiring decisions. But how much across-firm variation is there in network-based hiring? And what exactly does theory lead us to expect about this across-firm variation?

We have two primary goals in writing this survey. First, we hope to encourage personnel economists to redirect their efforts (at least partially) away from explaining firms’ choices

---

1We are certainly not the first to make the point that the demand side of the labor market — and hiring, in particular — is understudied. See Devine and Kiefer (1991), for example.
with regard to incentive compensation and toward firms’ choices with regard to recruitment strategies. We believe such a shift will benefit organizational economics generally, but may also have important spillover effects on other fields of economics. Macroeconomists have, for example, long focused on frictions in labor markets as an important source of business cycles (see Oi, 1962). Second, we offer a catalog of research on firm-level recruitment strategies, in the hope that economists will work to improve this literature.

This isn’t to say that the marginal social return to research on incentives in organizations is zero. Many unanswered questions remain in that area — especially, we think, surrounding the use of relational incentive contracts and subjective assessments of employee performance — and we look forward to reading that research in the future. But we believe the social return to research on hiring is much larger, both because we know less about hiring, and because hiring the right employee is potentially as important or more so than motivating the employee to take the right action after the employee has been hired (at least for some firms).

We organize this survey as follows. First, we offer a short review of the literature on incentives in organizations. Our aim here is not to be exhaustive or complete (see Lazear and Oyer, 2010 for a fuller discussion); instead, we hope to point out some successes and attribute some of the failures to measurement problems that will be hard to solve in future research. Our two broad conclusions here are: (a) we have good answers to many of the big questions, and (b) some of the unanswered questions are likely to be very hard to answer well.

Then, we turn our attention to the question of how firms hire. We outline the basic models of matching, search, and asymmetric information. We discuss the empirical evidence that speaks to the importance of each of these factors. Then we review the somewhat scattered literature on firms’ actual hiring practices. Specific hiring practices we discuss include hiring risky workers, use of labor-market intermediaries, raiding other employers, hiring CEOs, use of various screening techniques, accessing employees’ social networks, and the influence of firing costs on hiring choices. We conclude with a call for new research.

2 Incentives in Organizations

The broad economic question surrounding incentives involves distributed benefits and costs in the presence of asymmetric information. Most employees take actions that lead to direct
benefits to the firm but not to the employee; that is, employees do not directly capture the full marginal benefit of their actions. Efficiency requires that employees’ actions be the ones that maximize total benefit less total cost, but the distributed nature of benefit and cost plus the potential asymmetry of information regarding these benefits and costs makes the problem of motivating efficient actions quite complex.

Over the past 40 years, the huge “Contract Theory” literature has developed around this set of issues. While contract theory is far broader than the narrow study of incentives in employment relationships, the insights developed there have been very useful in understanding pay-for-performance relationships in employment. We review this literature while making three main points. First, we argue that it is well established that financial incentives do change behavior in organizations. Second, we argue that Personnel Economists understand the broad outlines of how incentives in organizations work. Firms provide incentives through a great variety of means, and empirical research suggests that our models of incentive pay are (mostly) right. Third, we argue that two of the large unanswered questions in this literature — the risk/incentive tradeoff and whether CEO pay packages are structured correctly — are unlikely to be easily answered. We do not intend for our discussion below to be a complete survey of the field; see, for example, Gibbons and Roberts (2010) for a more detailed summary.

2.1 Financial incentives do change behavior

One does still hear the claim that financial incentives do not change behavior (see, for example, Ariely, 2008). If true, this claim would be very problematic for Personnel and Organizational Economics. The agency-theoretic view of behavior — which is foundational to these fields — is that an agent selects from a set of actions with the objective of maximizing his or her expected utility. In a basic agency model, the agent bears 100% of the disutility or cost of effort, but the principal captures 100% of the benefit. In the absence of some means for aligning the two parties’ interests, the agent will not select the efficient effort level. A solution to this problem is to tie the agent’s utility to the principal’s benefit, usually by varying the monetary payment made to the agent in response to variation in some measure that’s related to the principal’s benefit. By changing the mapping from actions to utility, the principal changes the set of utility-maximizing actions, and this means that financial incentives can change the agent’s action. If financial incentives do not change behavior, then economists err in writing down
incentive constraints in principal/agent models.

This point is so fundamental to both Personnel and Organizational Economics that it bears repeating. Empirical research shows that financial incentives do change behavior in organizations. This has been shown repeatedly by economists using controlled field experimental methods in real firms using real incentives. The most convincing of these studies examine simple jobs where researchers are allowed by the firm’s management to vary the firm’s incentive pay plans using an experimental design. Simple jobs are preferred because employees’ actions can easily be measured, and the experimental design eliminates issues surrounding the endogenous choice of compensation plans.

Lazear (2000) studies the implementation of a pay-for-performance plan for automobile windshield installers at Safelite Glass. He reports four main results. First, a switch from hourly to piece rate pay led to a 44% increase in on per-worker output. Second, about half of this gain came from an increase in individual-level productivity as a result of the stronger incentives. The remainder is attributable to changes in the composition in the firm’s workforce after the implementation of the pay plan. Third, it appears the firm is compensating the employees for higher effort costs as a result of the switch. Per-worker pay increased about 10% after the implementation of the piece rate plan. Fourth, piece rates increased the across-worker variance of output, as better workers faced a stronger incentive to differentiate from others.

Shearer (2004) examines data from a field experiment involving tree planters in British Columbia. In his experiment nine men were randomly selected from the firm, and then randomly assigned to be paid using piece rates or a fixed wage. Each worker was observed under piece rates for 60 days, and also under fixed wages for 60 days. Piece rates led to a 20% increase in individual-level productivity, a figure that is on par with that observed by Lazear (2000). Shearer goes on to estimate a structural model of the worker’s response to incentives.

Bandiera, Barankay and Rasul (2009) study a change in managerial incentives in an English fruit-picking operation. When managers are paid using fixed wages, they tend to favor those with whom they are “socially connected,” as measured by shared country-of-origin, shared living quarters, or whether the manager and worker arrived at the farm at the same time. Socially connected workers benefit from managerial favoritism in the form of higher output — by 9% — which leads to higher pay since workers are paid piece rate throughout this experiment. However, when managerial pay is switched from a fixed wage to bonuses based on the overall
output, managers change their behavior. They instead begin favoring the most able workers regardless of the social connections.

So, do pay-for-performance incentives work? As Besanko et al. (2009) point out, the answer to this question likely depends on what we mean by the word “work.” Pay-for-performance incentives surely change behavior in organizations, as the studies cited above make clear. But this is different from saying that pay-for-performance incentives improve organizational performance in all contexts. Pay-for-performance appears to induce employees to take actions that improve measured performance, but (as the multi-tasking literature discussed below emphasizes) there may be important-but-harder-to-measure aspects of performance that are ignored as employees work to hit measured-performance benchmarks. The studies cited above could omit such effects, as they focus on simple-to-measure aspects of employee performance. Further, it is easy to find cases where pay-for-performance does not improve performance (and see Freeman and Kleiner, 2005 for one such example in the context of shoe manufacturing). Broadly speaking, however, the available empirical evidence suggests that pay-for-performance incentives are associated with improvements in organizational performance; see Bloom and Van Reenen (2010) for a thorough review of this literature.

2.2 How do firms provide incentives?

Firms provide incentives in an astonishing variety of ways. In some jobs, pay is tied to a specific performance measure using a specific functional form. In others, supervisors make subjective judgments about the quality of an employee’s performance. Promotions are important in other cases, and in still other cases it is access to future labor-market opportunities that seem to drive choices.

Important early insights were offered by Holmstrom (1979). In his model, a risk-averse agent selects a level of costly effort \( e \) to maximize his expected utility. A risk-averse principal does not observe the agent’s effort choice, but does observe “output” \( x \) which is affected both by the agent’s effort choice and by a random state of nature. It is important to recognize the specific (and narrow) agency problem considered in this paper. The agent’s choice is simply how much effort to exert — not what kind of effort to exert — and the marginal return to effort is known. This turns the principal’s problem into one of statistical inference. Holmstrom shows that the optimal sharing rule — \( s(x) \), the share of output \( x \) that the principal pays to
the agent — is characterized by
\[
\frac{G'(x - s(x))}{U'(s(x))} = \lambda + \mu \frac{f_e(x; e)}{f(x; e)}.
\] (1)

Condition (1) is intuitive. The left-hand side is the ratio of the marginal utilities for the principal (\(G\)) and agent (\(U\)). The right-hand side is a Lagrange multiplier for the agent’s participation constraint (\(\lambda\)) and a multiplier for the incentive constraint (\(\mu\)) times the marginal effect of effort on the likelihood of obtaining that \(x\), scaled by the likelihood of obtaining that \(x\).

Efficient risk-sharing requires the ratio of the principal and agent utilities to be equated across all output levels. Motivating effort, however, requires the agent to be paid more — yielding lower agent marginal utility — when the output is indicative of high effort. The \(x\) for which \(f_e/f\) is large are those indicative of high effort, and so the agent is paid more for these states. The optimal contract is monotone in \(x\) if \(f\) satisfies the monotone likelihood ratio property (that is, \(f_e/f\) increasing in \(x\)). This model also offers guidance on what information (in addition to output \(x\)) should be used in an optimal contract. Holmstrom’s Informativeness Principle states that any information that is incrementally informative about \(e\) should be used. That is, if \(x\) is not a sufficient statistic for \(y\) with respect to effort \(e\), then \(y\) is part of the second-best contract.

The Informativeness Principle suggests, for example, that relative performance evaluation can improve contracts when two employees outputs are positively correlated.

It is difficult to identify specific employment contracts that are well described by Holmstrom’s model. The model suggests that optimal sharing rules should be quite sensitive to the shape of the underlying probability distribution \(f\). Monotonicity and even linearity of sharing rules are common in organizations, but are predicted by this model only in very special cases. Further, firms commonly appear to violate the Informativeness Principle by ignoring some performance-related information in determining pay.

Models by Holmstrom and Milgrom (1991) and Baker (1992) connect agency theory more closely to observed employment relationships. Holmstrom and Milgrom (1991) build on the linear-contracting model in Holmstrom and Milgrom (1987) to examine “multi-task” principal/agent models. In their model, the agent privately makes a vector of effort choices, where we can interpret the elements of the vector as efforts toward various tasks. The principal is concerned not just with the overall level of effort (as in Holmstrom (1979)), but also with how the agent allocates effort across the various tasks. When the components of the effort vector
are substitutes in the agent’s cost function, the principal needs to take account of how offering stronger incentives toward one task will affect the agent’s choices toward other tasks. To see how this works, consider a simple two-task version of their model in which effort $e_i$ toward task $i \in \{1, 2\}$ influences output $x$ according to

$$x_i = e_i + \tilde{e}_i,$$

where $\tilde{e}_i$ is a normal, mean-zero random variable. If the agent’s pay varies with $x_1$ and $x_2$ according to

$$\beta_1 x_1 + \beta_2 x_2,$$

then first-order conditions for effort are

$$\beta_1 = \frac{\partial c(e_1, e_2)}{\partial e_1},$$

$$\beta_2 = \frac{\partial c(e_1, e_2)}{\partial e_2}.$$

Assuming $\frac{\partial^2 c(e_1, e_2)}{\partial e_2 \partial e_1} < 0$, higher $\beta_1$ implies that the agent will exert less effort toward task 2. The principal balances concerns for overall effort levels, allocation of effort across tasks, and optimal risk sharing in designing the optimal compensation contract. One important comparative static in their model is that the strength of incentives for task 1 ($\beta_1$) can be decreasing in the noise in the measurement of task 2 ($\tilde{e}_2$). When there is more risk associated with rewarding task 2, the firm optimally shifts toward weaker pay-for-performance incentives on that task. As task 2 incentives weaken, the employee will, according to the first order conditions above, shift effort away from task 2 and toward task 1. If the firm values task 2 sufficiently, it may find it optimal to weaken incentives toward task 1. This weakens overall effort incentives, but improves balance.

This notion can be applied to, for example, the question of whether to pay school teachers for test results. “Teaching to the test” is easily measurable, but other tasks — such as fostering student maturity and higher-order thinking skills — are not. Poor measurement of “student maturity” means that rewarding testing will lead teachers to shift effort excessively in this direction. School administrators may, in this case, prefer not to use test scores to determine teacher pay. This finding illustrates that the Informativeness Principle need not apply when firms need to motivate the right kind of effort rather than simply motivating the level of effort.

Baker (1992) makes a similar point using a single-task model in which a risk-neutral agent is privately informed about how a performance measure reflects effort. In this model, the
marginal return to agent effort is constant, but the employee’s output cannot be measured directly. Instead, the principal observes an imperfect performance measure. The marginal effect of effort on the performance measure is random, and is privately observed by the employee. This randomness in the measurement of performance means the employee sometimes exerts more effort than the first best, and sometimes less. Convexity of the agent’s cost function means the agent’s expected effort cost is higher when the performance measure is worse. Because the principal must compensate the agent for expected effort costs, the principal ties pay less closely to performance when performance is less well measured.

Broadly speaking, the Holmstrom (1979), Holmstrom and Milgrom (1991) and Baker (1992) models suggest two main costs of using pay-for-performance incentives when performance measurement is imperfect. First, tying pay to badly measured performance shifts risk onto agents, and the firm must compensate agents for bearing risk. Second, tying pay to badly measured performance can lead agents to choose the wrong actions. We discuss the empirical research on the risk/incentive tradeoff below. There is ample evidence that problems with performance measurement can lead agents to “game” the measures by selecting inefficient actions. Oyer (1998), for example, examines the relation between firms’ fiscal-year ends, non-linear incentive pay, and seasonality in revenues. He finds that fiscal-year ends influence business seasonality in most manufacturing industries. Revenues are higher toward the end of a fiscal year and lower at the beginning, compared to the middle. This is consistent with the notion that salespeople facing year-end quotas work to pull sales from the beginning of fiscal year \( t \) to the end of fiscal year \( t - 1 \). Larkin (2007) offers more direct evidence on this point in his study of proprietary sales data from a large enterprise software firm. He finds the firm’s non-linear incentive plan induces salespeople to shift a number of deals into a single quarter and avoid making deals in other quarters. Salespeople also use their (limited) discretion over pricing to entice customers into buying during periods that yield greater returns to the salespeople.

Firms commonly attempt to combat the problems with “objective” performance measures — numerical quotas and targets — with “subjective” assessments of employee performance. Subjective measures can include things like supervisors’ assessments or 360-degree peer reviews. Academic economists are subject to subjective performance evaluation at the tenure decision. In most universities, tenure isn’t based on a formula combining publication counts and citations; instead, senior faculty subjectively assess the quality of a junior professor’s work.
Baker, Gibbons and Murphy (1994) examine the interplay between objective and subjective assessments of employee performance. An important distinction between objective and subjective measures is that the latter are non-verifiable. Thus, contracts based on subjective assessments cannot be enforced by courts, and instead must be self-enforcing. Building on a repeated-game model of implicit contracts model by Bull (1987), Baker, Gibbons and Murphy (1994) first show that bonuses based on subjective evaluations are limited due to the firm’s reneging constraint. A firm that pays a promised bonus today maintains the employee’s trust and, as a result, is able to make credible promises to pay bonuses based on subjective measures in the future. The firm therefore compares the cost of paying the bonus to the value of future cooperation. Firms that value the future more heavily than the present are better able to pay bonuses based on subjective measures of employee performance. Baker, Gibbons and Murphy (1994) then consider how the presence of an objective measure of firm performance affects the subjective bonus. They find that objective and subjective measures are substitutes; as the objective measure of performance becomes a better alternative for the subjective measure, the firm places less value on its reputation for paying bonuses based on subjective measures.

We believe there is a great need for more empirical research on the use of implicit contracts and subjective performance evaluation in employment relationships. Hayes and Schaefer (2000) offer evidence consistent with the use of implicit contracts and subjective performance evaluation. They argue that if boards of directors base pay for CEOs partially on information that is not publicly available, then current pay for CEOs should predict future firm performance. Following the reasoning in Baker, Gibbons and Murphy (1994), they argue that this link should be stronger when the available objective performance measures are weaker. Their empirical analysis is consistent with these hypotheses, but this is at best an indirect support for the Baker, Gibbons and Murphy (1994) model.

Promotion tournaments — which can be based either on objective or subjective measures of performance — are a common feature within many firms. Lazear and Rosen (1981) model a firm that will promote one of two employees to a new position. Each employee takes an action that translates into noisy output and the firm commits in advance to promote the individual with the higher output. For an employee participating in a promotion tournament, the first
order condition for effort is given by

\[
\frac{\partial p(e_1, e_2)}{\partial e_1} \Delta W = c'(e_1).
\]

Effort increases the probability \( p \) of winning the promotion tournament. The employee gets an increase in compensation of amount \( \Delta W \) if she wins, so the marginal benefit of effort is the marginal effect of effort on the probability of winning times \( \Delta W \). Employees equate this to the marginal cost. Note that employee 1’s probability of winning depends also on employee 2’s effort choice, so equilibrium effort levels \( e_1 \) and \( e_2 \) are where this first-order condition holds for both employees simultaneously. Lazear and Rosen (1981) document a number of features of tournaments. First, appropriate choice of \( \Delta W \) can lead the employees to select the first-best effort level. Second, holding \( \Delta W \) constant, more noise in performance measurement will reduce effort. This effect occurs because luck, rather than skill, becomes relatively more important in determining the tournament winner. This reduces \( \frac{\partial p(e_1, e_2)}{\partial e_1} \). Third, tournaments are a form of relative performance evaluation, so any common shocks to employee performance are netted out. Fourth, when more employees compete for a promotion, \( \frac{\partial p(e_1, e_2)}{\partial e_1} \) falls and the firm may need to raise the prize \( \Delta W \) to compensate.

Empirical evidence suggests that promotions are an important determinant of wage changes. In their study of 20 years of wage data from a large firm, Baker, Gibbs and Holmstrom (1994) find that promotions and wage growth are highly correlated. This appears, however, to operate in a somewhat different manner to that suggested by Lazear and Rosen (1981). Individuals who are promoted receive small wage premiums in the year of the promotion but tend to be those individuals who receive large wage increases even in years when they are not promoted. Further, wage levels in this firm are not tied directly to job levels, as there is substantial variation in wages even among individuals at the same job level. DeVaro (2006) shows that promotions seem to be determined by relative performance for workers in a cross-section of establishments. He also estimates a structural model of tournaments, finding support for the assertions that employers set wage spreads to induce effort and that workers are motivated by larger promotion wage spreads.

Employees may also be motivated by the possibility of receiving outside offers. Fama (1980) introduced this notion of “career concerns??, which was then studied in detail by Holmstrom (1982). Holmstrom develops a model of symmetric uncertainty regarding worker ability. An
employee’s output today depends both on his ability and his hidden effort. The employee’s wage next period depends on the market’s posterior belief regarding his ability, given observed output today. From this setting, Holmstrom derives a “rat-race” equilibrium. Employees exert effort in a futile (at least in equilibrium) attempt to boost the market’s assessments of ability. Effort incentives in the model are strongest when the market’s prior about employee ability is more diffuse. Chevalier and Ellison (1999) study data on mutual fund managers, and show that portfolio choices seem to be influenced by career concerns. They show first that termination is more sensitive to performance for younger managers. Younger managers appear to respond to the market’s updating regarding ability by selecting portfolios with less unsystematic risk and more conventional sector weights.

2.3 Some Important but Hard-to-Answer Questions

2.3.1 The Risk/Incentive Tradeoff

One of the oldest theoretical predictions in the agency literature has proved to be one of the most difficult for empirical researchers. The tradeoff between risk and incentives arises if (a) employees have convex disutility for both risk and effort, and (b) performance measures are subject to random variation. In this case, the marginal benefit of using incentives comes from the fact that the employee’s effort choice is closer to the first-best when incentives are stronger, but the marginal cost is the increase in the employee’s risk premium. Because greater risk increases the marginal cost of incentives without affecting the marginal benefit, we get a clear comparative statics prediction: greater risk should be associated with weaker pay-for-performance incentives.

As Prendergast (1999, 2002a,b) has pointed out, however, empirical research on this topic offers weak support at best. This challenge has led many empirical researchers to look for new evidence in support of the tradeoff between risk and incentives and has also led to the development of models that lead to the prediction that incentives and risk will be positively related. This literature has been useful but not fully satisfying. The central problem is that almost any moral hazard model suggests that a large number of unobservables will influence the strength of the pay-for-performance relationship. To identify these, we examine a simple linear-exponential-normal agency model (of the type studied by Holmstrom and Milgrom, 1991
and applied to CEOs by Schaefer, 1998). Let an agent have CARA utility with coefficient \( \rho \), and convex disutility of effort with monetary equivalent given by \( \frac{1}{2} e^2 \), where \( e \) is a real-valued effort choice. Suppose further that effort translates into value \( V \) to the principal as follows:

\[
V = ve + \bar{\epsilon},
\]

where \( v > 0 \) and \( \bar{\epsilon} \) is a mean-zero normal random variable with variance \( \sigma^2 \). Assuming a wage contract that is linear in \( V \),

\[
Wage = \alpha + \beta V,
\]

it is easy to show that the optimal wage contract maximizes the total certainty equivalent of the two parties, subject to the agent’s incentive constraint. The problem is

\[
\max_{\beta} ve - \frac{1}{2} e^2 - \frac{1}{2} \rho \beta^2 \sigma^2
\]

subject to

\[
e \in \arg\max_{e} \beta ve - \frac{1}{2} e^2.
\]

The solution is

\[
\beta^* = \frac{1}{1 + \frac{\sigma^2}{2 \rho^2 \sigma^2}},
\]

and note that \( \beta^* \) — the optimal slope of the pay-performance relation — is strictly decreasing in \( \sigma^2 \), consistent with the prediction of a risk-incentive tradeoff.

We now consider the measurement challenges in devising a test of this relation. First is the simple problem of measuring “risk.” The \( \sigma^2 \) parameter in the theory is the conditional variance of the output measure \( V \). If effort \( e \) and the marginal productivity of effort \( v \) could somehow be held constant, then the conditional variance would be equal to the unconditional variance, and a test could examine whether \( \text{var}(V) \) affects the slope of the pay-for-performance relation. In most tests of the risk/incentives tradeoff (see Aggarwal and Samwick, 1999), some measure of the unconditional variance of a performance measure is used to assess risk. But, of course, the marginal return to effort \( v \) and the agent’s effort choice \( e \) are typically not observed by the econometrician. It remains unclear whether \( \text{var}(V) \) truly reflects the conditional variance of output.

Second note that there are three parameters besides \( \sigma^2 \) in our expression for \( \beta^* \). Incentives are stronger when the marginal return to effort is higher (higher \( v \)), when the agent is more risk
tolerant (lower $\rho$), and when the agent is more responsive to strong incentives (lower $c$). Notably none of these parameters are easily observable by the econometrician, and any correlation between these unobservables and $\sigma^2$ can confound tests of the risk/incentive tradeoff. Suppose, for example, that the marginal return to effort tends to be high in exactly the cases where $\sigma^2$ is high. Then we may observe stronger incentives in exactly the cases where risk is greatest. Such a finding would not imply that the theory of the risk/incentive tradeoff is necessarily wrong, but instead could indicate that we are unable to make a ceteris paribus comparison. Prendergast (2002a) argues that such a pattern might arise if firms delegate more decision-making authority to agents — leading to a higher marginal productivity of effort — in exactly those cases where more risk is present. Following Prendergast, Adams (2005) and DeVaro and Kurtulus (2006) attempt to control for the degree of delegation in framing a risk/incentives test, but find both measuring delegation and identifying exogenous variation in delegation to be significant challenges. Further, delegation is just one of many potential avenues that could lead to a positive association between $v$ and $\sigma^2$.

We think it is not clear how to solve these measurement problems surrounding the risk/incentive tradeoff. An empirical design with agent fixed effects can likely control for variation in $\rho$. But the marginal return to effort $v$ and the second-derivative of the agent’s cost function $c$ are presumably match specific, and this means it will be difficult to control for these with agent or firm fixed effects. Match-specific fixed effects could help, but only if we could identify sources of exogenous variation in the within-match $\sigma^2$ and be confident that this variation in risk is not also leading to variation in the marginal productivity of effort. We think this is likely to be a very tough nut to crack and, absent some novel measurement technique that we cannot currently envision, we do not see this as a fruitful research area.

2.3.2 The Structure of CEO Pay

Another persistent question in the broad incentives literature surrounds the structure of CEO pay packages. US CEO paychecks have risen ten times as fast as those of average workers since the 1970s (The Economist, 2006). Further, the tie between CEO pay and firm performance has strengthened, as more and more of CEO pay has come in the form of equity-based instruments. Literally hundreds of studies — in economics, finance, accounting, and management — have studied the question of whether CEO pay packages are efficient. Our fear is that despite all
this research — some of it conducted by the authors of this survey — social scientists really have very little conclusive evidence on whether CEO pay is structured correctly.

As with the risk/incentive tradeoff, we think the problems boil down to one of measurement. Our agency-theoretic models of pay suggest that the efficient sensitivity of pay to performance and the efficient level of pay depend on many unobservables. To illustrate our concerns, we return to the linear contracting model we developed above. To add in a discussion of the level of pay, we assume that the agent’s reservation utility is given by \( \bar{u} \) and that the firm’s reservation profit level is \( \bar{\pi} \). Again assuming a linear wage contract

\[
\text{Wage} = \alpha + \beta V,
\]

we ask what this model can tell us about \( \alpha \) and \( \beta \). As above, the efficient pay-for-performance term \( \beta \) maximizes the total certainty equivalent subject to the agent’s incentive constraint. We again have

\[
\begin{align*}
e^* &= \frac{\beta^* v}{c} \\
\beta^* &= \frac{1}{1 + \frac{c^2}{\rho \sigma^2}}.
\end{align*}
\]

The term \( \alpha \) — which can be interpreted as the employee’s level of pay when \( V = 0 \) — must be large enough to satisfy the agent’s participation constraint, which is given by

\[
\alpha + \beta^* \text{E}(V | e^*) - \frac{c}{2} e^{*2} - \frac{1}{2} \rho \beta^* \sigma^2 \geq \bar{u}. \tag{2}
\]

The \( \alpha \) term must not be so large as to leave the employer worse off than its reservation profit level. This implies

\[
(1 - \beta^*) \text{E}(V | e^*) - \alpha \geq \bar{\pi} \tag{3}
\]

If this employment match is efficient, we must have

\[
\text{E}(V | e^*) - \frac{c}{2} e^{*2} - \frac{1}{2} \rho \beta^* \sigma^2 \geq \bar{\pi} + \bar{\pi} \tag{4}
\]

Note that if the inequality in (4) is strict, then there are rents. For CEOs in particular, matching and specificity of human capital would seem to make it likely that rents are present. Combining the inequalities in (2) and (3), we have

\[
\text{E}(V | e^*) - \pi \geq \alpha + \beta^* \text{E}(V | e^*) \geq \frac{c}{2} e^{*2} + \frac{1}{2} \rho \beta^* \sigma^2 + \bar{u}.
\]
In words, our basic contracting model suggests that the employee’s expected level of pay $\alpha + \beta \cdot E(V | e^*)$ must insure participation of both parties, but beyond that the level of pay simply splits any match surplus.\(^2\) We assuming this surplus is split according to Nash Bargaining where the employee get share $\gamma$. Given this, the CEO’s expected level of pay is given by

$$
\gamma \left( E(V | e^*) - \pi \right) + (1 - \gamma) \left( \frac{c}{2} e^{s^2} + \frac{1}{2} \rho \beta^2 \sigma^2 + \bar{u} \right)
$$

This model suggests that CEO pay should depend on

1. The marginal return to managerial effort $v$,
2. The second derivative of the manager’s cost-of-effort function $c$,
3. The manager’s degree of risk aversion $\rho$,
4. The conditional variance of output $\sigma^2$,
5. The manager’s reservation utility $\bar{u}$,
6. The firm’s reservation profit level $\pi$,
7. The manager’s bargaining power $\gamma$.

Not one of these seven factors can be easily measured by empirical researchers. On top of that, several features of this market make it difficult to control for these factors using manager or firm fixed effects. First, firms employ just a single CEO at a time, and CEO tenure is typically a number of years. This means firm fixed effects are useful only to the extent that we believe firm characteristics do not change very quickly over time. Second, managers change jobs infrequently and not for exogenous reasons. Third, as we noted in our discussion of the risk/incentive tradeoff, factors like the marginal return to effort and the shape of the manager’s cost-of-effort function are likely to be match specific, which means they cannot be conditioned out easily.

\(^2\)Note, however, that many papers on efficient CEO pay contracts assume that the CEO is on his/her participation constraint and the firm captures all the rents. See, for example, Aggarwal and Samwick (1999). Kuhnen and Zwiebel (2009) take a different approach, by modeling pay as being set by the CEO himself, subject to limits on his entrenchment.
Given this, we think it is very difficult to draw conclusions over whether CEO pay practices are well explained by our basic models of contracting and labor markets. It seems that practically any broad pattern appearing in CEO pay data can be rationalized by a clever theorist who reverse-engineers the unobservables in such a way as to fit the data (see Edmans and Gabaix, 2009). This literature has, as a result, followed a bit of a he-said, she-said spiral, with few conclusions drawn. As examples, Bertrand and Mullainathan (2001) point out that CEOs are paid for “luck,” which seems to be inconsistent with the agency-theoretic Informativeness Principle from Holmstrom (1979). Oyer (2004) replies that if the employee’s outside option ($\bar{\pi}$, in our model above) is correlated with industry-wide share prices, then arrangements that look like pay-for-luck can be part of an optimal employment contract. Bebchuk and Fried (2003) are the most vocal current academic critics of CEO pay arrangements; broadly, their argument is that the patterns in CEO pay are hard to reconcile with any model of efficient contracting or competitive labor markets. Tervio (2009), Gabaix and Landier (2008), and Gayle and Miller (2009) reply that changes in firm scale may have affected the marginal return to managerial ability, which in turn changes the reservation utilities of all managers through labor market competition, which leads to a system of interrelated changes in firms’ pay plans that broadly seems to fit the pattern of changes in CEO pay over the past 30 years. Hayes and Schaefer (2009) further complicate the picture by offering a model in which CEO pay signals rents in the CEO-firm employment relationship. If firms care about short-run share prices, they may inflate CEO pay (above full-information levels) in a futile (in equilibrium) attempt to boost market value.

Having read (and written some of) this literature, we feel simply stuck. Theory suggests a long list of unobservables that should matter for CEO pay arrangements. It is not clear how empirical researchers can control for all of these factors well enough to draw firm conclusions about the degree to which CEO pay arrangements are or are not in line with theory. We know that CEO pay is not fully efficient at all firms, given the problems at firms such as MCI and Tyco. We also know that CEO pay contracts typically have features that economists predict to be part of an optimal contract (pay-for-performance that varies with regulation, age, and governance in the ways we might expect). But it seems unlikely that economic research will ever tell us exactly where the typical CEO pay arrangement lies on the spectrum from completely inefficient to completely optimal.
3 Hiring

In this section, we argue that while the fundamental economic problem in hiring is well understood, the methods that firms use to solve hiring problems still need a lot more research.

The fundamental economic problem in hiring is one of matching with costly search and bilateral asymmetric information. Job seekers have varying levels of aptitude, skill, and motivation, and firms have varying needs for these attributes. Economic efficiency requires that the labor market identify the best matches of workers to firms. The matching problem is complicated by the fact that firms and workers cannot costlessly observe all relevant aspects of potential trading partners. This means search is a common feature of hiring. A further complication is that firms and job seekers may each be able to misrepresent their quality as a trading partner. Potential employees are known to polish resumes or fabricate credentials, and firms at times may choose to downplay or conceal unpleasant aspects of the job. Labor markets are, of course, heterogeneous, so the extent to which matching, search, and asymmetric information are prevalent is likely to vary across labor markets.

To review this literature, we first outline the basic structure of our models of matching, search, and asymmetric information in labor markets. We then critique these basic models by pointing out that firms are, for the most part, treated as a mere production functions. The objective function given to firms in these models is to maximize the difference between an employee’s productivity and his wage. Because firms, in these models, are completely homogeneous, these models are by and large not useful for understanding firm-level heterogeneity in hiring strategies. We then discuss the literature on specific hiring practices and issues. This literature lacks focus, to a certain degree, and so our treatment here reads a bit like a laundry list of unconnected issues. We discuss hiring risky workers, use of labor-market intermediaries, raiding other employers, hiring CEOs, firms’ use of various screening techniques, accessing employees’ social networks, and the influence of firing costs on hiring choices.

3.1 Black-Box Models of Hiring

As noted above, labor economists have long recognized that hiring involves matching with costly search and bilateral asymmetric information. Jovanovich (1979b) draws out the implications of matching for labor markets. In his model, the productivity of a given worker/firm match is
unknown at the time of hiring. Once on the job, the worker/firm match quality becomes known over time as the firm gains observations about worker productivity. Employment matches persist as long as the expected surplus in the current employment relationship exceeds the parties' outside options. Matches that are revealed, over time, to be poor are terminated. Good matches persist, which implies that the hazard rate of worker/firm separations decreases with job tenure. This empirical implication is strongly borne out by the data (see Farber, 1999). The implications of job matching are, however, difficult to distinguish from those of firm-specific human capital (Jovanovich, 1979a). Under the hypothesis of firm-specific human capital formation, job matches become good over time as workers invest in skills that are specific to the firm. The question of whether the decreasing hazard rate of job loss is due to matching or firm-specific human capital remains open.

A large literature examines the effects of costly search on labor markets (see Mortensen and Pissarides, 1999). In the basic search model, workers sequentially sample wages from a known distribution. An unemployed worker’s strategy is characterized by an optimal stopping rule. Job offers that pay wages above an endogenous reservation wage are accepted, while others are declined. The basic employee search model has been applied to study unemployment durations, which are a function of the exogenous wage distribution and the exogenous rate at which wage offers arrive. Equilibrium search models explicitly consider search on both sides of the labor market, and endogenize wage distributions, job offer arrival rates, and firm-level vacancy durations. Search models that explicitly model heterogeneity in worker/firm match quality have been applied to understand both job flows and unemployment, and equilibrium wage dispersion.

The canonical Spence (1973) signaling model begins with the presumption that workers are privately informed about their productivity, and may take costly actions that credibly convey information about productivity. Greenwald (1986) notes that a firm’s incumbent employer is likely to hold a significant informational advantage (over potential rival employers) with regard to a given employee’s productivity. If employers focus their efforts on retaining those workers they privately observe to be able, then the stream of job changes will be adversely selected. Employers hiring from the pool of the unemployed will hire at low wages only. Asymmetric information can therefore impede the efficient matching of workers to firms.
3.2 Firm-Level Hiring Strategies

Despite the obvious success and empirical relevance of the models above, we think there is much work yet to be done to understand firm-level hiring choices. The work to date does not provide a good picture of where employers spend their resources and efforts when hiring workers and which hiring investments have proven most successful in various circumstances. They have also not generated a good sense for how to advise managers on developing a comprehensive recruitment strategy for their organizations. We review the work to date and discuss opportunities for future research (while noting the considerable impediments to doing the research we propose.)

3.2.1 Sources of Match-Specific Productivity

Discussions of hiring often begin with a desire to hire the right worker. But what makes a worker “right?” To put this in Jovanovic’s terms, what are the sources of match-specific productivity?

**Complementarities** Firm-level heterogeneity can lead to match-specificity in productivity if there is a complementarity between firm attributes and attributes of the employee. The assumption of such a complementarity underlies the large literature on assortive matching in labor markets (see Rosen, 1982 and Sattinger, 1993), but most of this literature simply assumes a complementarity between, say, firm size and employee ability, and goes on to derive implications for equilibrium matching. But what specific attributes are complementary? And what sources of firm-level heterogeneity give rise to these complementarities?

One answer is that employee attributes may complement certain production technologies. Information technology, for example, may be most productively utilized by employees with high skill levels. Such a complementarity lies at the root of the large literature on skill-biased technical change. Real prices for computing power have fallen dramatically since the 1970s, and this period has also seen dramatic changes in skill differentials in wages, as skilled workers saw much faster wage increases than unskilled. A complementarity between skilled labor and information technology can explain these facts, if falling IT prices caused firms to shift labor demand toward skilled workers.

---

Katz and Autor (1999) review much of the work on skill-biased technical change as part of a review of the literature on changing wage inequality. Autor et al. (2006) update that review in light of critiques such as Card and DiNardo (2002).
To cite some examples from this literature, Berman, Bound and Griliches (1994) use the US Annual Survey of Manufactures to examine skill upgrading in the 1980s. They show first that employment of production workers in US manufacturing dropped 15 percent in the 1980s, while employment of nonproduction workers rose 3 percent. This occurred despite the fact that relative wages for skilled labor rose over this period. The shift in employment toward nonproduction workers was driven primarily by changes within industry rather than between industries, and was larger in industries that made larger investments in computer technology. Autor, Katz and Krueger (1998) show that within-industry skill-upgrading (where skill is measured by educational attainment) accelerated from the 1970s to the 1980s and 1990s, and that various measures of computer usage were higher in the industries where skill-upgrading was highest.

Falling prices for information technology cannot, however, serve as an explanation for firm-level match-specificity in employee productivity. Changes in relative prices hit all firms equally, and hence in the absence of other firm-specific factors, all firms would shift their demand toward skilled workers equally. Bresnahan, Brynjolfsson and Hitt (2002) argue that investments in IT require coinvention on the part of individual firms. That is, firms cannot benefit fully from investments in IT without also reorganizing work practices and rethinking product offerings, a process that requires experimentation. Using detailed firm-level survey data, Bresnahan et al. (2002) show that conditional on investments in computerization, firms that do workplace reorganizations are more likely to also adopt high-investment human resource policies, such as screening for education in hiring, training, and cross training. The authors also report evidence of a complementarity in production between skilled labor, workplace reorganization and IT, as there are positive interaction effects in a regression of log value added on these variables. Ichniowski, Shaw and Prennushi (1997) report similar findings in their study of steel-finishing lines. Selection and training of skilled workers are complementary to adoption and installation of IT investments. Taken together, these results suggest that worker characteristics can be part of a constellation of complementary firm-level attributes, as suggested by Milgrom and Roberts (1990, 1995).

Potential employees, of course, can vary on a large number of dimensions, and most of the literature on skill-biased technical change focuses simply on a single dimension of skill as measured by educational attainment or a production/nonproduction worker distinction. Using
detailed firm-level data, Abowd, Haltiwanger, Lane, McKinsey and Sandusky (2007) consider how various components of skill are related to firms’ technological inputs. As with the skill-biased technical change literature, the authors report a strong relationship between technology and skill. Different dimensions of skill interact with technology in different ways, however. Using methods from Abowd, Kramarz and Margolis (1999), the authors decompose worker skill into a time-invariant worker effect (“basic ability”) and an experience effect that varies in proportion to labor-market experience. Interestingly, firms that use advanced production technologies are more likely to use high-ability workers, but less likely to use workers with extensive labor-market experience.

Complementarities between potential employee characteristics and firm characteristics can extend beyond firms’ technology choices. Andersson, Freedman, Haltiwanger, Lane and Shaw (2009) use matched employee-employer data in the software industry to study links between product-market segment and hiring strategies. They develop a simple theoretical model in which firms’ relative demand for successful innovation depends on characteristics of the product-market segment. Firms operating in market segments where payoff distributions are highly variable — video games, where having a blockbuster game can be worth hundreds of millions of dollars, are one example — will place a greater value on innovative employees. Empirical analysis supports this assertion. Firms operating in market segments with highly variable payoffs pay higher starting salaries than other firms. These firms also offer greater rewards for employee loyalty or experience.

**Specific Weights on General Skills**  Lazear (2003) proposes a model in which all skills are general, but firms place different weights on various combinations of skills. As an example, Lazear offers the case of a Silicon Valley startup that does tax optimization. A managerial employee in this firm may need to know about tax law, economics, and Java programming. These skills are all general, in the sense that there are other firms that would value each of these skills. But there may be no other firm that values an employee who holds all of these skills. An accounting firm may, for example, value knowledge of tax law and economics, but not java programming. An employee with all three skills who suffered job loss may, depending on market thickness and search costs, be unable to find another job demanding his full basket of skills, and may therefore receive lower wages. This observation reconciles the difficulty one
often has with describing skills that are truly firm-specific with the empirical facts of wage reductions on job loss and positive tenure coefficients.

While Lazear’s model focuses largely on a number of questions related to human capital theory — such as employees’ incentives to invest in skill and who pays for such investments — he also interprets the model as generating a match-specific productivity as in Jovanovich (1979a,b). Suppose there are two dimensions of skills, A and B, and let potential employee i’s endowment of skill A (B) be given by A_i (B_i). Suppose skill endowments are heterogeneous in the population of potential workers. Let potential employers be heterogeneous in their demands for skills, with the output generated by employee i working at firm j given by

$$\alpha_j A_i + \beta_j B_i.$$  

The right employee for firm j is one whose skill endowment (A_i, B_i) matches with the firm’s skill-weights (\alpha_j, \beta_j). A firms with a high \alpha_j and low \beta_j (relative to the population of firms) will hire employees with large A_i and small B_i. Such a firm would still find skill B to be of value, but its willingness-to-pay for this skill would be smaller than that of other firms in the market.

Lazear does not address the question of what gives rise to firm-level heterogeneity in skill-weights. Presumably these derive from firm-level differences in endowments of other factors of production or product-market strategy, but connections between, say, product-market differentiation and skill-weight-driven labor-market differentiation have yet to be drawn out. It may be fruitful to connect Lazear’s skill-weights approach to the different dimensions of worker skill as measured by Abowd, Haltiwanger, Lane, McKinsey and Sandusky (2007). A natural story for the lower relative demand for experienced workers by firms with large technology investments is that experienced workers have invested in specific skills that are made obsolete by the investments in new technologies. Drawing such connections would require detailed firm- and employee-level data on specific skills required in jobs and held by employees.

Some of the main empirical predictions of Lazear’s model — on market thickness and firm size effects in tenure coefficients in wage regressions — have yet to be examined by empiricists. Geel, Mure and Backes-Gellner (2009) use data from the German BIBB/IAB Qualification and Career Surveys to test some implications of the model for occupational training. They argue first that when the skill requirements of an occupation are more specific, firms should
bear a higher share of training costs. Second, they argue that more specificity in occupational skill-weights should be associated with a smaller likelihood of changing occupations. The Qualification Survey studied by Geel et al. (2009) offers detailed survey evidence on skills possessed by individuals in different occupations, which allows the authors to construct an index of occupational skill-specificity. Greater skill-specificity is associated with both a larger investment by firms in training, and lower across-occupation mobility after skills have been acquired.

**Risky Workers** Lazear (1998) argues that potential employees may vary not just in their skill — that is, the first moment of the distribution of their productivity — but also in the degree to which the employee is risky (that is, the second moment of the employee’s productivity distribution). He develops an equilibrium model where potential employees vary in terms of their riskiness, and derives predictions about which firms are good matches for risky workers. If firms can easily terminate risky workers whose productivity is revealed to be low and earn rents on those with high realized productivity, then hiring a risky worker has option value. Some barriers to mobility — either from direct turnover costs or employer private information about worker ability — must be in place in order to give risky employees option value. Given these ingredients, risky workers will be preferred to safe workers at a given wage. In equilibrium, wages adjust so that the marginal firm is indifferent between hiring risky and safe workers. Firms that expect to be in business for a long period of time value risky workers more, since they will be in position to earn the full stream of possible rents. Firms facing high turnover costs or low information barriers to raids will find it more attractive to hire safe workers. Lazear’s theory suggests that good matches for a given employer can depend on the second moment of employee productivity and firm-level characteristics such as firing costs, expected firm lifespan, and the degree of private information.

Burgess, Lane and Stevens (1998) examine one prediction of this model, specifically that firms with short expected time horizons will hire safe workers and therefore have low turnover. The authors use establishment-level data to relate firm-level churning flows — that is, changes in a firms’ workforce that are not accounted for by growth or contraction of the firm itself — to industry growth rate and mean firm age. Results suggest that firms in growing industries (as measured by industry growth rate) do indeed have higher churning flows. Firms in industries with older firms have lower churning. It is not clear from the analysis whether industries with
older firms should have longer or shorter expected future life; one can imagine that effect going in either direction.

Lazear also suggests that younger workers — who have less history in the labor market and therefore greater uncertainty about future productivity — might have greater option value. If so, then increases in termination costs will reduce this option value, and cause employers to shift demand away from younger workers. This argument is developed further by Oyer and Schaefer (2002), who extend it by connecting termination costs to data on how propensity-to-litigate varies with age for members of protected classes. Their empirical analysis studies how the Civil Rights Act of 1991 (CRA91) — which increased termination costs for members of certain protected classes — affected returns to experience. Unlike some prior affirmative action legislation, CRA91 had small aggregate wage and employment effects. However, CRA91 does seem to have changed the wage/experience profile for members of some protected groups.

Both Burgess et al. (1998) and Oyer and Schaefer (2002) are somewhat indirect as tests of Lazear’s hypothesis, because neither study is directly able to measure across-worker variation in “risk”. Indeed, it is difficult to imagine how one might do this in a standard employment setting. Two papers, Hendricks, DeBrock and Koenker (2003) and Bollinger and Hotchkiss (2003), use sports as a laboratory to examine the impact of uncertainty on hiring. These measurement benefits do come with a cost, however, since sports leagues often restrict within-league competition for players. Bollinger and Hotchkiss (2003) study baseball, where detailed statistics on player performance permit the explicit calculation of both the level and variability of expected performance. They show that players who have not yet become free agents earn a premium for riskiness — players with performance variability that is higher by one standard deviation earn 7% more, holding all else constant. It is not clear to us what the source of this premium is, however. In their sample, reserve clause players (those in the early years of their career) were governed by strict monopsony rules that limited bidding for players. In Lazear’s model it is competition that drives wages up for riskier workers, and this effect is necessarily absent for young baseball players. It is possible that their results are driven by some feature of the salary arbitration process, rather than the effects in Lazear’s model.

Hendricks, DeBrock and Koenker (2003) study the US National Football League (NFL), which has a number of features that make is a useful setting in which to test Lazear’s model. First, the NFL features a “draft” in which teams select, in a predetermined order, players
who have completed their college football careers. Second, players vary in terms of the quality of their college teams; Hendricks et al. (2003) use this variation to proxy for risk under the hypothesis that a player from a minor college team will not have faced strong competition during his college career, and therefore professional teams will have less information regarding productivity. Third, players are bound to teams for a fixed period of time after the draft and players can be terminated costlessly. The authors find that conditional on being selected in an early round of the draft, players from less prominent colleges tend to have better careers. In later rounds of the draft, the reverse is true, which suggests that teams might place a higher value on uncertainty in later rounds. The authors conclude that there is support both for Lazear’s model and for various forms of statistical discrimination.

Employee Preferences or Beliefs  The right employee might also be one who has the right beliefs or preferences. Under a standard agency model, all employers would prefer to have employees who are tolerant of risk and effort. It is also straightforward to show that employers will benefit from selecting employees who are intrinsically motivated to perform the task required for the job. But organization theorists have recently begun to develop models in which firms might heterogeneously demand employees with certain beliefs or preferences that differ from those of the firm’s owners.

Van den Steen (2005) starts by eliminating the common priors assumption that is standard in much of economic theory. He considers a firm that must first identify a project and decide whether to invest, when the state of the world is unknown. There are two potential states of the world, and the firm’s project will succeed only if it matches the actual state. An employee must choose one of the potential states, and then exert costly effort that increases the probability that the employee will identify a project. Direct pay-for-performance effort incentives are ruled out. If a project is identified, then a manager sees the (random) cost of implementing the project and decides whether to implement. If the project is implemented, then it succeeds if the employee’s initial choice of state was correct, and fails otherwise. Both manager and employee receive a deterministic benefit for project success.

Van den Steen shows that expected profits (relative to a reference belief) are higher when the firm hires a manager who is a “visionary,” in the sense that the manager has a much stronger prior about the true state of the world than does the firm. As an example, if a firm thinks
the states A and B are equally likely, it can still profit by hiring a manager whose prior is that state A obtains with probability 1. There are two reasons for this. First, hiring a visionary manager induces a beneficial sorting in the labor market. Firms with visionary managers will hire workers who agree with the manager’s vision. Second, agreement within the organization about the correct course of action is valuable because it encourages the employee to exert high effort in searching for a project. An employee who knows his manager agrees with him about the likely state will exert more effort looking for a project, because conditional on finding a project the likelihood of implementation is higher. Thus, the right employee for a firm to have is one who agrees with the vision laid out by the firm’s top management.

One implication of the Van den Steen model is that turnover of manager and subordinate should be temporally linked; a manager with a strong vision will attract a subordinate with similar beliefs, and these subordinates will not be the efficient matches for a successor manager who holds different beliefs. This implication is supported by the results of Hayes, Oyer and Schaefer (2006), who study top executive teams for evidence of complementarities among co-workers. Their main results are that the probability of non-CEO turnover increases markedly around times of CEO turnover, and that this effect depends on how long the CEO and non-CEO managers have worked together. Their evidence, however, is consistent with any source of complementarity among co-workers. Employees who have complementary skills, or invest in co-worker-specific human capital, or who simply enjoy working together will exhibit a similar pattern in turnover.

Prendergast (2007) begins from the premise that providing monetary incentives for “bureaucrats” is difficult. He observes that incentives for effort must then be provided through selecting the preferences of the bureaucrat, but that bureaucracies seem to differ in their selection. While some bureaucrats — teachers, firemen, and social workers — appear to be selected for preferences that are biased toward their clients, others — police and tax agents — appear to be biased against. Prendergast develops a model in which a social planner hires a bureaucrat to generate information used to assign a treatment to a client. As in Van den Steen’s model (where the agent’s incentives to develop a project are shaped entirely by his belief congruence with the manager), here the bureaucrat’s effort incentives are driven by his preferences regarding the client. Social surplus is highest when the bureaucrat puts more effort into acquiring information, so the key effect in the model is how bias influences effort choice. When client and
social interests are aligned (as in the case where the bureaucrat’s job is to assess whether cardiac surgery is warranted), a bureaucrat who is altruistic — that is, biased in favor of the client — exerts greatest effort. If client and social interests are not aligned, as in the case where the “treatment” is a prison sentence, then clients benefit from a less informed bureaucrat and hence an altruistic bureaucrat would exert too little effort. In this case, a bureaucrat who is actively hostile to the client’s interest can yield higher social surplus. Prendergast then shows that selecting the appropriate employees into a bureaucracy is likely to be difficult when potential employee preferences are not publicly observed. One possible outcome is bifurcated selection, where bureaucrats are either those most preferred by the principal or those least preferred by the principal.

Prendergast (2009) develops another model in which the inability to contract on output holds implications for the selection of workers. An institution carries out two tasks and employs two agents to do so. Each of the two workers is primarily responsible for one of the tasks, but contributes to the institution’s success with regard to both. An example is a university’s faculty and its administration. The university aims to both serve students and alumni and conduct research, and each party (faculty and administration) impact both. Incentives in the model come both from direct monetary payments from the institution to the agents and from career concern incentives. Notably, however, incentives from career concerns are biased toward one of the tasks. As the institution’s ability to contract on output falls, it substitutes toward the biased career-concern-based incentives. This yields a benefit in the form of higher effort, but generates costs in that the two agents will frequently fail to cooperate.

### 3.2.2 Inducing Self-Selection

The problem of hiring the “right” employee is further complicated by the possibility that employees may be privately informed about relevant personal attributes. Employees may have an incentive to misrepresent qualifications and overplay experience. Since Salop and Salop (1976), labor economists have thought about how firms might solve informational problems by inducing employees to reveal information prior to the hiring decision. Salop and Salop point out that compensation policy is one tool firms can use to induce self-selection. In their model, potential employees are privately informed about their exogenous short-run probability of quitting the job. Firms incur training expenses, so quits are costly and firms prefer to hire only workers with
lower quit probabilities. Firms can induce self-selection by asking employees to post a bond up front, and then making a payment to the employee that is conditioned on the employee remaining with the firm. See, also, Lazear (1979) on the role of mandatory retirement in settings where firms overpay relative to productivity late in an employee’s career. In general, firms will want to offer forms of compensation that are most valuable to the type of employee the firm wishes to attract.

Inducing self-selection is one of the leading explanations for the use of performance-based pay in organizations. As Lazear (1986, 2001) notes, employees who believe themselves to be productive will expect to earn larger payments in a pay-for-performance scheme. It is straightforward for firms to structure “incentive” compensation such that the participation constraint is met for a high-ability worker, but not for a low-ability worker. Such pay plans can be profit-maximizing even if there are no hidden-action problem. In his Safelite study (discussed earlier), Lazear (2000) shows that the implementation of a piece-rate incentive system was associated with an increase in the quality of newly hired workers. Janssen (2002) shows how compensation policy can be used to encourage employees to signal. He shows that firms may want to raise wages when job openings attract an excess supply of applicants. Raising the posted wage offer increases the return associated with being hired, and thus encourages the best of the potential applicants to engage in costly signaling.

More recent research has focused on how firms can induce self-selection on dimensions other than simple ability. Oyer and Schaefer (2005) argue that inducing self-selection is one of the leading potential explanations for the recent rise in stock-based pay for lower-level employees. Because lower-level employees have little, if any, impact on a firm’s share price, a grant of stock options to a lower-level employee cannot induce selection on ability. But if potential employees’ valuations of option grants are heterogeneous, the requirement that employees accept options as part of pay will tend to select those individuals who value the option grant most highly. Thus, option grants will tend to select employees who are less (a) risk averse and (b) more optimistic regarding the firm’s prospects. Firms can benefit from attracting such employees in a variety of ways; for example, employees who are optimistic about the firm’s prospects will tend to be those who agree with the “vision” proposed by top management (as in Van den Steen, 2005), and hence will be more willing to make investments that are specific to that vision.

Using survey data on actual option grants made by US firms, Oyer and Schaefer (2005)
show that the magnitudes of and risk-premia generated by these grants are consistent with this sorting explanation. For example, they show that a somewhat risk-averse employee who believes the firm’s stock will appreciate by 25% annually would prefer the observed option-plus-salary package to a cash-only compensation plan that yields the same net cost to the employer. Magnitudes of the option grants they observe are also consistent with a retention story of stock option use (see Oyer, 2004), but are hard to reconcile with an incentives-based explanation of stock-option usage. Oyer and Schaefer (2006) derive conditions under which options perform better than stock grants at inducing this selection. Bergman and Jenter (2007) note that optimistic employees can purchase equity-based instruments on their own accounts, and derive conditions under which the firm can extract some rents by making direct grants of options. They show that firms extract rents if employees prefer equity that has been granted by the firm to that sold in the market, or if the firm’s equity is overvalued in the market. Their empirical analysis suggests that firms may use options when boundedly rational employees are excessively optimistic about future share prices. Given the changes to the accounting treatment of stock options (in 2005) and the stock market crash of 2008, we think it possible that firms may change their stock-option-granting behavior going forward, and we encourage further research in that direction.

Delfgaauw and Dur (2007) study a model in which workers are heterogeneous in their intrinsic motivation to work for a firm. Offering higher wages increases the pool of potential applicants, but results in lower profits for the firm. When the firm can observe worker motivation, it faces a time-inconsistency problem. The firm wants to offer high wages to ensure a large pool of applicants, but then will be tempted to renege on this commitment (to grab rents) once it identifies a worker with high intrinsic motivation. A solution is to commit to a high minimum wage \textit{ex ante}. When the firm cannot observe worker motivation, paying a high minimum wage would induce workers with low intrinsic motivation to apply. In this case, the firm ensures a good match by offering a lower wage, but this leads to a higher likelihood of having the vacancy go unfilled.

\subsection{Labor-Market Intermediaries}

Autor (2009) summarizes a recent volume of collected papers on labor market intermediation, and argues that intermediation arises primarily to solve problems of costly search, asymmetric
information, or collective action. (We discuss some of the papers in this volume when discussing recruiting on the internet, below.) We think the field could benefit from continued research in this area, specifically in the area of how employers interface with intermediaries.

One stream of literature focuses on the role of temporary help firms as an intermediary. Autor (2001a) notes that temp agencies often provide free training in general skills such as computer software packages, in apparent violation of standard human capital theory. He argues that this training both induces worker self-selection and allows these firms to privately screen on worker ability. Temp agencies can then exploit this short-run information advantage about unobserved worker ability to recoup the costs of training. Notably, Autor reports that firms are increasingly using temp agencies to identify candidates for permanent employment. Thus, a firm hiring through a temp agency essentially outsources the screening function to a specialist intermediary. Despite this, Autor and Houseman (2005) report that low-skilled workers in Michigan did not find better permanent employment matches when placed initially with temp agencies.

One intermediary about which very little is known (at least by labor economists) is the executive search firm. While search firms surely account for a very small fraction of overall hiring, they are commonly used to fill important positions at the top of large organizations, both for-profit and non-profit. Bull, Ornati and Tedeschi (1987) offer a theoretical model of executive search firms, in which employers must undertake costly screening to determine whether a prospective employee is a good fit. In their model, executive search firms incur the same costs as employers when screening potential employees. Search firms have two advantages in this model: (1) they can diversify sampling risk, and (2) they can screen potential employees in advance, and therefore fill vacancies more quickly than the employer could do on its own.

It has now been more than twenty years since the preliminary step of Bull et al. (1987), and there has been essentially no follow-on literature in economics on executive search. While some economists have succeeded in getting data from prominent search firms about their businesses,

---

4Kaplan et al. (2008) study a related labor market intermediary — a firm that assesses the talent of candidates for top management positions. They show that certain skills are particularly valuable to Private Equity and Venture Capital firms choosing managers for their portfolio companies.

5There is some discussion of search firms in Simon and Warner (1992), which we discuss below, but they treat search firms as simply another source of employee referrals.
there has been little progress toward understanding the economic role of the search firm itself. Cappelli and Hamori (2006), for example, use detailed records from a search firm — including whether target executives contacted by the search firm have declined or pursued offers of employment — to assess factors that affect employee loyalty.

While it is certainly possible that search economies of the type considered by Bull et al. (1987) are valuable, our sense is that other factors play a more important role in explaining executive search. There are small literatures in management and sociology examining executive search. Khurana (2002), for example, conducts a series of interviews with search firm employees, CEOs who had been recruited by search firms, and directors who had participated in CEO selections. He argues that the CEO labor market is characterized by “few buyers and sellers, high risk, and institutionalized gaps between buyers and sellers,” and that these factors together give rise to intermediation. Buyers and sellers in this market, he argues, are often well aware of each other’s availability as a trading partner. That is, an aspiring CEO or business school dean may well know which jobs are open or likely to come open, and a board of directors or university provost may well know which candidates are looking. But Khurana argues direct communication between the parties is difficult, and the search firm plays an important role in legitimating the relationship. Finlay and Coverdill (1998) describe a number of “soft” dimensions on which search firms attempt to assess fit with client firm needs, while Beaverstock, Faulconbridge and Hall (2009) stress legitimization of potential trading relationships in their study of European search firms. A richer understanding of this process would help provide some nuance and realism to basic models of search and matching.

3.2.4 Firms Accessing Social Networks

For decades, social scientists have studied the role of informal social connections on labor markets (Granovetter, 1974). While it is by now well documented that social networks matter in labor markets, it is somewhat less clear why they matter. Do networks simply facilitate the search for job openings by potential employees? To what extent do network connections play a role in screening? To what extent does firm-level decision-making influence the process of matching through employees’ social networks?

To illustrate this question, consider the recent study by Bayer, Topa and Ross (2008), which uses detailed Census data on individuals’ precise Boston-area residence and employment
locations. They report that a pair of individuals who live on the same block are a third more likely to work together than individuals who live close to each other, but not on the same block. Thinness of local housing markets means it is hard to imagine this result is driven by a correlation in unobservables at the block level, so this study is a convincing identification of a local network effect on labor market matches. The study is unfortunately silent on the precise mechanism through which the neighborhood effects operate. Do neighbors simply mention job openings to each other when waiting at the bus stop? Or are the social connections between neighbors sufficiently strong that employers can rely on a current employee’s recommendation to hire a neighbor? If so, under what circumstances should an employer trust a current employee’s recommendation?

Many of the models of how social networks impact labor markets are of the “black box” variety, and focus solely on the role of the network in transmitting information about job openings. Calvo-Armengol and Jackson (2004), for example, use a network model of job search to offer an explanation for persistent race-based differences in labor market participation. Individuals randomly receive information about job openings, and can either act on those opportunities themselves or pass information to contacts. This mechanism gives rise to a positive correlation in employment status among individuals who are connected to each other. If there are costs associated with remaining in the labor force, those connected to a good network will be more willing to bear these costs than those who are not. See Ionnides and Loury (2004) for a recent survey of research on job networks and inequality.

We focus our attention on the part of the literature that features an active role for employers in accessing social networks. Saloner (1985) offers a model of a firm that uses an “old boy” network as a screening mechanism. Job seekers have heterogeneous ability, and are assumed to be unable to signal. Each employee has access to one of two “referees,” who privately observes a signal that is related to the job seeker’s productivity. Referees then communicate a hire/don’t hire recommendation to the employer. Saloner shows that despite the difference in objective between the employer and the reference — employers are profit-maximizers, while referees care about both placement rates and the quality of their contacts who are placed — the equilibrium features truth-telling and first-best hiring choices.

Montgomery (1991) offers a model in which a firm can screen by hiring the social contacts of its current employees. Current employees are randomly endowed with social ties to potential
future employees, and future employees are exogenously assumed to be likely to have the same type (productive or not) as the current employee. In the first period of this model, employers hire with no knowledge of worker productivity, so the market clears at a single wage. Employers then learn the productivity of their current employees and can make wage offers to potential employees who are tied to current employees. Firms with good-type first period employees make high wage offers to the ties of those employees, and earn profits due to some informational monopsony power.

Casella and Hanaki (2006, 2008) move this literature — in what we think is a very productive direction — by explicitly considering two channels through which firms can hire, and modeling job-seekers’ incentives to reach employers through each channel. They begin with the Montgomery (1991) model, and endogenize the formation of networks. At cost \( \lambda_N \), a young worker can establish a connection with an older worker. As in Montgomery, young workers are disproportionately likely to establish links to older workers of the same type: conditional on forming a link, the probability that the two connected workers have the same type is \( \alpha_N > \frac{1}{2} \). Unlike Montgomery’s analysis, however, employees who don’t form a network link still have options. At cost \( \lambda_S \), workers can attempt to signal — by, perhaps, acquiring a credential such as schooling. High (low) types succeed at acquiring the signal with probability \( \alpha_S (1 - \alpha_S) \). Importantly, both networks and signals have the potential to assist the employer in screening worker types.

Casella and Hanaki then ask where employers will prefer to look for workers. For the case where networking is free (\( \lambda_N = 0 \)), they find that referral-based hiring is almost always preferred by employers, even when, in equilibrium, certification is more informative. The key intuition here is simply that referrals allow the firm to privately screen, while certification is, by nature, public. Certified job-seekers are known by all employers to be disproportionately likely to be high ability, and therefore their wages are bid up and all rents flow to the worker. Referred job-seekers are somewhat less likely to be high ability than certified in this case, but because of the firm’s informational monopsony it is able to capture rents.

When both forms of information transmission by workers are costly, the workers will compare the expected informational rent from pursuing either to the up-front cost. Networking always necessitates rent-sharing with the firm, so workers only pursue it if it is either less costly or if it is more precise as an indicator of ability. We see, in the work of Casella and Hanaki, the seeds of
a theory of firm-level hiring strategy. Models of hiring to date have tended to focus on a single hiring practice and examine its efficacy. But real firms face a portfolio of choices over how to recruit. If we are to understand firm-level variation in hiring strategy, we need more models in which firms must choose how to access the labor market. We believe that future work along these lines can proceed in a number of directions. It would be useful, first, to expand the range of hiring venues available to firms. Second, researchers will need to introduce various forms of firm-level heterogeneity (both observable and unobservable). Third, ability is completely general in the Casella and Hanaki model, and it would be useful to know how match-specificity of ability impacts hiring channel choice. Fourth, search is missing from this model. Finally, it would be useful to understand to what extent rents associated with good employment matches can be captured by the referring employee.

The empirical literature on firm-level choices when accessing employees’ social networks is still in its infancy. Most large, nationally representative datasets contain little information that could be used to do within-firm comparisons of employees hired via different means. This makes it difficult to rule out unobserved firm effects as an explanation for differences in wages across hiring method. And studies of firms’ personnel records — often performed by sociologists — suffer from potential limits on generalizability. Much work remains to be done in this area, but we next survey the extant research.

Simon and Warner (1992) develop a search model of referrals from various sources (employees among them). Their model assumes the role of referrals is to reduce the employer’s *ex ante* uncertainty about worker productivity. Workers exogenously receive an offer either through a referral or through non-referral means. Because the better information results in better *ex ante* matches, Simon and Warner predict that referred workers will receive higher up-front wages, have lower wage growth (conditional on continued employment), and lower turnover probabilities. Support for all three assertions is found in data from the 1972 Survey of Natural and Social Scientists and Engineers. However, the data does not permit any within-firm analysis of different hiring practices, so it is difficult to rule out the hypothesis that differences in the firms hiring via referrals (rather than the referral method itself) is driving the results.

Kugler (2003) builds efficiency wages into a model of referrals, and derives an equilibrium in which industries that pay efficiency wages prefer to hire through referral, both because connected workers can engage in peer monitoring and because markets where referrals are
important are likely to be thin. Empirical evidence suggests that high-wage sectors do tend to engage in more referral-based hiring, and individuals hired by employee referral earn higher wages. Again, however, it is difficult to rule out various forms of firm-level heterogeneity as an alternative to the efficiency wage explanation.

Mosca and Pastore (2008) study the personnel records of organizations — public, private for-profit, and private non-profit — that provide social services in Italy. Unusual in this literature, they find that being hired through informal networks bring a substantial wage penalty (6.5%) for employees hired to public agencies. Interestingly, those hired through “public competitions” — which, presumably, could play the same role as signal in the Casella and Hanaki model — earn a 7 to 32% wage premium. Differences in hiring method account for a large fraction of the overall variation in wages across organization types.

Antoninis (2006) studies the personnel records of 209 employees of an Egyptian manufacturing firm. Conditional on observables such as experience and education, workers hired on the recommendation of someone holding direct knowledge of the new employee’s productivity (an old work colleague) earned higher wages, on average, and this effect was larger for employees hired into higher skilled jobs. New workers hired on the recommendation of friends or family received no wage premium, and may have taken a wage discount in lesser skilled occupations.

Organizational sociologists have conducted some excellent single-firm studies of social networks in hiring practices. Fernandez and Weinberg (1997) study hiring processes used for four entry-level jobs at a retail bank in the early 1990s. Because the firm rewarded current employees for referring new hires, the source of initial contact between the firm and each job candidate was carefully tracked. Fernandez and Weinberg (1997) document several ways in which referred candidates differ from non-referred. First, referred candidates were more likely to fit the skill profile desired by the firm. Second, referred candidates disproportionately applied for jobs where there were fewer applicants, which suggests the firm was relying on its employees’ networks more heavily when it was having trouble drawing applicants through other means. Third, referred candidates are both more likely to get interviewed and more likely to receive a job offer, even conditional on observables. Thus, it appears that this firm is using employee referrals in matching, search, and screening. It would be interesting to know whether the firm still uses referrals extensively when it is not having trouble finding applicants; presumably search costs have been affected by the possibility of internet recruiting for entry level positions like these.
Fernandez, Castilla and Moore (2000) study similar data from a phone center. Strikingly, they find that new hires identified through a referral are no less likely to turn over than hires identified through other means, suggesting that better job matches are not uncovered through referrals. Employees hired through referrals are, however, more likely to turn over when the referrer turns over, suggesting that social connections matter both pre- and post-hiring. Fernandez et al. (2000) close with a call for a greater dialogue between economic theory and case-study-based empirical work, which we endorse.

3.2.5 Employer-to-Employer Transitions and “Raids”

To what extent do firms look to other firms’ employees — rather than the pool of unemployed job seekers or those just entering the labor market — to fill vacancies? And what factors facilitate and impede employer-to-employer job transitions?

Employer-to-employer job transitions appear to be fairly common. Fallick and Fleischman (2004) use the Current Population Survey’s dependent interviewing techniques to measure the extent of employer-to-employer job transitions in the United States. They report that 2.6 percent of employed persons change jobs each month. This figure is comparable in magnitude to the number of people moving from employment out of the labor force, and is twice as large as the number of people who move from employment to unemployment. Fallick and Fleischman further report that almost 40% of new jobs started between 1994 and 2003 were employer-to-employer transitions, a figure that supports the importance of on-the-job search.

The efficacy of “raids” as a source of new employees depends on two main factors. First, to what extent is employee productivity firm-specific? In settings where firm-specific matching or skill acquisition are important, we would expect to see less employer-to-employer movement of employees. Second, how does the labor market learn about the productivity of employees? If learning is symmetric — in that the current employer and potential future employers observe the same information regarding worker productivity — then raiders can bid for other employers’ workers without fear of a winner’s curse. If learning is asymmetric, then informational problems can impede employer-to-employer worker flows (Greenwald, 1986).

Models of symmetric employer learning have been used to explain a number of facts about wage growth over time. Farber and Gibbons (1996) develop a model in which education or other employee characteristics observable to labor market participants convey information about
worker productivity. Employers also learn about productivity over time by observing output, and update accordingly. Consistent with the model, they show that time-invariant unobservables (the AFQT, for example) explain more of the variability in wages as workers gain labor market experience. Altonji and Pierret (2001) enrich this framework by showing that if firms statistically discriminate using education, then education should hold less explanatory power for wages as workers gain experience. This assertion is supported in NLSY data.

Lazear (1986) considers a model of raids under symmetric learning. An incumbent firm hires a worker with ex ante unknown ability to a job with a downward rigid wage. The worker’s productivity at the incumbent firm is revealed. The worker’s productivity at a potential raider is given by his productivity at the incumbent plus a random positive or negative shock. The main result is that it is the good workers who are raided. This is because raids happen under two conditions: (1) the worker’s productivity is higher at the raider than at the incumbent, and (2) the worker’s productivity at the raider is higher than the worker’s wage at the incumbent.

Tranaes (2001) allows firms to endogenously select whether to attempt to hire a currently employed worker or hire from the pool of unemployed. He assumes symmetric learning — all potential employers can observe the abilities of all employees — but assumes that it is not possible for firms to observe abilities of workers who are not currently employed. Hiring an employed worker is costly because the firm must offer a higher wage than the incumbent, but hiring from the pool of unemployed is likely to yield a worker of low quality. Equilibrium in the model features unemployment even for good workers (separations happen exogenously). As raiding becomes more difficult (proxied by an exogenous friction), unemployment falls as hiring from the pool of unemployed becomes more attractive. Because raids impose a negative externality on unemployed job searchers, social welfare is strictly improved if raids are prohibited. Note the model does not included any gains from matching.

The literature on asymmetric learning has focused to a large extent on how incumbent employers might exploit their informational advantage. Waldman (1984) considers how the desire of an incumbent firm to preserve its informational advantage can distort job assignments and promotions within a firm. In Waldman’s model, rival employers cannot observe a worker’s productivity directly, and instead make inferences about productivity from job assignment. The main findings are that the information asymmetry causes the incumbent firm to tie wages to jobs rather than productivity and also to distort job assignments away from first-best.
The basic insight that asymmetric learning can distort choices inside organizations has been applied in a variety of theoretical models. Milgrom and Oster (1987) show that asymmetric learning can distort incentives for human capital investments for workers with high unobserved ability. Bernhardt (1995) applies Waldman’s idea to promotion decisions, and compares the cost (from inefficient job assignment) and benefit (from reduced labor market competition) of delaying promotions. This cost is higher for more able workers, which may explain fast-track promotion paths in organizations. Scoones and Bernhardt (1998) show that if wages are attached to promotions (due to asymmetric learning), then employee human capital investment incentives may be distorted toward skills that will lead to promotions. If a general skill investment is efficient, employees may prefer to make a firm-specific investment if doing so is the surest path to promotion.

Of greater interest here is how asymmetric learning affects strategies pursued by raiding firms, and research on this question is somewhat more limited. Waldman (1990) shows how up-or-out promotion decisions — where a retention decision signals high productivity — can provide incentives for general-purpose human capital investments even if such investments are not directly observable to raiders. In Waldman’s model, a crucial role of the retention decision is to spur labor-market competition from raiders, as this improves \textit{ex ante} incentives. Bernhardt and Scoones (1993) consider a raider who must make a decision about whether to invest in learning whether an incumbent firm’s employee is a good match. The incumbent can deter such an investment (and hence deter wage competition in the event the raider determines the employee is a good match with the raider) by making preemptive wage offers.

Empirical evidence on symmetric vs. asymmetric learning is, perhaps not surprisingly, broadly in support of the notion that both models hold some explanatory power, depending on context. As noted above, Farber and Gibbons (1996) and Altonji and Pierret (2001) offer evidence consistent with symmetric learning. DeVargo and Waldman (2009) show that promotion decisions in the Baker, Gibbs and Holmstrom (1994) data appear to conform to signaling theories stemming from asymmetric learning; specifically, they show that wage increases associated with promotions are smaller when worker education levels are higher. Schonberg (2007) develops a learning model with endogenous mobility that allows for both forms of employer learning. She argues that under symmetric learning, job movers and stayers should have the same average ability (as proxied by AFQT), while movers should be lower ability under asymmetric learning.
Further, asymmetric learning suggests that education should better explain wages for movers than stayers (because the raider lacks the incumbent’s direct observations of productivity). She estimates the model using the NLSY data studied by Farber and Gibbons (1996) and Altonji and Pierret (2001) and finds support primarily for symmetric learning, with some support for asymmetric learning for more educated workers. Kahn (2009) also uses the NLSY but identifies the relative importance of symmetric and asymmetric learning using layoffs, economic conditions at the time of starting a job, and differences across occupations. Like Schonberg (2007), Kahn (2009) finds evidence of both types of learning. However, in contrast to Schonberg (2007), she concludes that asymmetric learning is economically and statistically more important than symmetric learning. Pinkston (2009) offers a model in which raiders receive noisy private evaluations of potential hires. Because raiders receive a signal the incumbent employer does not, the raider can profitably bid for the employee even in the absence of matching. The key empirical prediction of the model is that as experience increases wages reflect evidence of public learning, while as job tenure increases wages reflect private learning. Analysis of the NLSY supports both forms of learning.

Our view is that this literature needs industry studies, of the type commonly seen in industrial organization economics. While important progress has been made on understanding employer learning using aggregate data sets like the NLSY, it is likely the case that different markets vary dramatically in the extent to which asymmetric learning and matching are important. It is not hard to imagine that hiring another firm’s CEO would present a markedly different set of issues from hiring another firm’s retail clerk. We think the broad outlines of the raiding problem are well understood, but what is needed is an understanding of how matching and learning play out in specific labor markets. There are studies of a few such markets — which we review next — but much more is needed.

Fallick, Fleischman and Rebitzer (2006) use the CPS to examine how employer-to-employer flows vary by industry and geographic region. Specifically, they examine interfirm mobility of college educated men working in the computer industry, and find the rate of such flows to be significantly higher in California than elsewhere. Employer-to-employer transitions among this group average 1.95% monthly nationwide, but are around 3% in California. The authors note that rates of employer-to-employer job changes are no higher in Silicon Valley than in the rest of California, which suggests that California law (which does not permit the enforcement of
noncompete clauses) may play a role in explaining this effect. Notably, employer-to-employer transitions outside of the computer industry are no higher in California than elsewhere in the US. Thus, it appears that it is the interaction of California law with specific features of the computer industry that drives this higher mobility.

There is a reasonably large literature on across-firm mobility of Chief Executive Officers. One advantage of using this setting as a laboratory is that in the US both pay and performance for these employees are publically observed, albeit noisily. Murphy and Zabojnik (2006) document a striking trend in employer-to-employer flows in this market. For firms appearing in Forbes compensation surveys, between 1970 and 1979, just 14.9% of newly appointed CEOs were hired from outside the firm. This figure rose to 17.2% in the 1980s, and 26.5% in the 1990s. Murphy and Zabojnik develop a model in which managerial technology became more general over this period. That is, due to standardization of various accounting, management and IT practices in the US economy, it became easier for outsiders to manage large US firms. An associated increase in wage competition for the most able managers is offered as an explanation for the large increases in CEO pay over the same period.

Fee and Hadlock (2003) study a sample of managers who were hired from outside to either the CEO position or a senior non-CEO managerial role. They report that the prior employers of outside CEOs tended to have above-average stock price performance prior to the manager’s departure, as measured by stock returns. This pattern does not hold for outsiders hired to non-CEO positions. They also report that the existence of an “heir-apparent” manager at the incumbent firm — a non-CEO manager who has been tagged as the likely next CEO — increases the likelihood that other managers at that firm will depart for better employment opportunities elsewhere. Huson, Malatesta and Parrino (2004) document that outsider CEOs are associated with larger subsequent improvements in operating performance.

Hiring from outside can also have implications for the incentives of current employees. Chan (1996) considers the effects of this choice in the context of promotion tournaments. External recruitment reduces the likelihood that an insider will win the promotion tournament, and hence weakens incentives. A firm can restore incentives by increasing the promotion-based wage differential, but this may lead to increased rent-seeking efforts on the part of insiders. Alternatively, the firm can commit to hire from outside only if the external candidate is substantially better than the insider. Thus, there is a tradeoff between attracting good candidates from outside
and *ex ante* effort incentives for insiders. Agrawal, Knoeber and Tsoulouhas (2006) provide empirical evidence consistent with the hypothesis that firms hire external CEOs only when the external candidates are considerably superior to internal candidates on observable dimensions. The findings of Hassink and Russo (2008) run counter to this, however. Using matched employee-employer data from the Netherlands, they find that candidates hired internally into open jobs earn a 15% wage premium (conditional on observables) compared to those hired from outside.

Securities analysts — stock market observers who offer forecasts of corporate earnings — offer another market where employer-to-employer job mobility can be easily tracked. This setting has both advantages and disadvantages relative to the CEO context: analysts wages are not commonly disclosed, but it is relatively easy to devise measures of employee performance that can be compared pre- and post-mobility. Groysberg, Lee and Nanda (2008) find that star analysts who change jobs show a long-lasting reduction in job performance. This suggests that, despite claims of industry observers, there may be substantial firm-specificity in analyst skills that is lost upon job mobility. It is also possible that this is evidence of a winner’s curse stemming from asymmetric learning. It is not clear how this set of facts is consistent with equilibrium behavior by market participants, unless there is some gain that offsets the losses due to reduced performance.

### 3.2.6 Employer Search

While there is an enormous literature on employee job search, the literature on search by employers is less developed. An early contribution by Rees (1966) focuses attention on two channels of search. Employers can expand search on the extensive margin, by gathering more applications. On the intensive margin, employers can expand search by gathering more information on each potential applicant. More recently, researchers have examined choices over the range of search activities a firm may engage in.

Barron, Bishop and Dunkelberg (1985) study the 1980 Employer Opportunity Pilot Project (EOPP) survey, for which employers who had recently hired were asked about staff hours spent on recruiting, the number of applications received and interviews conducted, and wages for the new hire. The paper is perhaps best known for a simple but striking fact: 90% of the job offers made by firms in this sample were accepted. The authors then relate these search choices to firm
and job characteristics. A primary result is that firms spend more time on the search process when the job requires larger training expenditures by the firm. Human capital theory suggests firm-specific training generates rents, so firms will want to insure a good match prior to making such investments. Firms also spend more time recruiting when the educational requirements for the job are greater. Jobs that feature higher training expenditures by the firm are associated with more hours spent interviewing. Finally, firm size appears to be correlated with recruiting expenditures; a doubling in firm size increases the number of applicants per offer by 10%.

Another early contribution is Holzer (1987). Again using the EOPP data (this time from the 1982 wave), he shows that advertising openings to current employees is the most commonly used recruitment method.\textsuperscript{6} Interviews and reference checks are overwhelmingly popular as screening mechanisms. There is some evidence that employees hired through personal referrals have higher productivity, lower turnover, and lower screening costs.

There is a fairly sizable empirical literature on vacancies. The approach taken here is from search theory. Employers sequentially sample from a pool of potential employees with a known ability distribution. A vacancy duration is simply the amount of time until the vacancy is filled. Typical empirical design involves estimation of a hazard rate of vacancy filling, and examining how this rate varies with characteristics of the job or firm. Data sources are mostly from one-off surveys, which makes it difficult to examine how the mechanism by which the vacancy is filled is related to job tenure, wage growth, turnover and other measures of match success one would like to examine.

In a series of papers, van Ours and Ridder (1991, 1992, 1993) examine a pair of Dutch surveys on vacancies. van Ours and Ridder (1991) examine data from the Dutch Bureau of Statistics and study how vacancy durations and vacancy flows — defined as the rate at which new vacancies are created — varies over the business cycle. The authors estimate the vacancy duration hazard rate and find that vacancies fill more slowly when the vacancy flow rate is higher. Jobs that require more education also fill more slowly. Further, vacancy flow is more sensitive to the business cycle for low-education openings.

Using data from the Organization for Strategic Labor Market Research, van Ours and Ridder (1992) again find that higher educational requirements are associated with longer vacancy

\textsuperscript{6}It is not clear in this entire literature whether the new hires are all external hires. If a firm advertises to its current employees, the firm may want them to refer friends or apply for the job themselves.
durations. Vacancies also exhibit positive duration dependence. The main contribution here, however, is to examine the timing of applications. van Ours and Ridder (1992) show that the acceptance probability (that is, the probability a candidate accepts a job offer) is zero for the first two weeks after the vacancy’s posting. The probability rises after the first two weeks, and stays high for another ten weeks. The arrival rate of applicants is very high for the first two weeks, essentially zero for the next two weeks, before rising again. The evidence points to a non-sequential search strategy, where firms cast a net, draw in a pool of applicants, simultaneously screen, then make an offer or draw again. Thus, while employee search appears to be sequential, employer search is not. van Ours and Ridder (1993) find that employers spend far more time on selection than on search. Mean application period — estimated off the hazard rate of applications — is 3.1 weeks, while the mean selection period is 14.6 weeks (21.1 weeks when university education is required). Thus, vacancy durations should be thought of not as the duration of search but rather as a combination of search duration and screening duration. This raises the question of whether the internet has affected screening times; presumably the process of screening on observables is considerably faster now than it was in 1992. Abbring and van Ours (1994) also partition vacancy durations into search and selection period, and find that the ratio of unemployed to vacancies affects search durations — longer when the ratio is smaller — but not selection periods.

Barron, Berger and Black (1997) develop new theory on employer search when screening expenditures (intensive search) are endogenous. They predict that vacancies for jobs requiring more training will see both more applications before an offer is made, and greater screening expenditures per application. Analysis of four data sets from the US supports this contention. Barron et al. (1997) also show that job offers are rarely rejected, and that there seems to be no pattern in the data on rejected offers.

Burdett and Cunningham (1998), studying the 1982 EOPP data, echo many of the findings of Barron et al. (1985). Firms undertake longer searches when the training expenditures associated with the job are higher. Larger firms fill vacancies faster, presumably because of easier access to pools of employees. There is some evidence that jobs with high skill and/or education requirements take longer to fill. The primary distinction between the results in Burdett and Cunningham (1998) and Barron et al. (1985) is that the later study examines vacancy duration as its measure of selectivity in recruiting, while the earlier study used surveys of staff hours.
Burnett and Cunningham go on to report evidence of a non-monotonic hazard function for filling vacancies. When a vacancy is “young,” the hazard function is increasing, but as the vacancy gets old the likelihood of filling it drops. Finally, the authors report that vacancy durations are longer when the employer has advance notice that the vacancy will occur. Brencic (2009c) confirms this advance notice finding using detailed data on all employment vacancies in Slovenia, and Brencic (2009a) finds (using Slovenian and US data) that firms also adjust hiring standards when advance notice periods are expiring.

Another series of papers examines the Dutch “How Do Firms Recruit?” survey (conducted for the Dutch Ministry of Social Affairs). Gorter, Nijkamp and Rietveld (1996) study the relation between recruitment strategy and vacancy duration. They first estimate a multinomial logit model of first recruitment method choice — advertisements, informal channels, labor office, or other — and then estimate vacancy duration. Advertisements are used more heavily when the vacancy stipulates work experience, and the labor office is commonly used for low-education jobs. Hazard rates for vacancy duration show that when informal recruitment is the first choice, the vacancy tends to fill immediately or not at all. Further, advertised vacancies fill more slowly when education is required. Russo et al. (2000) use the same data set to examine recruitment method choice and the rate of applicant arrival. They report that the flow of applicants is, not surprisingly, related to overall labor market conditions. Further, when conditions are tight, employers adjust recruitment methods to use additional advertisements. This method of recruitment is shown to have the largest impact on the flow of applicants Russo et al. (2001) build on this work to show that firms are less likely to hire currently unemployed workers when conditions are tight. Finally, Van Ommeren and Russo (2008) argue that sequential search implies that the number of rejected applicants should be proportional to the number of vacancies. As an example, suppose a firm searches sequentially to fill a single vacancy, and must perform a costly screen on each candidate. If one-third of sampled applicants pass the screen, then on average the firm will reject two applicants for each one it hires. This reasoning implies the elasticity of the number of rejected applicants to vacancies should be one. If, on the other hand, search is sequential, then there is no requirement that this elasticity be unity. This observation forms the basis of a test. Sequential search is not rejected for informal methods (social networks, school recruiting, and temp agencies), but is rejected when firms access employment agencies and run advertisements.
Manning (2000) studies proprietary data from a small set of firms employing low-wage workers, and reports three main facts. First, most vacancies generate few applicants (less than three, on average). Second, wage and non-wage aspects of the job influence the number of applicants. Third, firms are more likely to grant an interview to workers who are currently employed, but conditional on making it to the interview stage, currently unemployed workers are no less likely to receive a job offer. A small literature examines the duration of vacancies.

Andrews, Bradley, Stott and Upward (2008) note that while employees infrequently reject job offers, employers often allow vacancies to go unfilled, simply by withdrawing the vacancy from the market. Using detailed data on vacancies from the UK’s Lancashire Career Services Agency, they show that the hazard rate of filling a vacancy exhibits positive duration dependence, while the hazard for withdrawal exhibits negative duration dependence.

Brencic (2009b) interacts the vacancies literature with that on Employment Protection Legislation. She reasons first that firms facing costly search — measured by the immediacy with which a vacancy needs to be filled — may respond by relaxing hiring standards. Doing so, however, can result in poor matches and hence greater likelihood of termination. This implies a complementarity between relaxed hiring standards and temporary employment, especially in cases where firing is costly. Brencic examines this relation using very detailed data on vacancies in Slovenia. During her sample period, there is no variation in firing costs — although she reports that changes in these costs occurred just after her sample period — so the test simply examines the cross-sectional association between costly search, relaxation of hiring standards, and temp employment. She finds that employers facing costly search tend to relax hiring standards, but only when the position being hired for is temporary. She does not find evidence that employers switch permanent positions to temporary when reducing hiring standards, as theory might suggest.

DeVaro (2005) adds both a new data source and an emphasis on starting wages. He studies the Multi-City Study of Urban Inequality (MCSUI), a large cross-sectional survey of employers in four metropolitan areas of the United States. He first documents that recruitment strategies vary with firm and job characteristics. As prior literature finds, firms rely more heavily on referrals when hiring for professional occupations. Next, he includes recruitment method in a standard wage regression. While Casella and Hanaki (2006, 2008) would suggest that personal referrals should be paid less conditional on human capital observables, there is little evidence
to suggest this is the case. In his richest specification, (personal, firm, and industry controls), DeVaro finds that referrals from friends and current employees are associated with small wage premiums relative to those recruited via help-wanted signs, but there is essentially no statistical difference between personal referrals and newspaper ads or school referrals. Clearly there is work to be done to reconcile screening-based theories of network hiring with the facts.

DeVaro and Fields (2005) use the MCSUI as well, and regress worker performance (on a 100-scale, employer-reported) on indicators for recruitment and screening method, and firm characteristics. Data show very little support for the assertion that recruiting and screening methods are related to performance. Conditional on firm, worker and job characteristics, method of recruitment bears little relationship to subjectively assessed performance. The approach — using employer-evaluated performance as a dependent variable — is subject to some criticism. Presumably employers would be pleased to hire a less able worker at a sufficiently discounted wage. It is not clear how one might combine wages with this numerical performance score to get the net impact on the firm’s bottom line.

DeVaro (2008) uses the same data to estimate a dynamic, discrete-choice structural model of recruitment. The firm is modeled as choosing recruitment strategies and wage offers over time to try to fill a vacancy. The primary advantage of this structural approach is that it permits the study of the effects of various counterfactuals on firms’ recruitment policies. It does come at a cost of some generality, however. Here a firm is modeled as choosing either formal or informal recruiting methods at the beginning of the recruiting period, but Holzer (1987) and others show that firms tend to use multiple recruiting methods simultaneously. The analysis suggests that, due to higher offer-acceptance probabilities, the firm will tend to offer lower wages when engaged in informal recruiting. Note that this is a different rationale for the informal-recruitment/lower-wages channel than that offered by Casella and Hanaki (2006, 2008). Counterfactual simulations suggest that wage subsidies, “information policies” that improve match qualities through formal methods, and changes in the degree of employee heterogeneity can have important effects on firms’ recruitment decisions and wages. Wage subsidies shift firms toward informal means, in part because the subsidy makes employment more attractive and employees identified through informal means are (assumed to be) more likely to accept. Information policies, modeled on the Workforce Investment Act of 1991, push recruitment toward formal channels to take advantage of improved matching. Increases in employee heterogeneity shift recruiting to formal methods.
through an order-statistic effect. Formal methods allow quicker sampling and reviewing many applications quickly and selecting the best is most valuable when the variance of match quality is high.

Finally, some papers consider a potential complementarity between various recruitment and screening strategies. Under the complementarity hypothesis, one might expect positive interaction effects between screening and recruitment methods. In the MCSUI data, DeVaro and Fields (2008) find very little support for the complementarity hypothesis. Interaction terms in performance regressions are typically negative.

Bartling et al. (2009) argue that choice of recruitment method — screening, in particular — can have complementarities with discretion, rent-sharing, and wages. In an experimental setting, they document the endogenous emergence of two markedly different organizational design strategies. One focuses on control, with little discretion, no rent sharing, low wages, and no screening, and one focuses on trust, with discretion, high wages, rent sharing, and screening. A key driver of this choice is the information available to employers about potential employees pre-hiring. When employers can observe a signal about employee past performance, many employers conditioned wages on this signal; this gives employees a reputational incentive to reciprocate employer trust, which facilitates discretion. When no such signal is available, employers cannot condition wages on past performance — so no career concerns operate — and employers tend not to trust.

### 3.2.7 Recruiting on the Internet

As rates of Internet adoption rose in the 1990s, participants in labor markets began to experiment with ways to use this tool in economic activity. Early attempts in this direction are summarized by Autor (2001b), who hypothesizes a number of ways that the Internet might affect labor markets, including the processes through which employers and employees match. He points out that the Internet may improve aggregate match efficiency by reducing search costs, thus allowing both firms and workers to consider many potential trading partners much more quickly than before. This effect may be ameliorated by concerns about adverse selection. Because electronic communication makes it easy for workers to apply for many jobs — even jobs they (privately) know themselves to be poorly suited for — recruiting on the Internet may lead to sharply increased screening costs for employers. Thought of another way, applying for
a job imposes some match-specific costs on the applicant. As the internet has lowered this application cost, the signal generated by the application process (as in the Spence (1973) model) has become less meaningful and the application process is less good at inducing applicants to self-select efficiently.

The years since 2001 have seen a trickle rather than a flood of research about online job matching. Much of the existing research focuses on search by potential employees, rather than on recruiting by employers. Kuhn and Skuterud (2004), for example, use the Current Population Survey Computer and Internet Supplements to examine selection into internet job search and unemployment durations among online job searchers. They find that internet search is associated with lower unemployment durations, but that this effect is entirely explained by worker observables such as education and occupation. Once these observables are controlled for, internet job search is associated with similar, or in some specifications, longer unemployment durations. Kuhn and Skuterud conclude that either Internet job search is ineffective at reducing unemployment durations, or that the pool of Internet job searchers is adversely selected.

Stevenson (2009) points out that 22% of workers who began a new job in mid-2002 cited the Internet as the primary means through which they found the job. She further reports that state-level Internet penetration is associated with a reallocation of job search activity, and an increase in the overall level of job search. Finally, Stevenson documents that most Internet job-searchers are employed, and that employed workers who search online are more likely to experience an employer-to-employer job transition than those who do not search online.

Bagues and Labini (2009) analyze the impact of the Italian AlmaLaurea online job board, and exploit the fact that different universities joined the board at different times to construct a difference-in-differences estimator of the impact of the online board on labor market outcomes. Notably the AlmaLaurea is unlike the US-based internet job boards studied by Stevenson (2009) in that it provides employers with detailed records — supplied by the colleges not the students — on students’ academic careers. This compulsory disclosure may serve to mitigate adverse selection in the labor market. Bagues and Labine find that the AlmaLaurea reduces the likelihood of unemployment by 1.6 percentage points, increased wages by 3 percentage points, and also increased the likelihood of regional mobility for university graduates. The study is silent, however, about whether the job board improves employers’ ability to find high ability workers or those with high idiosyncratic match values. Understanding this distinction has important im-
lications for the effects of internet hiring on inequality and whether some potential employees will be made worse off.

Some researchers have studied the question of how firms incorporate the internet into recruiting strategies. Hadass (2004) uses proprietary data from a US-based multinational firm to examine the impact of internet recruiting on various aspects of job match. Online recruiting at the firm accounted for just 0.2% of the firm’s hires in 1996, but grew to 20% by 2002. A Cox duration model shows that Internet recruiting leads to job durations that are statistically identical to that found for print advertising, but durations that are significantly (both statistically and economically) shorter than those for employee referral and college recruiting. Employee referrals and college recruiting lead to job durations that are 1.7 times as long as those found for internet and print.

Nakamura, Shaw, Freedman, Nakamura and Pyman (2009) offer a descriptive discussion of modern e-recruiting services, and give a list of five key facts. First, they point out that the main commercial jobsites are not standalone corporate entities. Monster, CareerBuilder and HotJobs are all parts of larger firms that engage in multiple activities. Second, online recruiting allows firms to search more widely and consider a larger variety of applicants. Third, there are substantial returns to scale in online recruiting. The costs to a large employer of advertising to find ten employees of a given type is not substantially larger than the cost of advertising to find one. This fact would seem to explain why large US retailers account for such a large fraction of overall internet job search traffic. (Target’s career site by itself accounted for more internet traffic in early 2007 than the site that hosts all US federal government jobs.) Fourth, in line with the results in Stevenson (2009), online methods allow firms to access currently employed workers who are only passive job seekers. Fifth, US-based firms currently dominate online job search worldwide.

Brencic and Norris (2009) focus on some specific choices facing employers when posting job openings online. They collect a sample of job postings made to Monster.com and compare job openings that the employer reports must be filled immediately to openings where employers do not make this statement. When employers report greater impatience, they also tend to list more information about the job application process and less information about hiring requirements. Further, openings for jobs that need to filled immediately tend to be withdrawn from the job board more quickly.
Finally, Kuhn and Shen (2009) use online job postings in China to provide direct evidence on employer preferences for gender, age, height, and beauty. Labor-market discrimination based on such characteristics is illegal in much of the world, but, strikingly, remains legal in the world’s largest labor market. Surprisingly, online job postings in China often contain explicit requirements on such “US-prohibited” characteristics: 90% of firms posting 50 or more ads in the Kuhn and Shen sample expressed at least one such preference. Notably, firms are less likely to express such a preference when educational requirements for the job are more stringent.

3.2.8 Organizational Demography

A developing literature in economics examines organizational demography; that is, factors that influence the demographic characteristics of those hired by the firm. Giuliano, Levine and Leonard (2009) use personnel data from a large US retail chain to show that the race or ethnicity of the hiring manager appears to be an important determinant of the racial composition of new hires. While store fixed effects — which capture both store characteristics and characteristics of the store’s local labor market — are the largest determinant of a store’s racial mix of hiring, Giuliano et al. (2009) find that the race of the hiring manager matters as well. Specifically, they report that non-black managers hire more whites and fewer blacks than black managers. Estimates suggest that the race of the store manager shifts the black employee share by around four percentage points. On average, black managers hire workforces that are 21 percent black, while non-black managers hire just 17 percent black. Similar effects are found for Hispanic vs. white managers, when restricting attention to stores where Hispanics make up 30 percent or more of the local population.

It is not clear what accounts for this propensity for racial match of hiring. It is possible that hiring managers are accessing their social networks (which may be partially segregated by race) to identify promising employees. Another hypothesis is that there are direct productivity effects, if, for example, black managers communicate better with black subordinates. This pattern could also be accounted for by preferences of managers or employees. It is difficult to disentangle these effects, but Giuliano et al. (2009) show that store-level sales does not appear to be significantly impacted by the racial match of the manager and employees.

Oyer and Schaefer (2010) examine the organizational demography of large US law firms. Using lawyer biographies posted to firms’ web sites, they document substantial across-firm
variation in hiring strategies. Law firms are found to be somewhat concentrated with regard to the law schools where they hire. Oyer and Schaefer find that for the average firm, the probability that two lawyers selected at random attended the same law school is about six percent higher than this probability for two attorneys selected at random from the sample. Some firms pursue a fairly unconcentrated hiring strategy, and the distribution of law schools attended by their attorneys is comparable to the overall sample distribution of law schools. Other firms appear to hire from a narrow set of law schools located in close geographic proximity to the firm’s offices. Still others hire nationally, but only from the very top law schools.

Oyer and Schaefer find that about a third of the observed variation in office-level law-school shares can be accounted for by simple geographic proximity. Law offices, by and large, disproportionately tend to hire from close law schools. Higher ranked law firms also tend to hire from high-ranked law schools, and this effect explains a small additional amount of variation in hiring practices. Even conditional on geography and reputation match, they report a strong association between associate-level law-school shares and partner-level shares. They estimate that when a firm’s partner law-school share is higher by one percentage point, the firm’s associate law-school share is higher by around 0.6 percentage points. As with the study by Giuliano et al. (2009), it is difficult to say whether social networks, production complementarities, or employee preferences are driving this relation. Wage and/or productivity measures are required to more specifically identify the causes and effects of these relationships.

3.2.9 Hiring, Agglomeration, and Firm Location

The optimal matching and searching processes that firms engage in will vary across local labor markets. Larger and more concentrated populations typically lead to thicker labor markets which may reduce search costs and can also lead to better average matches between firms and workers. While this will tend to increase surplus, thicker labor markets also lead to greater competition in the labor market. So, while total surplus may be greater in thicker labor markets, firms may have to settle for a smaller share of that surplus because it is more difficult to generate monopsony power.

There is a large literature on firm location and, more specifically, agglomeration economies, that studies the relationship between firm co-location and other variables. One of the factors that often lead similar employers to locate near each other is a source of certain types of
worker (such as skilled workers near university towns.) However, firm location decisions and the supply of labor in a given labor market will clearly affect one another, so it is difficult to generate credible causal statements about the effect of either one of these.

Several recent studies have analyzed the relationship between labor market thickness, firm location, and worker/firm matching. Wheeler (2001) develops a model where capital and worker skill are complements. In this model, thicker labor markets lead to higher productivity, greater wage inequality, and higher returns to skill. He cites and generates empirical evidence that is consistent with all these ideas. Andersson et al. (2007) put this same basic idea to somewhat more rigorous scrutiny. They use matched employer/employee data from California and Florida to show that there is more assortative matching between “high quality” workers and “high quality” firms in thicker labor markets. They show that establishment-level productivity is related to match quality and argue that the relationship between better matching and thick labor markets can explain a substantial portion of the urban productivity premium. Freedman (2009) looks at similar issues, with a focus on the software industry in a single (unnamed) state. He first derives a model where firm/worker match quality is based on differences in the Human Resources packages offered by firms and variation in employee preferences. Empirically, he finds that agglomeration of software firms is associated with higher wages, bigger firms, and less wage dispersion, as his model would suggest.

Garicano and Hubbard (2007) and Garicano and Hubbard (2009) analyze how market thickness affects the organization of law firms and the kind of work lawyers perform. In Garicano and Hubbard (2007), they show that bigger markets allow lawyers to specialize more, to be more likely to work in a hierarchy, and to have more leveraged hierarchies. That is, labor market thickness affects the matching process in that more senior attorneys can better leverage their unique skills. Garicano and Hubbard (2009) find that lawyers become more specialized as market size increases. The analysis drop the largest legal markets, so that they can be sure they are isolating the effect on the organization of work (rather than the type of work.) But they also show that some the most skill-intensive and expensive work gets done in a few big cities, indicating that firms and lawyers locate there to do certain types of work.

---

7 Local labor markets are just one of several important reasons firms agglomerate. See Ellison and Glaeser (1997) and Ellison and Glaeser (1999) who look at other sources of “natural advantage” and “spillovers” that drive co-location of similar firms.
This section is titled “Hiring” because no labor market matching can take place without an initial employment contract being formed. But labor market matching is constantly going on as firms decide who to retain and workers decide whether to engage in on-the-job search. While there is a large literature on the effects of job loss, there is relatively little on firms’ choices about retention and worker displacement. We have already touched on the post-hiring matching issues in our discussions of firing “risky” workers that turn out to have low productivity, raids, and up-or-out systems. We now add a short discussion of the literature on firms’ strategies regarding retaining and displacing workers.

Gibbons and Katz (1991) extended basic adverse selection models of hiring to the firing context. They derive a model where firms, when faced with a negative shock to productivity, lay off their less productive workers. Other firms draw inferences about workers’ ability when a prior employer selectively chooses who to lay off, while no such inference is possible for workers who lose their job because their firm shuts down altogether. Consistent with the model, they find that workers who are laid off from a continuing operation suffer more from job loss than workers who lose their jobs when their establishment shuts down. Though others have called the empirical results of Gibbons and Katz (1991) into question (see, for example, Krashinsky (2002)), the important conceptual distinction between “layoffs” and “plant closings” has been widely accepted in the literature.

When the match-specific component of productivity is important and firms pay workers something close to the marginal product of their effort, firms can (at least in principle) let employees efficiently separate from the firm when they need to reduce staff. That is, under the right conditions, there should not be a substantial adverse selection problem in offering voluntary severance packages and doing so may enhance the firm’s reputation and/or be necessary to honor written or implicit labor contracts. Pencavel (2001) studies this issue by looking at which University of California employees accepted buyouts in the early 1990s. As one might expect, larger severance benefits increase the probability of a worker accepting a buyout. But

---

8 Oyer and Schaefer (2000) further break down layoffs into those that are driven by economic issues and those driven by an individual’s poor performance (“firings”). They show that firms appear to choose which of these ways to displace workers at least partially based on the firing costs associated with each.
Pencavel also shows that it is otherwise very difficult to predict who will accept buyouts. Kim (2003) shows that the University of California got positive selection in one set of layoffs, because faculty whose productivity had been on the decline were more likely to accept buyouts.

There is one group that has been closely studied (over-studied?) in terms of job dismissal — CEOs of large American corporations. As with incentives, CEOs present an unusually public group in terms of observability of dismissal and performance. However, CEOs are also not very representative of employees more generally because of the large cost of not replacing a bad CEO, the fact that they may earn substantial rents, and the inability to move them to a new job within the firm if they are not working out as CEO.

Two recent papers have taken a careful look at the drivers of CEO dismissal. Both Jenter and Kanaan (2008) and Kaplan and Minton (2009) find that CEO turnover is sensitive to firm performance and that the effects of industry and total stock market performance are not completely filtered out of the relationship between performance and forced turnover. Kaplan and Minton (2009) also show that forced turnovers have become more responsive to turnover over time and question whether turnover decisions made by boards of directors are generally efficient. While these results are very useful for corporate governance scholars, it is hard to draw broad conclusions about displacement from these workers.

The availability of large employer-employee matched datasets has the potential to generate insights into displacement strategies of firms, though these data do not generally distinguish between layoffs, firings, and quits. Matching to the next job or to unemployment records may help narrow down the potential reasons for leaving, however. In the absence of being able to make such distinctions, even these rich datasets have little to add relative to the factors we already know to be associated with job loss (see, for example, Kletzer (1998).)

3.4 Do Hiring Practices Matter?

We conclude this section by asking a big question that has received far too little attention: Do hiring practices matter for the performance of an organization or a business unit?

There is limited evidence on this question, mostly coming from the growing literature on

---

9One minor complication is determining whether a departing CEO is leaving voluntarily, but researchers have developed credible methods for separating voluntary and involuntary departures. See, for example, Parrino (1997).
productivity effects of firm-level human resource management choices (and see Bloom and Van Reenen, 2010 in this Volume for a thorough review of this literature). As we noted above, the Ichniowski, Shaw and Prennushi (1997) study of steel finishing lines considers “extensive selection procedures” as one of a set of complementary HR practices that appear to have boosted productivity. Bresnahan, Brynjolfsson and Hitt (2002) show that one hiring practice — screening for education — appears to be part of a bundle of organizational practices that complement investments in information technology. Adoption of this set of complementary practices is associated with higher organizational value added. Bloom and Van Reenen (2007) conduct a survey of management practices and show that a variety of performance metrics — productivity, profitability, and sales growth, among others — are associated with good management practice. Their survey questions address a broad range of management practices, and among them is a series of questions about the firm’s attitude toward attracting human capital.

We think highly of these papers, but also believe this literature can be moved forward in a couple of ways. First, these papers are not, primarily, about hiring. Rather, they are about how a much broader set of human resource management choices are related to organizational performance. As a result, these papers do not generally make careful distinctions between, say, hiring practice A and hiring practice B. In some sense, the level of detail with regard to hiring choices is limited to simple questions about whether the firm thinks hiring is important, as opposed to gathering information about firm-level differences in specific hiring strategies. Second, this literature is subject to questions about causality. Firms in these studies are clearly choosing one set of organizational practices over another, and it remains unclear whether hiring choices are driving good performance, or whether there is some third factor that explains both.

We think this literature needs a series of carefully constructed hiring-related field experiments. Personnel Economics now has a very solid tradition of incentives-related field experiments, and we are eager to see this toolkit applied to hiring decisions. The lack of hiring-related experiments is, we think, evidence of the great importance of hiring in modern firms. What manager, after all, would allow an academic economist to experiment with the firm’s screening, interviewing or hiring decisions? We hope this concern will not prevent economists from performing Safelite-style experiments on hiring practices in the future.
4 Conclusion

We can summarize our view of the last few decades of Personnel Economics research as, “Incentives matter. Getting them right is important for firms. Measurement limitations and other challenges to employers’ ability to implement incentive programs are well explained by recent agency models.” The primary drivers of this success in studying incentives have been theorists’ ability to use advances in information economics to generate realistic and detailed models of employment relationships and the availability of new firm-level and matched employer-employee datasets.

We hope that the authors of the review of Personnel Economics in the next volume of this Handbook are able to conclude, “Recent research has generated important and practical insights into the ways firms and workers generate economic surplus by matching appropriately. Firms’ strategic decisions about how to source appropriate workers and how to craft attractive job packages have advanced significantly in the last few years. This is primarily driven by the development of more nuanced models of hiring and the creative use of firm-level and employer-employee datasets.” We recognize, however, that that is more easily hoped for than done. There are significant challenges to the matching research we call for. While data limitations are the most obvious challenge (it is very rare to have a data source with information both on people that a firm hires and those that the firm does not hire, for example), there may well be others such as the possibility that the heterogeneity of optimal hiring strategies may simply be much greater than the heterogeneity of optimal incentives. We hope that, on balance, these potentially higher costs of doing research on matching and hiring do not discourage researchers from undertaking work in this area and we hope that at least a few others will join us in trying to advance this research area.
References


