Modern agency theory begins with Mirrlees (1976) and Holmstrom (1979) with a general prescription for how compensation can be used to alleviate agency issues. It proposes that an agent’s pay should vary whenever there is information about her effort and that all information on performance should be used. Furthermore, the ability to resolve agency problems is limited only by risk-sharing considerations associated with random variation in performance measures. With some parametric restrictions, this also has the implication that greater randomness (uncertainty) reduces the intensity between pay and performance measures.

This work has been appropriately feted as the root from which modern agency theory derives. It does, however, suffer from one important problem: most people do not get paid this way. Instead, the current pay of most workers is insensitive to pretty much anything; relevant information is consciously not used, and the relationship between uncertainty and the
intensity of pay for performance if anything goes the wrong way. Because of this, the literature published in the journal has instead originated a series of avenues that—while respecting the logic of these contributions—at times better accord with observed practices. These avenues largely fit into two categories: (i) there are other ways to skin the cat of motivating workers, and (ii) in many settings, some of the assumptions of the canonical model do not hold. I consider each in turn.

Other Ways of Motivating

Workers are often embedded in firms, in markets, and in long-lasting relationships. Each of these issues has been explored to offer alternatives to the pay for performance logic of the canonical model.

First, most of us work in organizations that are hierarchical, with workers sorted into positions on the basis of performance and ability. Given this, it should not be surprising that an alternative source of motivation for workers is the possibility of promotion. The earliest and most important contribution here is Lazear and Rosen (1980), which argues for the role of tournaments in providing incentives. Tournaments are settings in which a group of agents compete for a set of fixed prizes rather than face a pay schedule that varies explicitly with their actions. Agents exert effort to change the probability of getting a better prize. In contrast to an individual piece rate setting, here what matters is relative performance in a particularly stark way, as only rank order performance determines pay. In simple settings, Lazear and Rosen show how tournaments can induce the first-best outcome without explicit pay for performance and offer some results on when tournaments dominate piece rates based on only individual performance. This logic has subsequently been extended by Green and Stokey (1983) and Malcomson (1984).

One difficulty that agency theory has faced has been the absence of clean empirical testing. Tournament theory is a welcome exception. There are two natural tests. First, do bigger prizes cause agents to work harder? In a number of (mostly sports) settings, this has been shown to be so. The second test—do perceived marginal probabilities of winning affect effort?—is both more subtle and more directly related to the notion that agents compete in probabilities of winning. The strongest work in that vein is Brown (2011), which showed that the presence of Tiger Woods in golfing tournaments acted as a disincentive for other golfers. They perform worse as they perceived the marginal value of effort to be lower.

Second, workers are embedded not only in firms but also in markets. Another theme that agency theory has explored successfully over the last three decades is how labor markets can act as an alternative to contracts. This logic derives from Fama (1980), which described how external audiences can constrain opportunities for moral hazard through the process...
by which outside opportunities evolve. A simple example would be a professional baseball player. Despite the plethora of performance measures on his performance, pay for performance is rare. The reason, of course, is that there is an external market that sets the market pay of baseball players and a player who shirks can expect his reputation and future opportunities to worsen.

Fama does not claim that markets always offer enough discipline to solve agency problems. This issue has been elegantly formalized by Holmstrom (1999), and these two papers have now helped to develop a subfield of agency theory known as “career concerns.” A more general characteristic of the orientation of the JPE is a focus on theoretically informed empirical testing. A particularly notable example of this in this field is the paper by Gibbons and Murphy (1992). They study the interaction between career concerns and formal pay for performance in the market for CEOs. Using a tightly specified model of career concerns interacting with formal pay for performance, they show how greater career concerns incentives reduce the need for formal pay for performance, as arises in their empirical results.

Finally, workers are located not only in hierarchies and markets but often in long-term employment relationships. This allows temporal considerations to enter optimal contracting, where long-term employment relations allow the possible use of deferred compensation. This is where an agent may not be rewarded for current performance today, but rather sometime in the future. There are by now many dynamic contracting papers whose optimal outcome is to hold back some component of pay until late in the agent’s career. The logic is usually simple: by deferring a performance-related bonus until a worker is older, incentives for older workers improve while maintaining incentives for the younger worker (because by working hard now, she can be in line for that bonus later). One of the earliest formalizations of this logic is the study by Lazear (1979). His interest in that work is the need for mandatory retirement in optimal employment relationships. However, the reason why mandatory retirement is needed is that older workers are overpaid relative to their contemporaneous marginal productivity, which itself derives from the desire to defer compensation for the reason above.

The Assumptions Do Not Hold

Another series of extensions to the canonical model have arisen from the realization that many of its assumptions do not hold in many important settings. Here I provide a number of important examples.

The canonical model assumes that more pay for performance changes behavior only in ways that benefit the principal. However, by now it is well known that incentive pay can also induce the kind of dysfunctional
behavioral responses that cause pay for performance to backfire. Chevalier and Ellison (1997) offer a nice example. This line of research has become known as multitasking and typically mutes the use of performance pay. An elegant and tractable example of this is Baker (1992).

Other early contributions in the *JPE* on dysfunctional responses have focused on another characteristic of organizations, namely, the use of rules to allocate resources over allowing discretion. A series of papers have related this to dysfunctional behavioral responses. Consider the motivating example in Milgrom’s (1988) important contribution on what he terms “influence activities.” American Airlines needs to staff routes, and flight attendants have preferences over which routes they are given. Rather than allowing supervisors to assign “shifts” on the basis of the idiosyncratic preferences of employees, it uses a much simpler rule: routes are assigned by seniority, where workers with the most seniority pick first. (As a more substantive example, many firms use “last in, first out” rules for layoffs, despite the fact that some more senior workers could be less productive for the firm than their junior counterparts.) This arises in Milgrom’s work as a way of deterring dysfunctional lobbying behavior by workers, where time is spent influencing superiors for resources for themselves rather than spending time on more productive activities. Said another way, while such bureaucratic rules may be inefficient at the point at which the decision is made, it may save sufficient resources at an earlier point to be worthwhile. Similar logic underlies Prendergast and Topel (1996) in a setting in which favoritism arises.

Public agencies are often accused of being unaccountable to their constituents. Foremost among these is the behavior of police forces. Another example of how ex post inefficient rules can be part of optimal oversight arises in Prendergast (2003). In that setting, legitimate consumer complaints are ignored. This is done because if public agents believe that their behavior is likely to be investigated on the basis of such complaints, they will simply capitulate to those consumers in settings in which they should not. (In the police example, they become more resistant to arresting suspects.) Once again, the response to dysfunctional behavioral actions is to ignore valuable information.

Probably the most important parallel exploration to agency contracting has been to understand governance in settings in which outcomes are noncontractible. There have been two lines of inquiry. The first has derived from the seminal work of Grossman and Hart (1986), where ownership of assets (or, in later work, the control of assets) can be used to mitigate agency concerns. This has led to the modern theory of the firm. This contribution and its extensions are described elsewhere in this issue by Rob Vishny and Luigi Zingales. The second area dealing with noncontractibility has addressed relational contracts, where repeated interactions can potentially resolve noncontractibility issues.
Work on noncontractible environments has also led to consideration of other instruments than compensation to align incentives. Contributions to the *JPE* have played a central role in this. An important example is Aghion and Tirole’s (1997) work on real and formal authority. Consider a setting in which a worker cares not just about current pay and effort but also about the kinds of activities that they engage in. Furthermore, by exerting effort, she can identify which activity she prefers. Yet she will exert effort only if that preferred outcome is likely to be implemented. The problem for her is that her boss may not agree with her and instead overturns her recommendation. The principal will overturn, though, only if he is sufficiently sure of the right action. In settings in which he is not sure, the agent has real authority even though the agent may be subject to the formal authority of her boss. However, if the fear of being overturned is sufficiently salient to the agent, she will not exert any effort. In these settings, the firm may delegate formal authority to the agent. This work has been influential not just for its elegant and tractable modeling but also as the picture it paints—of organizations characterized by conflicts, with individuals vying for control—resonates with reality.

A feature of many institutions is conflict. Recent work in noncontractible settings has focused on the value of using workers who do not share the beliefs and preferences of their superiors. An early example is Che and Kartik (2009). The authors consider a setting in which an expert collects information to determine the right course of action, but she may be biased for or against that action (compared to the beliefs of her principal). Suppose that the principal could choose the bias of the agent: should she share the preferences of the principal? The intuitive suggestion that their beliefs should be aligned turns out not to be right. The reason is that an agent who shares the principal’s belief realizes that if she exerts no effort, the principal is likely to do what the agent already thought was the right answer. Instead, the optimal strategy is to introduce disagreement between agent and principal, because an agent who believes that the principal’s prior is wrong is more likely to work hard to dissuade him. Yet this is not costless, as she is less likely to reveal her information clearly.

Many workers, arguably most, are not rewarded on output measures. Instead their inputs are monitored, where they follow instructions provided by their superiors. Showing up on time and doing what is asked of them is the reality for most workers. Imagining input monitoring as an alternative to pay for performance has also helped to make progress on understanding one of the empirical difficulties faced by the literature. Specifically, there is little empirical support for what has become known as the trade-off between risk and incentives, where more uncertain environments would result in less pay for performance. Instead, the evidence seems more supportive of greater uncertainty leading to more
pay for performance. Prendergast (2002) addresses this by noting that for many workers, the alternative to pay for performance is a situation in which a superior tells a worker what to do. Now consider a setting with more uncertainty. In the canonical model, this only adds measurement noise to the ability to infer agent effort. This alone would attenuate pay for performance. However, in more uncertain settings, a superior may now additionally be less able to tell the agent what to do (as he knows less about what is going on). If so, the principal may need the agent to decide the right course of action in these uncertain settings, which likely leads to more pay for performance. Said another way, uncertainty may indeed render pay for performance costly; it may render the alternative even worse.

When a CEO increases the earnings of a company or a sales agent sells more, we can be pretty sure that this is a good thing. A final area of exploration in the field has been to consider settings in which it is not clear what output means. For example, when an auto mechanic tells you to have your car repaired at some expense, it is unclear if this reflects good or bad performance by the mechanic. Taylor (1995) studies this problem. A beautiful example of this inability to interpret performance measures is the work of Dewatripont and Tirole (1999) on advocates. Once again, consider a setting in which information needs to be collected on whether to carry out an action. The canonical agency model relies on a monotone likelihood ratio property (MLRP), where if it holds, pay will strictly increase in output (or profits). The difficulty here is that information can be positive or negative and can be offsetting. Specifically, an agent who has worked hard and found one piece of positive information and one piece of negative information finds herself in the same position as one who collects no information. Formally, this means that the MLRP of the canonical model fails, and it can render it impossible to have one person collect all the information. The alternative is advocacy: where one person collects only information that is positive and the other collects only the negative information, and they are rewarded if the outcome reflects the kind of information they collect. As such advocacy is so pervasive inside and outside organizations, this paper offers an insight far from the canonical model, but does so with a minimum of additional assumptions.

To conclude, it should be clear from this short essay that the literature on agency has come far from its original focus on the shape of compensation functions, and the central role played by the JPE in that development. It is perhaps worthwhile to conclude by noting the paucity of empirical work among the contributions above. Fields can thrive only when numbers are added to the Greek alphabet, and it is hoped that the JPE can play a significant role in promoting such work going forward.
References


