Abstract

This article surveys the literature on principal-agent problems with moral hazard that gained popularity following the seminal works of Mirrlees (1976), Holmström (1979), and others. This literature is concerned with designing incentives to motivate one or more workers—typically by paying for performance—in settings where his actions cannot be directly contracted on. I begin with the canonical framework, and then in Section 2 I categorize the subsequent literature that relaxes various assumptions of this framework. Section 3 focuses on the empirical research, and Section 4 discusses other strands of the literature. My goal is to provide a broad and concise overview—sometimes at the expense of depth, and focus on the economic insights (rather than methodological contributions).

1 Foundations

This article provides a broad but concise overview of the agency literature under moral hazard, which is concerned with designing incentives—typically performance pay—to motivate one or more workers.

In the canonical framework of Holmström (1979) the game begins with the principal offering a contract $w$ that specifies the agent’s pay as a function of output produced. The agent then decides whether to accept the contract, in which case he chooses a costly action (effort), or he rejects the contract and collects his outside option, which corresponds to seeking employment elsewhere. Finally, output is generated according to a distribution function that depends on the agent’s action, payoffs are realized, and the game ends. The principal’s objective is to maximize her expected profit, while the agent chooses an action that maximizes his expected payoff. Mathematically, this problem can be formulated as follows:

$$\max_{w(\cdot),a \geq 0} \mathbb{E}[x - w(x)|a] \quad \text{(P)}$$

subject to

$$a \in \arg \max_{a' \geq 0} \mathbb{E}[u(w(x))|a'] - c(a') \quad \text{(IC)}$$

$$\mathbb{E}[u(w(x))|a] - c(a) \geq u, \quad \text{(IR)}$$

1. I am grateful to Enrique Ide, Niko Matouschek, Michael Powell, as well as to Peter Traber for excellent research assistance.

2. Kellogg School of Management, Northwestern University. Email: g-georgiadis@kellogg.northwestern.edu.
where \( x \) denotes output, \( a \) the agent’s action, \( u \) his strictly increasing and concave utility function, \( c \) his strictly increasing and convex cost of choosing each action, and \( y \) the value of his outside option. The expectation is taken with respect to a family of probability density functions \( f \) that map each action into output; i.e., \( x \sim f(\cdot|a) \). All parameters are common knowledge. The principal here chooses a contract and a recommended action to maximize her profits subject to two constraints: (IC) stipulates that the recommended action is incentive compatible given \( w \), and (IR) requires that the agent prefers the offered contract to his outside option.

As a benchmark, consider the first-best outcome; i.e., the one that maximizes total welfare. This corresponds to the solution of (P) when (IC) is omitted. Because the agent dislikes risk while the principal is risk-neutral, the first-best contract features constant pay.\(^1\)

To solve this program, it is standard to replace (IC) with its first-order condition, and decompose the problem into two steps (Grossman and Hart, 1983).\(^2\) In the first, one finds for each action the contract that minimizes the principal’s expected cost. This has been shown to satisfy

\[
\frac{1}{u'(w(x))} = \lambda + \mu \frac{f_a(x|a)}{f(x|a)} \quad \text{for all } x,
\]

where \( f_a \) is the derivative of \( f \) with respect to \( a \), and \( \lambda \) and \( \mu > 0 \) are the Lagrange multipliers associated with (IR) and the first-order condition of (IC), respectively. In the second step, one solves for the profit-maximizing action.\(^3\)

Intuitively, the cost-minimizing contract attaches large payments to informative-but-rare outputs; i.e., outputs for which \( f_a(x|a) \) is large, signifying that it is highly likely that the agent indeed chose the target action \( a \) rather than \( a - \epsilon \) (which is the relevant deviation given the first-order approach and \( \mu > 0 \)), and outputs for which \( f(x|a) \) is small so the principal must make these large payments only rarely. Moreover, the variation in payments is decreasing in the curvature of \( u' \); that is, the optimal contract is shaped by a trade-off between providing incentives to the agent and insuring him from volatility in his pay which, being risk-averse, he dislikes.\(^4\)

A key insight from this model is the informativeness principle, which asserts that any signal that is informative of the agent’s action should be incorporated into the contract. This has several implications. If, for example, a CEO’s current efforts impact future profits, then incentive pay should also depend on lagged profits. In similar vein, if multiple executives are subject to common market-wide shocks, then it is valuable to evaluate them relative to one another. See also Holmström (2017) for a more detailed analysis of this framework and discussion.

---

\(^1\)This outcome can also be achieved in the presence of (IC) if (i) output is deterministic, or equivalently, effort is contractible, (ii) the distribution that maps \( a \) to \( x \) has shifting support, or (iii) the agent is risk-neutral. See Holmström (2017) for details.

\(^2\)Replacing (IC) with its first-order condition, \( E[u(w(x))f_a(x|a)/f(x|a)|a] = u'(a) \), is without loss of generality if and only if the agent’s expected payoff is differentiable and single-peaked in \( a \). Rogerson (1985a) and Jewitt (1988) provide sufficient conditions for this to be the case.

\(^3\)The second-step program, barring stringent conditions (Jewitt, Kadan and Swinkels 2008), is not convex but can be solved via line-search. Luckily most of the insights can be gleaned from the solution of the first step.

\(^4\)Note that ex post, the principal has no uncertainty about the agent’s action. Yet to provide him with ex ante incentives to choose a particular action, the contract must reward him generously for some output realizations and punish him for others.
2 Towards More Realistic Models

We now turn to the literature that has sought to relax various assumptions of the above framework.

2.1 Flexible Actions

In the canonical model, the agent chooses a one-dimensional action which determines the entire distribution of output. In reality agents have at least some flexibility in determining distributions over outcomes in settings including professional investing (Chevalier and Ellison 1997), management (Shue and Townsend 2017 and Rahmandad, Henderson and Repenning 2018), and sales (Oyer 1998 and Larkin 2014).

Holmström and Milgrom (1987) were one of the first to model flexible actions. Take a salesperson who is paid according to quarterly sales. Each morning, he checks his sales in the current quarter and chooses how much effort to spend. Under certain conditions, the optimal contract is linear in quarterly sales and induces constant effort over time. Intuitively, if the contract were nonlinear, the agent would be able to exploit it at the principal’s expense. Suppose for example that it pays a fixed wage plus a lump-sum bonus if sales exceed a quota. If the agent hits the quota prior to the end of the quarter, he has no incentive to keep working; if he deems it unlikely to meet the quota, he likewise gives up. A linear contract is robust to this type of behavior because each sale is rewarded the same regardless of when it was realized. This result is often used to motivate restricting attention to linear contracts in static models, which allows one to solve the entire problem analytically and obtain testable comparative statics; see for example Prendergast (1999).

Others have studied static models where the agent can influence the entire distribution of output. In Palomino and Prat (2003) where the agent controls the first two moments of his output distribution, a bonus contract is optimal; i.e., one with binary payout. In Diamond (1998) the agent can choose any distribution at a cost that depends on its mean, whereas in Barron, Georgiadis and Swinkels (2020) the agent can costlessly add mean-preserving noise to his output. In both cases the optimal contract is linear. Hebert (2017) microfounds a model where the agent can choose any output distribution at a cost that is proportional to its Kullback-Leibler divergence relative to some default distribution and shows that a piece-wise linear contract akin to a call option is optimal, while Georgiadis, Ravid and Szentes (2022) studies a model without functional form assumptions on the agent’s cost of choosing a distribution.

2.2 Multiple Tasks

Most jobs comprise many tasks; e.g., salespeople selling multiple products, employees controlling both quantity and quality of output, managers splitting their time between their own work and training.

Formally, this can be modeled by letting the agent choose the output distribution $F$ to maximize $\int u(w(x))dF(x) - C(F)$, where $C(F)$ is the agent’s private cost of choosing $F$. Researchers have explored different restrictions on $C$ to suit particular applications while maintaining tractability.

Another way to enrich the agent’s action space is to let the resolution of uncertainty be observed by the agent before he chooses his action; e.g., he might learn how difficult the task is (Edmans and Gabaix 2011). As in other models, this leads to simple contracts being optimal.
others, and so on. Moreover, sometimes employees can improve their performance measures with legitimate effort as well as with undesirable or illegal tactics.

Holmström and Milgrom (1991) proposes the following framework to model such situations. The agent now chooses a vector of \( n \) actions, denoted by \( a \), by bearing cost \( c(a) \). The principal has a vector of \( m \) performance measures denoted \( x = Ma + \epsilon \), where \( M \) is an \( m \times n \) matrix and \( \epsilon \) is \( m \)-dimensional Gaussian, zero-mean noise. Assume the principal’s profit is \( \omega \cdot a \) net of the agent’s pay, where \( \omega = (\omega_1, \ldots, \omega_n) \) represents how much the principal values each action, and following Holmström and Milgrom (1987), she restricts attention to linear contracts of the form \( w(x) = b + \alpha \cdot x \), where \( b \) is a wage and \( \alpha \) is the commission rate for each dimension of performance.

This model can be solved analytically with some additional assumptions, and generates several insights. First, if some actions are measured with more noise than others, then it is optimal to provide lower-powered incentives on all measures; i.e., to choose smaller \( \alpha_i \)’s. As an example, let the actions represent the quantity and quality of output, where the latter can only be measured with noise. Given the risk-incentives trade-off, other things equal, the principal would like to offer higher-powered incentives for quantity than for quality. However, because the tasks are substitutes, doing so will lead the agent to substitute effort away from quality and into quantity at the detriment of the principal. In such cases, it is optimal to mute incentives for both quality and quantity.

The second insight applies to settings where none of the performance measures are perfectly aligned with what the principal truly cares about; i.e., none of the rows of \( M \) are collinear with \( \omega \) (Baker, 1992). In such cases, the optimal power of incentives (i.e., \( \alpha \)) increases in the alignment of the performance measures with the principal’s preferences. Take for example Volkswagen, whose leadership set out a plan in 2008 to drastically increase sales in the United States, and promised executives and engineers lucrative incentives for meeting certain targets. Those targets could be achieved with product improvements, or alternatively, by cheating. As it turned out however, product improvements were much harder than the (perceived) disutility associated with cheating, and the high-powered incentives led employees to mask the emissions of their diesel engines. Other examples include doctors cherry-picking patients (Dranove et al., 2003), the Wells Fargo fake-accounts scandal, the BP Texas City refinery explosion, and the V.A scandal involving manipulated medical waiting lists. See also Holmström (2017) for further details and Bonham and Riggs-Cragun (2021) for a multitasking model with flexible actions in the spirit of Section 2.1.

2.3 Dynamics

An extension which has received a lot of attention is one where the players interact over multiple periods. Dynamic models come broadly in two flavors: ones in which the principal can commit to a contract at the beginning of the relationship that specifies payments to the agent in each period as a function of current and past output, and models in which she cannot—typically referred to as

7The interesting case arises when \( c \) is a supermodular function. For example, if \( a_1 \) and \( a_2 \) represent quantity and quality, respectively, this assumption implies that the tasks are substitutes; i.e., the marginal cost of providing quality increases in quantity, and vice versa.

8Such a situation can be modeled by letting \( a_1 \) and \( a_2 \) represent effort towards product improvements and cheating, respectively, \( x = a_1 + a_2 + \epsilon \) denote sales, and \( \omega_1 > 0 > \omega_2 \) represent the firm’s weights on the two dimensions of effort.
relational contracting models.

**With commitment.** Early contributions are due to Rogerson (1985b) and Spear and Srivastava (1987), who study a two-period and infinite-horizon model, respectively. An important insight is that the optimal contract has memory and moreover, it backloads incentives: good performance in the current period is rewarded with a larger immediate payment, and larger payments in future periods. DeMarzo and Sannikov (2006) and Sannikov (2008) reconsider this problem with a risk-neutral and a risk-averse agent, respectively, whose effort controls the drift of a Brownian motion. Their continuous-time formulation permits a sharp characterization of the optimal contract and several insights—see Athey and Skrzypacz (2017) for an overview. Others including Myerson (2015) and Green and Taylor (2016), analyze this problem with Poisson production technologies, which leads to optimal contracts with somewhat different properties.

A strand of this literature studies models in which the agent can manipulate output intertemporally. Edmans et al. (2012) considers a tractable model in which the agent can engage in short-termism by inefficiently increasing current stock price returns at the expense of future returns. To deter this behavior, the optimal contract has the property that the return in any given period affects log pay in future periods at an increasing rate. In DeMarzo et al. (2014) and Wong (2019), the agent can increase current output at the risk of a low-probability but large-in-magnitude loss the next instant. A takeaway is that deterring risk-taking following a period of poor performance may be unavoidable. See also Makarov and Plantin (2015) for a model with risk-taking where the principal learns about the agent’s ability over time.

**Without commitment.** Performance evaluations are often based on “soft” information such as a subjective evaluation by a supervisor. In these cases, performance-contingent contracts cannot be enforced by a court, so firms must provide incentives on a relational basis. To capture such a situation, MacLeod and Malcomson (1989) considers a discrete-time model in which the principal can voluntarily pay the agent every period after observing his effort. As long as players are sufficiently patient, there exists a stationary equilibrium where in each period, the principal “recommends” a particular effort, the agent follows the recommendation, and receives a “bonus” in return. If the principal ever reneges on paying the bonus—off the equilibrium path—the agent will choose the cheapest effort in every subsequent period; i.e., he plays a grim-trigger strategy. Intuitively, it is in the principal’s interest to pay the bonus as long as its cost is smaller than her value of maintaining the relationship. This framework has been extended to incorporate multiple agents (Levin 2002), unobservable efforts and noisy output (Levin 2003), endogenous feedback (Fuchs 2007), multiple agents and relational incentives coupled with formal contracts (Rayo 2007), asymmetric information (Halac 2012 and Li and Matouschek 2013), multiple agents with bilateral monitoring (Andrews and Barron 2016 and Barron and Powell 2019), and limited liability (Fong and Li 2017).

If the game comprises a single period, then the only way to motivate the agent is with a scheme that involves money-burning (MacLeod 2003).
2.4 Unknown Production Environment

Another strand relaxes the assumption that the production environment is perfectly known by all parties. These models fall in three categories: those in which the principal designs a mechanism to elicit the agent’s private information, those where the principal chooses a contract to maximize her worst-case payoff, and career concern models where both players learn about the production environment over time.

**Moral Hazard & Adverse Selection.** The first features models where the principal is oblivious of one parameter; e.g., the agent’s cost function is $c(a, \theta)$ where $\theta$ is the agent’s private information. Guesnerie and Laffont (1984) studies this problem assuming output is deterministic so that conditional on $\theta$, the agent’s action is observable. Chade and Swinkels (2022) incorporates noisy output and gives conditions such that the principal’s problem can be disentangled into a pure adverse selection and a pure moral hazard problem. The optimal mechanism involves a menu of output-contingent contracts with each type self-selecting into his designated one; see also Foarta and Sugaya (2021). When the agent is cash constrained, all types are offered the same contract under general conditions if he is risk neutral (Gottlieb and Moreira 2022), and under more stringent conditions if he is risk averse (Castro-Pires and Moreira 2021).

**Robust Contracts.** When multiple parameters are unknown, the mechanism design approach employed above becomes intractable, and even forming a prior might be impractical. Researchers have therefore studied “robust contracts” that provide the principal with the largest payoff guarantee—this is as if an adversarial third party can choose the unknown aspects of the environment after observing the offered contract to minimize the principal’s payoff; so the principal seeks a contract to maximize this worst-case payoff. If the principal knows only a subset of the actions available to the agent and their costs, then a linear contract is optimal (Carroll 2015); under monotone-likelihood-ratio restrictions on the adversarial party’s action space, the optimal contract is a mixture of equity and debt (Antic 2022). Walton and Carroll (2019) provides general conditions for linear contracts to be optimal, Chassang (2013) and Miao and Rivera (2016) embed these ideas in dynamic settings, and Dai and Toikka (2022) incorporates team production. See also Carroll (2019) for an overview.

**Career Concerns.** A key source of one’s motivation is often tied to their reputation: what others think of their abilities, where a favorable perception can lead to promotions, outside offers, and pay raises. Holmström (1999) considers a model with two-sided uncertainty where both the principal and the agent are learning about the agent’s ability over time from past performance, and the agent’s wage increases in the principal’s perception of his ability. Absent incentive contracts, career concerns induce an inefficient pattern of effort over time: the agent works inefficiently hard early on while the principal is learning about their ability, and not hard enough later on when the principal’s belief is hard to influence. See also Gibbons and Murphy (1992) for a contracting model with career concerns,

10 Also related is Garrett et al. (2021) where the agent designs his production technology—by choosing the cost of each output distribution—before contracting with the principal.
2.5 Endogenous Performance Measures

To design an incentive plan, one must first identify appropriate performance measures, and then specify the relationship between pay and those measures. Dye (1986) considers a model a-la Holmström (1979) where after observing a signal of the agent’s effort, the principal can acquire another costly signal. In Georgiadis and Szentes (2020), the principal chooses a strategy for sequentially acquiring signals and a wage scheme as a function of the acquired signals, and shows that a bonus contract is optimal. Hoffmann, Inderst and Opp (2020) studies deferred compensation: signals about the agent’s one-shot effort arrive over time and the principal decides when to stop and pay the agent according to the observed signals. The central trade-off in these models is that acquiring more signals is costly, but provides the principal with more accurate information, reducing agency costs.

Khalil and Lawarree (1995) considers a pure adverse selection model where the principal can monitor either the agent’s input (e.g., hours worked) or his output. When these monitoring technologies are equally informative, the former is preferable if and only if the principal is the residual claimant of output; see also Baker (2002). Datar, Kulp and Lambert (2001) shows that if one restricts attention to linear contracts, then it may be optimal to ignore some informative signals. In Li and Yang (2020) the principal chooses a partition of the signal-space at a cost that increases in the fineness of the partition and a wage scheme as a function of the cell of the partition in which the signal lies.

2.6 Many Agents

Another significant strand studies models with multiple agents and externalities. These can be positive, if agents work in teams for example, or negative, if they are evaluated and paid relative to one another, as in a tournament or a stack ranking system.

Teams. When people work in teams, disentangling each person’s contributions is often difficult (Alchian and Demsetz, 1972). This may then necessitate incentivizing the group as a whole. In a seminal contribution, Holmström (1982) shows that when only joint output is contractible, there exists no sharing rule that implements the efficient output. The intuition is that because agents don’t internalize the marginal benefit of their efforts on others, they work too little from the perspective of joint surplus; i.e., they freeride. When efforts are complementary, Winter (2004) shows that the cheapest wage scheme that induces identical agents to work in the unique equilibrium necessarily treats each agent differently; Halac, Lipnowski and Rappoport (2021) further shows that this design hinges on wage schemes being public. Admati and Perry (1991) and Fershtman and Nitzan (1991) extend these ideas to dynamic settings and show that freeriding is reduced if but only if players’ efforts are strategic complements across time.11

11See also Georgiadis (2014) for a continuous-time formulation which provides testable comparative statics, and Cvitanic and Georgiadis (2016) for a mechanism to eliminate freeriding.
Relative Performance Evaluations & Contests. The productivities of workers are often subject to common shocks, that is, the external factors that affect one worker’s performance often affect others’ performance in similar ways. For example, the productivities of executives across firms or of salespeople within a firm may be jointly affected by the broader economic conditions. In such cases, relative performance evaluation and pay can serve to filter out these shocks hence reducing the riskiness of pay (Holmström, 1982). This observation initiated the literature on tournaments starting with Lazear and Rosen (1981), who characterizes optimal incentive schemes that condition pay on each agent’s ordinal rank rather than absolute performance. Green and Stokey (1983) shows that rank-order tournaments dominate individual contracting if the common shock is sufficiently diffuse. In general however, by Holmström’s Informativeness Principle, conditioning pay on both absolute and relative performance is optimal (Mookherjee, 1984). Rosen (1986) studies sequential pairwise elimination tournaments commonly encountered in athletic events, while Gradstein and Konrad (1999) shows that depending on the relative importance of effort to random factors, the optimal structure either comprises a single stage, or it is a pairwise elimination tournament. Another benefit of tournaments is that they may facilitate commitment when performance evaluations are subjective: it is easier for the principal to commit to award a fixed prize to the best-performing agent no matter what than to commit to an evaluation-contingent contract (Malcomson, 1984).

3 Empirical Literature

The empirical literature on contract theory can be broadly classified into three buckets. The first comprises studies that evaluate the degree to which workers respond to incentives. The second features papers that examine whether observed contracts have the properties predicted by theory. The third bucket includes papers that operationalize different theories, for example by building structural models, recovering the unknown parameters, and computing an optimal contract targeted at real applications.

Responsiveness to Incentives. Several papers use quasi-experimental or experimental variation to show that higher-powered incentives cause individuals to work harder, at least on the dimensions that are highly rewarded. These effects have been found in a variety of settings, including windshield repairers (Lazear, 2000), tree planters (Shearer, 2004), bicycle messengers (Fehr and Goette, 2007), gig workers (Chen and Sheldon, 2016), salespeople (Chung and Narayandas, 2017), workers on Amazon’s Mechanical Turk platform (DellaVigna and Pope, 2018, 2022), creative tasks (Englmaier et al., 2018), and day laborers (Güteras and Jack, 2018) in developing countries. An important finding is that incentives affect not only motivation, but also selection. In particular, when a windshield installation firm switched from hourly wages to piece rates, productivity increased by 44%, half of which was attributable to increased motivation and the other half to selection; that is, more productive workers...

---

12Another way to filter out common shocks is by evaluating performance relative to a benchmark; e.g., portfolio managers’ returns are often evaluated against an index such as the S&P 500.

13See also Harris and Vickers (1987) for a dynamic tug-of-war, Taylor (1995) for an optimal dynamic contest, and more recently, Halac, Kartik and Liu (2017) and Ely et al. (2022) for dynamic contests where the principal can curate the agents’ beliefs about their own and their rivals’ progress.
joining the firm and less productive ones leaving (Lazear 2000). See also Paarsch and Shearer (2000) and Lo, Ghosh and Lafontaine (2011), as well as Lazear (1986) for a theoretical argument.

Another strand studies multitask settings and how imbalanced incentives can lead to adverse outcomes. Chevalier and Ellison (1997) and Shue and Townsend (2017) presents evidence of mutual fund managers and CEOs taking excessive risks, respectively. Oyer (1998) and Larkin (2014) find that nonlinear incentives led salespeople to manipulate the timing of their sales, and being paid according to sales revenue, they offered excessive discounts. See also Jensen (2001) for additional examples. Tzioumis and Gec (2013) presents evidence that mortgage officers approved marginal applications with higher delinquency rates when close to making their quota at the end of the month. Dranove et al. (2003) finds that when the states of New York and Pennsylvania began publishing health care report cards with surgeon-specific mortality rates, doctors stopped doing certain procedures for very ill patients. In a manufacturing setting, Hong et al. (2018) finds that higher piece rates had a positive effect on quantity at the expense of a large increase in defect rates.

A manifestation of the multitask problem concerns settings where a worker’s pay depends both on his and others’ productivity; e.g., under relative performance or team pay. Bandiera, Barankay and Rasul (2005) presents evidence that relative performance pay for agricultural workers led to lower productivity vis-a-vis individual piece rates if (but only if) workers could easily monitor each other. Bandiera, Barankay and Rasul (2007) on the other hand, finds that paying managers a performance bonus if aggregate productivity was sufficiently high led to them working harder to increase the productivity of their subordinates; see Itoh (1991) for a theoretical treatment. Berger, Harbring and Sliwka (2013) finds positive effects of a forced distribution system in a real effort experiment.

Are Observed Contracts Optimal? The risk-incentives trade-off asserts that, other things equal, pay-for-performance sensitivity should be lower if the performance measure is more volatile or the worker is more risk averse. However, evidence of this trade-off coming from various occupations is, at best, weak (Prendergast 2002). Prendergast posits three reasons why pay-for-performance sensitivity may be positively correlated with risk. The first has to do with performance pay being more prominent in more uncertain environments where the choice of what tasks to work on is delegated to the employee—as opposed to a manager telling them what to do and paying fixed wages. The second is that performance reviews are often subjective and are more likely to be biased when stakes are higher. If the environment is noisier, this bias is less costly for the firm, and so higher-powered incentives are optimal. The third reason is that performance investigations are often only triggered if poor decision-making is suspected; e.g., following an accident. If the initial impression that triggers an investigating is noisier, then the worker suffers less from the investigation and so their ex ante incentives are muted. In response, the firm optimally provides stronger incentives.

A second prediction has to do with the shape of the pay-for-performance relationship. In practice, this is typically simple—for example efficiency wages, bonus contracts, linear and piecewise linear

---

14 Another example comes from the FBI: In the years after 9/11, leadership discovered that the field offices that did the best intelligence work also wrote the most reports on intelligence matters. So in 2007 the FBI started evaluating field offices based on pages of intelligence produced. By 2008 the positive relationship between reports produced and quality had disappeared—agents were simply filling their word-count quotas (Fisman and Sullivan 2015).
contracts are most common (Prendergast 1999). In contrast, the shape of the optimal contract in canonical models depends on the primitives and is simple only under specific assumptions (Grossman and Hart 1983). However, linear contracts have been shown to be close to optimal in the canonical Holmström model under relatively mild assumptions (Bose, Pal and Sappington 2011). Moreover, researchers have proposed richer models that can explain the optimality of simple contracts. In particular, the optimal contracts in Holmström and Milgrom (1987), Carroll (2015) and Barron, Georgiadis and Swinkels (2020) are linear, in Antic (2022) and Hebert (2017) they are piecewise-linear, whereas in MacLeod (2003) and Georgiadis and Szentes (2020) they comprise a fixed wage and a single bonus.

Operationalizing the Theory. Misra and Nair (2011) estimates a dynamic structural contracting model and uses it to evaluate counterfactual contracts for the salesforce of a large contact lens manufacturer; see also Chung, Steenburgh and Sudhir (2014). Chung, Kim and Park (2021) estimates a structural model that includes performance pay, as well as training, recruiting and termination decisions, and applies it to a sales dataset to compute counterfactual contracts. Georgiadis and Powell (2022) establishes conditions under which the information provided by an A/B test of incentive contracts is sufficient for determining how best to improve a status quo incentive contract in the classic Holmström model, and assesses the empirical relevance of these conditions using data from a real effort experiment.

4 Discussion

We conclude with a discussion of additional related literatures and an outline of previous surveys.

CEO Pay. Several researchers have sought to develop and test models tailored to CEO pay. Edmans, Gabaix and Landier (2009) proposes a model with multiplicative preferences, and derives an empirical measure of incentives that is calculated as the dollar change in CEO wealth for a percentage change in firm value divided by annual flow compensation. This measure is independent of firm size, and thus comparable across firms and over time. See also Murphy (1999), Murphy (2013), and Edmans, Gabaix and Jenter (2017) for surveys of theoretical and empirical findings on CEO pay, and Edmans, Gosling and Jenter (2021) for a survey of firm directors and investors about the practical considerations in designing and evaluating pay packages.

Efficiency Wages. Efficiency wage theories argue that firms may raise wages above competitive levels to motivate their workforce, to attract and retain more productive workers, or to foster workers’ 15

---

15 See also D’Haultfoeuille and Fevrier (2020) who uses variation in contracts to estimate the losses associated with using linear contracts when workers are risk neutral.
16 For details on common estimation methodologies, see Hotz and Miller (1993), Bajari, Benkard and Levin (2007), and Arcidiacono and Miller (2011).
17 See also Antic and Georgiadis (2022) who reconsiders this question in a setting where the principal is oblivious to the agent’s action set and their costs, and evaluates contracts by their worst-case payoff. In this case, the optimal contract is a convex combination of one of the original contracts and a linear one.
18 This measure is available for a large number of firms at https://alexedmans.com/data/. The authors argue that multiplicative preferences are necessary to deliver empirically consistent predictions for the scaling of various incentive measures with firm size.
loyalty toward the firm (Raff and Summers 1987 and Krueger and Summers 1988). In particular, paying supra-competitive wages increases one’s cost of losing their job which, coupled with appropriate monitoring and firing policies, can motivate higher effort (Shapiro and Stiglitz 1984). A critique however is that efficiency wages are typically dominated by other contractual forms (Murphy and Topel 1990).

Behavioral Frameworks. A substantial literature investigates behavioral responses in contracting settings. Psychologists and sociologists argue that performance pay is counterproductive because it undermines “intrinsic motivation”; see for example Gerhart (2017) and the references therein. To rationalize this view, Benabou and Tirole (2003) shows how performance incentives offered by an informed principal can adversely impact an agent’s perception of the task or of his own ability, and ultimately undermine his motivation. For example, the agent might rationally infer that he is being paid a bonus because the task is unenjoyable, and this inference can crowd out his intrinsic motivation. See also Benabou and Tirole (2006) for an alternative model. Another strand studies agency problems with non-standard preferences. For example, in Englmaier and Wambach (2010) there are many inequity-averse agents, whereas in Englmaier and Leider (2012) the agent has reciprocal motivations. See also Fehr, Goette and Zehnder (2009) for a survey of the literature with fairness concerns, and Kamenica (2012) for a review of empirical findings on the anomalous impacts of incentives.

Payments in Kind. Pecuniary incentives such as bonuses, commissions, piece rates, and promotions are only one way to motivate workers. Firms often complement them with various non-monetary incentives. The “delegation literature” explores the motivating effects of granting subordinates decision-making power; see, for example, Aghion and Tirole (1997), Bester (2009), Li, Matouschek and Powell (2017), and Lipnowski and Ramos (2020) for theoretical analyses, and Fehr, Herz and Wilkening (2013) for an experiment. Others have studied the motivating effects of paying (partially) with knowledge such as apprenticeships (Fudenberg, Georgiadis and Rayo 2021). A more recent strand studies how feedback and information can complement financial incentives to reduce agency costs (Fuchs 2007 and Ely, Georgiadis and Rayo 2022). Another strand explores the motivating effects of symbolic awards; see for example Gallus (2017) for an experiment at Wikipedia, and Frey and Gallus (2017) for a survey.

References


