Essays on Incentives for Effort Provision in Principal-Agent Settings

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Economics

by

Troy A. Kravitz

Committee in charge:

Professor Joel Sobel, Chair
Professor Gordon Boyack Dahl
Professor Gary Charles Jacobson
Professor Branislav Lazar Slantchev
Professor Joel Watson

2013
The dissertation of Troy A. Kravitz is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

Chair

University of California, San Diego

2013
# TABLE OF CONTENTS

Signature Page ........................................ iii

Table of Contents ...................................... iv

List of Figures .......................................... vi

Acknowledgements ....................................... vii

Vita ....................................................... viii

Abstract of the Dissertation ........................... ix

Chapter 1  Incentives for Spot Market Labor When Output is Unverifiable 1
1.1 Introduction ........................................... 2
1.2 Literature ............................................. 5
1.3 Example ............................................... 8
1.4 Model .................................................. 12
  1.4.1 Statement of the problem ......................... 12
  1.4.2 The formal model .................................. 12
1.5 Analysis .............................................. 19
  1.5.1 Implementation .................................... 20
  1.5.2 Optimization ...................................... 24
1.6 Discussion ........................................... 28
  1.6.1 Hierarchy ......................................... 31
  1.6.2 Virtual monitoring ................................. 32
  1.6.3 Mediated contracts and the “Gold Standard” .... 33
  1.6.4 Alternative models and robustness ............... 34
  1.6.5 Other extensions ................................... 36
1.7 Conclusion ........................................... 39

Chapter 2  Lobbying for Influence with Strategic Lawmakers .......... 41
2.1 Introduction .......................................... 41
2.2 Background ........................................... 45
2.3 A model of lobbying with endogenous valuations ........... 46
2.4 Example ............................................... 49
2.5 Analysis .............................................. 50
  2.5.1 Endogenous valuations but no discrimination ... 50
  2.5.2 Endogenous valuations and endogenous discrimi-
       nation .............................................. 53
2.6 Related literature .................................... 58
2.7 Conclusion ........................................... 59
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>The relationship between tasks, worker-tasks and jobs.</td>
<td>13</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Distribution on ${R^i_{kj}}_{k=0,...,J,j\leq k}$ induced by $(\sigma^i, \overline{\sigma}^i)$ and $Q^i$.</td>
<td>21</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Composition of per-task wage bill.</td>
<td>25</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Optimal monitoring probability as a function of the job size.</td>
<td>26</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Efficiency loss as a function of the job size.</td>
<td>27</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Bidding strategies with $\gamma$ set optimally when there is no cap.</td>
<td>54</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Comparing $\gamma$ set optimally versus an increase in $\gamma$, with binding cap.</td>
<td>56</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Bidding strategies with $\gamma$ set optimally when there is a binding cap.</td>
<td>57</td>
</tr>
<tr>
<td>Figure A.1</td>
<td>Example job posting on Amazon’s Mechanical Turk.</td>
<td>63</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I’d like to thank my advisor, Joel Sobel, and committee members Gordon Dahl, Gary Jacobson, David Miller, Sebastian Seigh, Branislav Slantchev, and Joel Watson for helpful guidance throughout this process. I feel like I’ve learned more from conversations with Joel about “science” than I have from any economics studies. I’ve benefitted greatly from the UCSD theory group, as well as from countless discussions with my various coauthors: Aislinn Bohren, Tim Keller, Chulyoung Kim, and Branislav Slantchev. In particular, Chapter 1, “Incentives for Spot Market Labor When Output is Unverifiable,” is coauthored with J. Aislinn Bohren. Chapter 2, “Lobbying for Influence with Strategic Lawmakers,” is coauthored with Chulyoung Kim. I look forward to continued common endeavors.

This was a long process. The support and encouragement of my family and friends has meant a lot to me during the journey. Sandra has put up with me being – essentially – a perpetual student since we met. (Unfortunately, being a student is not just about discounted movie tickets.) Charlotte and Penny, through trips to the aquarium and walks along the beach, have benefitted from the same. My cohort at UCSD provided continual fun (often disguised as intellectual engagement) and intellectual engagement (sometimes disguised as fun). No problem set could ever provide intrigue, back-stabbing, or conversations superior to those provided by board game nights.

Financial support from UCSD, the Department of Economics, and the Institute for Humane Studies is gratefully acknowledged.
## VITA

<table>
<thead>
<tr>
<th>Year</th>
<th>Education/Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>Bachelor of Arts in Economics <em>summa cum laude</em>, Emory University</td>
</tr>
<tr>
<td>2007–2013</td>
<td>Teaching assistant, Department of Economics, University of California, San Diego</td>
</tr>
<tr>
<td>2013</td>
<td>Doctor of Philosophy in Economics, University of California, San Diego</td>
</tr>
</tbody>
</table>
ABSTRACT OF THE DISSERTATION

Essays on Incentives for Effort Provision in Principal-Agent Settings

by

Troy A. Kravitz

Doctor of Philosophy in Economics

University of California, San Diego, 2013

Professor Joel Sobel, Chair

Situations in which multiple parties with competing preferences interact are endemic throughout society. These essays consider the problem of a principal who seeks to induce self-interested agents to exert effort or promise contributions. The principal’s task is to create an incentive structure that aligns the agents’ preferences with his own.

Chapter 1 considers a principal seeking to induce agents to exert costly, unobservable effort when the output they produce is unverifiable. The firm’s solution is to hire multiple workers for some tasks and compare the output produced by the agents. The optimal mechanism bundles multiple tasks together to reduce the cost of monitoring and conditions wage payment for any task upon satisfactory completion of all tasks. The optimal mechanism is not efficient as the principal
prefers greater duplication of tasks in exchange for reduced worker rents. Asymp-
totically, as jobs grow large, the firm approximates its first-best payoffs from the
contractible effort benchmark. It is only in the limit that the optimal mechanism
is also efficient.

Chapter 2 studies the campaign finance landscape following recent changes
to the law. A discriminatory all-pay contest model with a contribution cap is used
to study legislative and lobbying behavior. The principal, a strategic lawmaker,
is central to the analysis. The lawmaker both designs the prize the lobbyists are
competing to obtain – and, in doing so, determines their valuations for the prize
– and determines the terms of the prize’s allocation. Contrary to existing work,
expected contributions always increase as the contribution cap is relaxed. The
effect on policy is more nuanced and depends on whether the cap is binding. The
analysis highlights that a decrease in competitive forces, for example, from one
lobbyist being richer or valuing the prize more, can be only partially offset by
providing a discriminatory benefit to the other lobbyist. The strategic lawmaker
endeavors to prevent such an imbalance from arising in the first place.
Chapter 1

Incentives for Spot Market Labor When Output is Unverifiable

Abstract

This paper studies an agency problem in which a firm employs workers to perform a task. The worker selects an effort level that determines the probability that she completes the task successfully. The firm cannot directly observe effort or verify whether the task is successful. To encourage effort, the firm hires several workers to perform the same task and bases compensation on the degree to which the workers’ output agrees. The setting differs from traditional agency theory along two dimensions: no signal about whether effort was exerted is available, and the firm is unable to threaten workers with negative wage payments. In the optimal mechanism, the firm endogenously monitors workers with nontrivial probability and bundles multiple tasks together to reduce the per-task cost of monitoring. By assigning more tasks to each worker and conditioning wage payment for any task on satisfactory performance on all tasks, the firm approximates its first-best payoffs even when firing and large punishments are unavailable.
1.1 Introduction

Technological advancement is changing the way modern companies interact with their workforce. Traditional in-house employees only recently gave way to outsourced workers contractually tied to the firm. Further changes are afoot: new information technology creates the opportunity for firms to access a flexible and inexpensive pool of workers on-demand. One example is crowdsourcing, where firms hire temporary workers through open calls online. Millions of potential employees are available around the clock and able to start work immediately. Neither a pre-existing relationship nor a continuing relationship is required. In the absence of conventional methods of supervising employees, tapping into this global workforce presents a host of new incentive issues.

Much of the allure of spot labor markets is that interaction between the firm and its employees is minimal. This is also a liability as reputation mechanisms have little bite. Workers are compensated for their effort, but the exertion of effort is unobservable and no direct signal about whether effort was exerted is available. The firm must guard against shirking, but how?

Traditional agency theory offers limited guidance for firms interacting in these spot markets. Economic theories of crime, like those of Becker (1968) and Kolm (1973), for example, suggest optimal enforcement regimes combining infinitesimal monitoring with arbitrarily harsh punishments. Mirrlees (1974) shows that this principle applies broadly in principal-agent settings. Yet actual firms operate quite differently from the theoretical recommendations (Dickens et al., 1989). Considerable resources are devoted to monitoring employees and firms are unwilling or unable to impose sufficiently harsh punishments to stamp out undesirable employee behavior. This suggests the problem is appropriately framed with constraints on the firm’s available punishments.

We study the challenges faced by a firm hiring in a spot market by viewing the situation within a principal-agent context. The firm has a series of items to categorize. Each item can be thought of as having a state that is unknown to the firm but that can be observed by a worker who exerts costly effort. Worker payoffs are independent of the state and protected by limited liability clauses,
effort is non-contractible, and output is unverifiable. Each worker is offered a contract consisting of tasks to be completed, a monitoring protocol specifying how many workers are to be employed for each task, a verification technology to review output, and a payment schedule.

The setting is similar to multilateral principal-agent problems except that the firm lacks a direct signal about worker quality. In the agency literature, the principal typically receives a direct signal about output that is related to unobserved effort. With no such signal available, the firm draws inferences about the quality of worker input by hiring multiple workers to complete the same task. In effect, the firm generates a costly signal about effort by monitoring workers with other workers. Correlation in agents’ information is used to discipline behavior without resort to large negative payments as threats.

The main result of the paper derives the firm’s optimal organization of incentives. The optimal mechanism exhibits two features absent in standard contracting solutions: endogenous monitoring and bundling. Effort cannot be induced without hiring additional workers. The firm assigns an agent a collection of tasks to complete. It then hires another worker to duplicate a subset of these tasks and pays the agent only for satisfactory performance on all tasks. Producing evidence of shirking is costly for the firm and the amount and efficacy of employee monitoring is determined endogenously.

Wages depend on the frequency of monitoring: higher wages are needed as the firm monitors less frequently. At the optimum, the firm employs monitors frequently enough to keep wages down but infrequently enough to prevent excessive duplication of assignments. The mechanism is wasteful compared to the observable effort benchmark: not only must additional workers be hired, but each must be paid a wage greater than his cost of effort.

Alchian and Demsetz (1972) suggest that there should be specialization in monitoring. This is not the case in our setting: the optimal incentive organization of the firm treats workers symmetrically. The distinction between monitors and subordinates is solely expository. Armen monitors Harold while, at the same time, Harold monitors Armen. No hierarchy of monitors is created.
Information aggregation is often emphasized as a motivation for hiring multiple agents (Rothschild, 1974; Aghion et al., 1991; Grossman et al., 1977). In the canonical setting of statistical decisionmaking, a gambler tests a slot machine multiple times to acquire improved information about the machine’s unknown odds. The question is when to move on to another machine if the gambler’s objective is to maximize winnings. As both the principal and the agent, the gambler does not need to provide incentives for the agent to undertake the principal’s desired behavior: experimentation is solely for learning reasons.

In the model of Section 1.4, by contrast, there is no learning justification for the duplication of tasks: multiple workers are employed solely for incentive reasons. The setting without noise can be interpreted as the opposite pole to the setting of Robbins (1952). The paper thus provides a complementary explanation for why a principal may consult multiple agents before taking a decision.

Another contribution of this paper is to highlight the role played by combining multiple tasks into one job. Bundling tasks together permits the principal to reduce the cost of monitoring. Specifically, the firm assigns multiple tasks to each worker and hires another worker to perform only a few of these tasks. The worker’s payment is then conditioned on his performance on the tasks that are monitored.

The firm cannot punish workers by assessing penalties for poor performance. The most it can do is withhold wages. So while the firm generates incentives by stochastically duplicating an agent’s work, tying the agent’s wage for one task to his performance on all tasks strengthens these incentives. Monitoring and bundling are strategic substitutes: the firm hires monitors less frequently as the number of tasks assigned to an agent increases. The efficiency loss relative to the contractible effort benchmark vanishes asymptotically.

The theoretical setting is presented within the context of spot labor markets in general and the online crowdsourcing marketplace in particular. Spot markets for labor are already big business: in crowdsourcing, where firms and workers interact through the web, worker earnings are in the billions, the revenues of vendors matching firms to workers were estimated at $500 million in 2009 (Frei, 2009) and
venture capital firms injected almost $300 million in 2011 (Sanders, 2011). It is not surprising that industry insiders believe one-third of the global workforce could be hired online by 2020 (Vanham, 2012).

The mechanisms we identify offer potentially significant improvement over those currently in practice. Firm can reduce their monitoring expenditures by structuring contracts so that individual workers check each other. By shifting from piece-rate payment schedules to schemes requiring satisfactory performance on all tasks, firms can recreate the same incentives at lower cost. The theoretical performance improvements can be empirically tested through field experiments carried out online.

The main findings of the paper apply to other settings as well. For example, considering a multidimensional chore instead of a series of tasks, the firm monitors individual components of the chore and punishes workers across all dimensions for poor performance on any dimension. Tax auditing practices, hiring rules-of-thumb and the monitoring of rider provisions are examples of the main ideas of the paper applied to this setting.

The paper proceeds as follows. Related literature is discussed in Section 1.2 and a representative example is presented in Section 1.3. The example isolates key properties of the setting. The formal model is introduced in Section 1.4 and the firm’s optimal incentive organization is derived in Section 1.5. Discussion and extensions follow in Section 1.6. All proofs are relegated to Appendix A.2 following the conclusion. Appendix A.1 contains a brief overview of a setting of interest, the crowdsourcing marketplace.

1.2 Literature

In the standard agency model, an agent selects action $a$ and the principal observes $s$, an informative signal about $a$. The principal cares about output $y$, which depends on $a$ and $s$. The principal’s problem is to generate incentives for a particular $a^*$ using only a function of $s$. Actions can be interpreted broadly; $a^*$ is

\footnote{Mirrlees (1975), Mirrlees (1976) and Holmstrom (1979) are seminal papers in the literature on moral hazard.}
commonly an optimal level of investment or effort. The signal $s$ may be something like sales volume or peer evaluations. As long as the conditional distribution of the signal $p(s|a)$ varies with $a$, the firm is able to align incentives so that $a^*$ is the employee’s voluntary action choice.\(^2\)

The characterization of optimal contracts in the general setting is limited.\(^3\) Instead, the literature provides a host of elegant solutions for specific settings. Many of the mechanisms identified are variations on a common dynamic: if for each $a$ there exists $s$ such that $p(s|a^*) \neq p(s|a)$, the firm is able to statistically discriminate among worker actions. When $s$ is realized, the firm inserts a wedge until $a$ is no longer attractive. Bad signals – those that are more frequent when $a$ is taken instead of $a^*$ – are punished while good signals are rewarded.

Sometimes the observable signal $s$ only reveals aggregate information about actions instead of individual action choices. The firm is now in a multilateral contracting environment and must guard against the free-rider problem. Holmstrom (1982) emphasizes the role of group penalties: all workers are punished whenever bad signals obtain. Group penalties are natural here since the firm cannot discern which worker deviated. They continue to play a role elsewhere, especially when identifying the deviator is more costly than identifying that a deviation occurred.

A practical, though not theoretical, complication arises if the conditional distribution of $s$ given $a$ is very similar to that given $a^*$. Here is it difficult for the firm to distinguish whether $a$ or $a^*$ was chosen by the worker. In order to dissuade $a$, the firm must threaten large punishment in the off-chance it believes $a$ more likely than $a^*$. Larger and larger punishments are required as the signals triggering punishment become more rare.

The incentive contracts above break down if the firm is limited in how severely it can punish an employee. One response is that instead of punishing a worker by paying him $-x < 0$ when a bad signal obtains, the firm transfers $x$ to the worker at the outset and simply takes away this transfer upon observing a bad signal. Such an arrangement preserves incentives, so if the contract without

\(^2\)This often introduces a distortion away from the optimal contract with contractible effort since the worker needs to be compensated for the risk in his wage payment.

\(^3\)See Bolton and Dewatripont (2005, Ch. 4).
limited liability is able to induce $a^*$, then so is the modified contract satisfying limited liability. But note that the firm’s expected payments have now increased by $x$: hiring workers may no longer be attractive at all.

In standard agency models, the principal cannot achieve the first-best equilibrium obtained when actions are contractible. Correlation in agents’ valuations or information provides a way for the principal to recover first-best. 4 Viewed in isolation, an agent’s action or signal thereof is uninformative about the underlying uncertainty. But viewed in conjunction with those of other agents, the principal is able to draw inferences about the hidden object.

Legros and Matthews (1993) study a partnership problem between the poles of individual and aggregate signals. Each partner privately devotes effort to a common project. Allocating credit for success is impossible except when only one partner is contributing to the project. Randomly with small probability all partners but one shirk, perfectly revealing the action of the working partner, who is then punished according to his effort provision.

The signal is commonly produced without cost as a byproduct of action choices. 5 In this respect, Legros and Matthews (1993) is an example of costly monitoring. Rahman (2012) also considers a costly monitoring setting; in an ideal arrangement, a worker exerts effort and the firm never monitors the worker. This would be the action profile with contractible effort. When payments cannot be made contingent on actions – perhaps because actions are unobservable – this profile cannot be achieved. Rahman (2012) shows that the profile can be approximated arbitrarily well by using the dynamic described earlier. The signal $s$ is now the report of a monitor periodically hired to verify the worker’s action. 6 The worker is punished severely whenever the monitor reports that he shirked.

The signal structure can be viewed as the firm’s monitoring technology. In much of the literature, regardless of whether signals are a costless byproduct of ac-

---

4 The surplus extraction literature illustrates, applying the dynamic described earlier to the bid profile from an auction where bidders have correlated valuations of the item being sold (Cremer and McLean, 1988; McAfee and Reny, 1992; Riordan and Sappington, 1988; Bose and Zhao, 2007).

5 See, for example, the subjective evaluation literature (Prendergast, 1999; MacLeod, 2003).

6 The relationship of the present paper to Rahman (2012) is discussed further in Section 1.6.
tions, the monitoring technology is exogenous. Early studies in which monitoring is a choice variable include Becker (1968), Kolm (1973) and Mirrlees (1974);\(^7\) these papers suggest combining infinitesimal monitoring with arbitrarily harsh punishments.

The observed incentive structures of firms bear little resemblance to these arrangements (Dickens et al., 1989). Even if arbitrarily harsh punishments are possible in practice, they may still be undesirable since they flatten the penalty gradient between minor and major offenses (Stigler, 1970). Becker and Stigler (1974), Carr-Hill and Stern (1979) and Carmichael (1985) provide other reasons to question the appropriateness of unlimited liability.

This paper identifies the optimal incentive organization for a firm \((i)\) bound by limited liability and in the absence of repeated interactions and \((ii)\) in a setting where costly monitoring is necessary to produce informative signals. Punishments entail taking away rents when a deviation from \(a^*\) is observed. The punishments are strengthened by tying the payment for one part of the job to performance on all parts of the job. This dynamic is similar to that identified by Fuchs (2007) in a repeated setting; there, Fuchs (2007) shows it is optimal for the firm to withhold payment until the final period.\(^8\)

### 1.3 Example

Each day users upload thousands of images to an auction website. The firm controlling the website needs to ensure the images are not obscene. A workforce of thousands is available around the clock in an online labor market.

As new item listings are created, the uploaded images pass through a central database for verification. Multiple images can be collected together and sent to workers as a package, but the firm is unwilling to delay approval of prospective listings and so limits each worker to reviewing no more than 10 images at a time.

---

\(^7\)The efficiency-wage theory of Shapiro and Stiglitz (1984) is sometimes portrayed as an example of endogenous monitoring. See, for example, Bolton and Dewatripont (2005, § 4.1.3).

\(^8\)Abreu et al. (1991) emphasize the reusability of punishments: one punishment can simultaneously provide incentives across many periods.
Workers verify content by viewing the image. This process takes time, the cost of which is $0.02, but accurately reveals objectionable content without fail. Workers and the firm know that most images are acceptable and that, on average, only 5% will be flagged as obscene. They also know that it is not cost-effective for the firm to verify a worker’s suggestion about the status of an image; that’s why the firm is hiring them to act on its behalf in the first place.

Payoffs are as follows: the risk-neutral firm earns a payoff of 1 for correctly classifying an image (i.e., forbidding obscene content or approving acceptable content) and a payoff of 0 otherwise, less any wages promised to workers. Risk-neutral workers do not care whether the firm treats an image correctly: payoffs are their received wages less the cost of their effort.

We now consider various settings and identify how the firm optimally structures incentives within each setting. A labor contract consists of the number of images a worker must review, a wage for the job and the conditions for payment.

**Observable effort**

First, suppose that the firm can observe whether a worker viewed the image. The firm pays a worker their time cost of effort $0.02 for each image he views and nothing for the images he skips. The firm does not need to hire additional workers for any task and there is no benefit through reduced wages per task to assigning a bundle of images to the worker instead of a solitary image.

**Unobservable effort and unlimited liability**

Now suppose effort is unobservable. Since the firm cannot verify the accuracy of their recommendations, workers have incentive to bypass viewing the images and randomly report their status. The firm must find a way to monitor the workers it hires.

If the firm assigns the same task to two workers, it can compare their recommendations. Negative wages are permissible with unlimited liability, and so the firm offers the following contract. For each image the worker is assigned, with probability $q > 0$ the firm hires a second agent to report the image’s status. The
worker is paid $0.02 unless his recommendation disagrees with that of the second agent, in which case the worker is assessed a penalty of $0.02(1 - 1/0.05q). The firm hires $1 + q$ workers in expectation for each image. As $q \to 0$, the firm hires approximately one worker and compensates him at his cost of effort. Again, there is no benefit to assigning multiple images to a worker.

**Unobservable effort and limited liability**

The firm can no longer threaten workers with negative wage payments. Suppose the firm follows the incentive structure from above, hiring a second worker with probability $q$ and paying him $0.02$ unless his recommendation differs from that of the second worker. Since the firm cannot penalize the worker with a negative wage, the worker will guess the status of the image and hope that either he is unmonitored or that his guess is correct. Thus, the worker requires a premium above his cost of effort in order to be induced to exert effort. Additionally, the firm must hire additional workers for each task with nontrivial probability. In the current example, a wage of at least $0.40$ per image is required by each worker. These losses relative to the observable effort benchmark are a general feature of limited liability when effort is unobservable.

The firm can do better by leveraging its ability to assign multiple images to each worker. Two situations are worth highlighting. First, if the firm can make the worker bear the risk that a second agent will not be called upon to verify his recommendation, the firm can offer the worker a contract that pays him if and only if a second agent is hired and verifies his recommendation on every assigned task. The firm promises to hire a second agent independently for each task with probability $q > 0$ and pays the worker $0.2/q^{10}(1 - .95^{10})$ if the criteria for payment are satisfied.

As $q \to 0$, the firm hires a second worker for each task less often and fewer workers meet the criteria for payment. The firm’s expected payments per image approach $0.02/(1 - .95^{10}) \approx 0.05$. When workers are protected by limited liability, the firm can do no better than this arrangement. Note that with unlimited liability there is no benefit to the firm of having workers bear the risk of monitoring.
This incentive structure is implausible for several reasons. It imposes upon workers potentially undesirable ex ante variation in wages. When the firm does not verify the worker’s recommendation, then no wages are paid, even though the failure to satisfy the payment criteria is not due to negligence of the worker. The firm has incentive to induce a breach of contract by not monitoring in this setting, seriously straining the credibility of the firm’s commitment to hire additional workers.

**Unobservable effort, limited liability and limited risk**

If, instead, workers cannot be made to bear the risk that the firm does not monitor their output, the optimal frequency with which the firm hires a second worker is endogenously determined. At the optimum, the firm’s cost per image is just below $0.10.

In sum, when effort is observable, the firm’s cost per image is $0.02, no additional workers are needed, and there is no benefit to bundling multiple images together. This expected cost can be approximated arbitrarily well with unobservable effort by threatening workers with large fines if their recommendation differs from that of another worker. Additional workers are required with arbitrarily small probability and assigning multiple images to workers provides no benefit to the firm. If the firm continues to hire workers for single images, the firm’s expected cost per image increases to $0.80 when workers are protected by limited liability.

Bundling provides a way for the firm to mitigate the costly effect of workers’ limited liability protection. By assigning multiple images to each worker, the firm is able to tie a worker’s compensation on one task to successful completion of all tasks. This strengthens the incentives it can offer regardless of whether workers can be made to bear the monitoring risk.
1.4 Model

1.4.1 Statement of the problem

A firm is faced with an infinite stream of tasks, and seeks to learn some unknown attribute of each task by delegating this work to agents. For example, a social media firm must moderate the content of user-generated images; the firm hires workers to determine whether each image meets content guidelines. An agent can either exert costly effort on an assigned task, in which case the agent learns the true attribute of the task (i.e. whether the image is acceptable), or not exert effort, in which case the agent learns nothing. The agent then sends a message to the firm about the relevant attribute.

Effort is unobservable, an agent’s message (in isolation) conveys no information about his effort choice and the agent’s preferences are independent of the task’s unknown attribute. Therefore, there is no incentive structure that will compel an agent to exert effort when he acts in isolation. However, delegating a task to multiple agents and comparing their messages will generate a signal of their effort choices, which can be used to effectively structure incentives.

The firm’s problem is to design a labor contract that optimally allocates tasks to workers and compensates workers based on their messages. The firm must decide how many tasks to assign to each worker, how many workers to hire for each task, and a compensation schedule that conditions payment on the profile of messages for each task assigned to a worker. This general multilateral contracting problem is presented formally in the next subsection.

1.4.2 The formal model

A firm is faced with a countably infinite stream of independent and identical tasks \( t = 1, 2, \ldots \) and has access to a countably infinite pool of workers \( i = 1, 2, \ldots \). The firm can design labor contracts to delegate tasks to workers. There are two formal components to the model: the task \( t \) and the job \( J^i \), or the set of tasks...
A task can be assigned to multiple workers, and multiple tasks can be assigned to a single worker. We use the phrase worker-task to refer to an individual worker’s decision problem on a each task within a job (equivalently, the individual decision problem for each worker assigned to a task). Worker-tasks partition both tasks and jobs. Figure 1.1 describes these components visually.

---

Figure 1.1: The relationship between tasks, worker-tasks and jobs.

---

9The following convention is maintained: objects pertaining to tasks are subscripted and objects pertaining to workers are superscripted.
The Task

On each task \( t \), there are two possible states of the world, \( \omega_t \in \Omega \equiv \{0, 1\} \) with common prior belief: \( \pi = Pr(\omega_t = 1) \in [1/2, 1) \).

The firm seeks to match the realized state of the world by choosing an action \( A_t \in \Omega \). It receives a payoff of 1 if its action matches the realized state, and a payoff of 0 otherwise.\(^{12}\) Workers’ payoffs are independent of the state.

Before selecting an action, the firm can hire workers to learn about the unknown state. For each task \( t \), the firm chooses a probability measure \( f_t \in \Delta(\mathbb{N}) \) specifying a distribution over the number of workers to hire for \( t \), where \( \Delta(\mathbb{N}) \) is the set of all probability measures over the nonnegative integers. Let \( q_t = \sum_{n=2}^{\infty} f_t(n) \) be the probability that multiple workers are hired for task \( t \) and \( n_t \in \mathbb{N} \) be the realized number of workers hired for task \( t \). Once the number of workers has been decided upon for a task, each worker-task is assigned to the next available worker. The set \( \mathcal{I}_t \) represents the set of workers hired for task \( t \). Workers and tasks are identical; therefore, we do not distinguish between the identity of different workers or tasks going forward.

The worker-task: A worker \( i \) assigned to task \( t \) chooses whether to exert high effort (\( e_i^t = 1 \)), in which case he perfectly observes the state of the world and incurs costs \( c_i \), or low effort (\( e_i^t = 0 \)), which yields no information about the state and is costless. Let \( s_i^t \in \{\Omega \cup \emptyset\} \) be the information worker \( i \) observes about task \( t \). Effort choices are unobservable by the firm and information is not verifiable. After making an effort choice and observing information about the state, the worker sends a message to the firm, \( m_i^t \in \Omega \).\(^{13}\)

Back to the task: Upon receiving messages from all hired workers on a task, the firm compiles a message profile,

\[ m_t = \{m_i^t\}_{i \in \mathcal{I}_t} \in \mathcal{M}_t \]

\(^{10}\)For example, whether an image meets content guidelines.

\(^{11}\)Section 1.6.5 extends the model to settings where all tasks are not identical.

\(^{12}\)Section 1.6.5 allows the firm to have asymmetric payoffs for matching each state.

\(^{13}\)With limited liability, it is without loss of generality to restrict the message space to \( \Omega \).
to inform its action choice, where $\mathcal{M}_t = \bigcup_{n=1}^{\infty} \Omega^n$ is the set of all possible sets of messages for task $t$.$^{14,15}$ The firm’s strategy on task $t$ is a mapping from the set of message profiles into the set of mixed actions,

$$\alpha_t : \mathcal{M}_t \to \Delta(\Omega).$$

Let $\alpha = (\alpha_1, \alpha_2, ...)$ represent the firm’s strategy profile.$^{16}$

**The Job**

A job $\mathcal{J}^i$ consists of the set of tasks assigned to worker $i$, where $J^i = |\mathcal{J}^i|$ denotes the size of the job. We restrict the maximum job size the firm can create to be $J < \infty$ to capture exogenous legal or technological constraints that the firm may face.$^{17}$ Once a worker receives his job, he completes each worker-task and sends a message profile,

$$m^i = \{m^i_t\}_{t \in \mathcal{J}^i} \in \mathcal{M}^i$$

to the firm, where $\mathcal{M}^i = \Omega^{|\mathcal{J}^i|}$ is the set of all possible sets of messages for a job.

The worker’s strategy specifies an effort level and a message for each task in his job. The effort strategy specifies the probability of exerting effort on each task,

$$\sigma^i = (p^i_1, ..., p^i_{|\mathcal{J}^i|}) \in [0,1]^{|\mathcal{J}^i|}$$

where $p^i_t = Pr(e^i_t = 1) \in [0, 1]$ is the probability worker $i$ exerts high effort on task $t$. The strategy $\sigma^i$ induces a distribution on the number of assigned tasks for which the worker exerts effort. Let $\zeta(k|\sigma^i)$ be the expected probability of exerting

$^{14}$Note $\Omega^n$ denotes the cross product of $\Omega$, repeated $n$ times.

$^{15}$As noted above, worker identity is irrelevant, and $\mathcal{M}_t$ is independent of the identity of the workers in $\mathcal{I}_t$.

$^{16}$In principal, the firm’s strategy on task $t$ could also depend on the message profiles from other tasks. We preclude this possibility by defining the firm’s strategy on task $t$ as a mapping from the messages for task $t$ to the action space. It is without loss of generality if all workers with the same contract choose the same effort strategy, and this effort strategy is symmetric across tasks within a job. Given such an effort strategy, messages on other tasks convey no information about a worker’s effort choice on task $t$. The firm cannot improve its action choice on task $t$ by examining the worker’s performance on other tasks. We are interested in characterizing the optimal high effort equilibrium, hence, the restriction is without loss of generality here. Note that we will allow the firm to design contracts specifying payments that depend on all tasks within a worker’s job.

$^{17}$Workers may not complete multiple jobs, so there is no scope for repeated interaction.
effort on \( k \) tasks when following \( \sigma^i \). Define \( \sigma = (\sigma^i, \sigma^{-i}) \) and \( \tilde{\sigma} = (\sigma^i, \sigma^{-i}) \) as the strategy profiles for always working and always shirking, respectively. The message strategy specifies a message profile as a function of the information observed about all tasks\(^{18}\)

\[
\rho^i : (\Omega \cup \emptyset)^{J_i} \rightarrow \Delta(M^i).
\]

Let \( \tilde{\rho}^i \) denote the strategy where (1) the agent reports signals truthfully for tasks on which he exerts effort and (2) the agent reports the state with a higher prior for tasks on which no effort is exerted.

**Contracts**

So far, we have discussed how a firm chooses the number of workers to hire for each task, and the number of tasks to assign to each worker (the job). These are the first two components of a worker’s contract; the third component is the payment scheme.

Effort is unobservable and information is not verifiable; therefore, the payment scheme can only condition on the message profile of a worker and the messages of the other workers assigned to each task in the worker’s job. The firm constructs a report for worker \( i \),

\[
r^i = \bigcup_{t \in J^i} m_t = \bigcup_{t \in J^i} \bigcup_{i \in I_t} m^i_t \in \mathcal{R}^i
\]

where \( \mathcal{R}^i = \mathcal{M}_t \times ... \times \mathcal{M}_t \).\(^{19}\) The firm offers a payment scheme as a function of this constructed report. Let

\[
W^i : \mathcal{R}^i \rightarrow \mathbb{R}_+
\]

specify the payment obtained by worker \( i \) for job \( J^i \) when report \( r^i \) is received. We assume that workers are protected by limited liability, and restrict attention to payment schemes \( W^i(r^i) \geq 0 \) for all \( r^i \). It will often be convenient to work with the per-task wage \( w^i(r^i) = \frac{W^i(r^i)}{J^i} \).

---

\(^{18}\)The information partitions the effort choice of the worker, so it is without loss of generality to define the message as a function of only information.

\(^{19}\)Reports for different workers will contain common elements when the workers have overlapping tasks.
The strategies \((\sigma^i, \sigma^{-i}, \rho^i, \rho^{-i})\) induce a distribution over \(\mathcal{R}^i\), which is denoted \(G(r^i|\sigma, \rho)\).

Define an implementation plan for worker \(i\) as the collection of probability measures over the number of workers hired for each task in worker \(i\)'s job,

\[ Q^i = \{f_t\}_{t \in J^i}. \]

Then we can formally represent a contract for worker \(i\) as a set of tasks, an implementation plan and a payment scheme,

\[ C^i = (J^i, Q^i, W^i(\cdot)). \]

It is useful to partition \(\mathcal{R}^i\) into subsets \(\{\mathcal{R}^i_{kj}\}_{k=0}^{J^i}, j \leq k\). \(\mathcal{R}^i_{kj}\) is the collection of reports from worker \(i\) where additional agents were hired for \(k\) tasks and all workers make the same recommendation on \(j \leq k\) of these tasks; define \(\overline{\mathcal{R}}^i = \bigcup_{k=0}^{J^i} \mathcal{R}^i_{kk}\).

**Payoffs**

A worker’s payoff for a job \(i\) depends on the transfer received from the firm and the cost of effort. We assume workers are risk neutral, so a worker’s expected utility, given a contract \(C^i\) and strategy profile \((\sigma, \rho)\), is represented as:

\[ U^i(\sigma, \rho; C^i) = \mathbb{E} \left[ W^i(r^i) - c \sum_{t=0}^{J^i} 1_{\{e'_t = 1\}} \right] \]

Note that a worker’s payoff is independent of the firm’s strategy, \(\alpha\).

The risk-neutral firm seeks to match the realized state of the world for each task. The firm must pay each agent the specified transfer for their employment. The firm’s expected flow payoffs from an incentive contract for \(i\), given worker strategy profile \((\sigma, \rho)\), firm strategy profile \(\alpha\) and contract \(C^i\), is represented as:

\[ V(\alpha, \sigma, \rho; C^i) = \mathbb{E} \left[ \sum_{t=1}^{J^i} \frac{1_{A_t = \omega}}{n_t} - W^i(r^i) \right] \]
The general problem

The firm maximizes its expected flow payoffs from an incentive contract by choosing a job size, monitoring technology, transfer scheme, and action profile for each possible report, all subject to the strategy profile of the hired workers being optimal. That is, the firm chooses \((C^i, \alpha)\) such that:

\[
\max_{C^i, \alpha} V(\alpha, \sigma, \rho; C^i) \tag{*}
\]

subject to

\[
\begin{align*}
(\hat{\sigma}^i, \hat{\rho}^i) & \in \arg \max_{(\hat{\sigma}^i, \hat{\rho}^i)} U^i(\hat{\sigma}^i, \sigma^{-i}, \hat{\rho}^i, \rho^{-i}; C^i) \quad \forall i \tag{IC} \\
U^i(\sigma, \rho; C^i) & \geq 0 \quad \forall i \tag{IR} \\
W^i(r^i) & \geq 0, \quad \forall i, \ r^i \in \mathcal{R}^i \tag{LL} \\
V^*(\alpha, \sigma, \rho; C^i) & \geq J^i \pi. \tag{FP}
\end{align*}
\]

The maximand of \(\ast\) is the firm’s objective function; the remaining expressions capture worker incentive compatibility and individual rationality, limited liability, and firm participation. FP states that the firm must prefer the equilibrium to not entering the contractual relationship.

Section 1.5 considers the class of contracts for which the firm lacks incentive for induced breach. In other words, a report without evidence of shirking is deemed acceptable for the purposes of wage payments. In particular, the firm’s failure to monitor workers is not just cause for withholding payments. The no-induced breach condition is stated as B. It can also be viewed as a limit on the firm’s ability to credibly commit ex ante to ex post costly and unverifiable actions when public randomization over the decision to hire additional workers is unavailable.\(^{20}\)

\(^{20}\)In the optimal equilibrium without B, the firm pays a worker if and only if multiple workers are employed on each of the worker’s tasks and the output they produce always coincides. (See Result 3 in Appendix A.2.) The contracts used in practice bear little similarity to this bounty-like structure (Dickens et al., 1989).

Limited liability alone does not sufficiently constrain the firm’s behavior, suggesting a relevant feature of the firm’s problem is omitted. The contract treats as justification for nonpayment both evidence of shirking and a lack of evidence of working. This is one reason why adhesion
\[ W^i(r^i) = t \forall r^i \in \overline{R}^i. \] (B)

It is worth noting that B places no restrictions on firm behavior when there is evidence of shirking.

**Definition 1 (Equilibrium).** An organizational equilibrium (henceforth equilibrium) is a solution to the firm’s problem satisfying B in which workers exert high effort and report truthfully.

The problem facing the firm remains quite general: the firm is able to stochastically hire workers for each task and offer non-linear payment schemes to reward the agents. Equilibrium continues to take a simple structure and, as Section 1.5 demonstrates, the efficiency loss relative to contractible effort vanishes as jobs grow large.

### 1.5 Analysis

It is useful to begin with the first-best equilibrium when effort is contractible. The equilibrium must only satisfy the participation constraints of the worker and firm (IR and FP). Worker participation is ensured by \( w \geq c \). The firm’s equilibrium payoff for each task is \( 1 - c \), which it prefers to guessing the state blindly for \( c \leq 1 - \pi \).

**Remark 1 (Contractible Effort Benchmark).** When effort is contractible and \( c \leq 1 - \pi \), \( Q^* \) is such that \( q_t = 0 \) for all \( t \) and

\[
W^*(r^i) = \begin{cases} 
J \cdot c & \forall r^i \in \overline{R}^i \text{ and } (\sigma^i, \rho^i) = (\overline{\sigma}^i, \overline{\rho}^i) \\
0 & \forall r^i \in \overline{R}^i \text{ and } (\sigma^i, \rho^i) \neq (\overline{\sigma}^i, \overline{\rho}^i).
\end{cases}
\]

The contractible effort benchmark is unattainable here and, more generally, there is no equilibrium with \( q_t = 0 \) for any task: since a worker’s report reveals contracts, like those considered here, traditionally minimize the actionable obligations of the firm to the payment of money. Posner (1972) and Farber (1980) note that courts will not enforce supercompensatory damages caused by a breach of contract. This reluctance is explained by a desire to avoid creating incentives for “induced breach.”
nothing in isolation about his effort choice, the firm cannot ascertain effort and workers will shirk upon accepting such a contract. The lesson is that the firm must hire additional workers to induce effort. The workers duplicate tasks, enabling the firm to compare output across agents. The firm generates incentives for effort by conditioning payment upon reports it deems acceptable.

Monitoring leads to unavoidable redundancy as multiple people are dedicated to the same project. There is no distinction between supervisors and subordinates in the model. Instead, the firm must create a monitoring apparatus by hiring more than one worker for each task and having the workers serve as monitors for each other.

The analysis follows by dividing the principal’s problem into two stages: first, a job size $J$ is given and for any monitoring technology $Q^i$ there is shown to be another monitoring technology with constant monitoring probability that the firm prefers. The optimal wage given this superior monitoring technology is then derived. Theorem 1 therefore describes necessary properties of the optimal incentive organization. The optimal monitoring technology is then determined by maximizing over all technologies and job sizes satisfying Theorem 1.

### 1.5.1 Implementation

Lemma 1 simplifies the exposition by establishing two results. First, the firm never hires more than two workers for any task in equilibrium so it is without loss of generality to describe implementation plans as $Q^i = (q_1, ..., q_J)$ instead of $Q^i = (f_1, ..., f_J)$. Second, it is without loss to restrict attention to truthful recommendation strategies $\rho$ in which workers report their signals truthfully and follow the prior when shirking. Recommendation strategies are omitted in the discussion that follows and attention is focused on inducing effort.

**Lemma 1.** For any equilibrium of the firm’s problem, there exists another equilibrium providing the same expected payoffs to the firm and all agents in which (a) the firm adopts an implementation plan $Q^i = (q_1, ..., q_J)$ with $q_t = \sum_{n=2}^{\infty} f_t(n)$ and $f_t(n) = 0$ for all $n > 2$, and (b) workers use truthful recommendation strategies $\rho = \bar{\rho}$. 
The firm wants to induce effort at the lowest expected cost. One worker is insufficient and hiring two workers deterministically is costly. The firm may be able to improve its payoffs by stochastically employing a second worker.

The firm needs to align worker incentives so that high effort is an equilibrium. Consider generic implementation plan $Q^i = (q^i_1, ..., q^i_J)$ and strategy profile $(\sigma^i, \bar{\sigma}^{-i})$. The implementation plan and the strategy profile induce a distribution on the partition $\{R^i_{kj}\}_{k=0,...,J,j\leq k}$. Let $Pr(R^i_{kj}) = \sum_{r^i \in R^i_{kj}} G(r^i|\sigma^i, \bar{\sigma}^{-i}, Q^i)$ be the probability of a report in partition element $R^i_{kj}$ when worker $i$ is following strategy $\sigma^i$ and all other workers are exerting effort. The induced distribution on this partition is represented in Figure 1.2. Similarly define $\overline{Pr}(R^i_{kj}) = \sum_{r^i \in R^i_{kj}} G(r^i|\bar{\sigma}^i, \bar{\sigma}^{-i}, Q^i)$.

The firm’s payment scheme details lump-sum payments as a function of the report. Denote by $W_{kj} = W^i(r^i)$ for $r^i \in R^i_{kj}$ as the transfer to worker $i$ when a monitor is employed on $k \leq J$ tasks and output matches on $j \leq k$ of these tasks.

The general incentive constraint can now be written. Incentive compatibility requires

$$W_{00} \overline{Pr}(R^i_{00}) + W_{10} \overline{Pr}(R^i_{10}) + ... + W_{JJ} \overline{Pr}(R^i_{JJ}) - Jc$$

$$\geq W_{00} Pr(R^i_{00}) + W_{10} Pr(R^i_{10}) + ... + W_{JJ} Pr(R^i_{JJ}) - c \cdot \sum_{k=0}^{J} k \cdot \zeta(k|\sigma^i) \quad (1.1)$$
for all strategies $\sigma^i$.

Regardless of the transfer scheme, hired workers can choose to shirk and report to the firm as if they exerted effort and acquired signals. There is always positive probability that such a strategy garners payment and so satisfaction of incentive compatibility implies satisfaction of individual rationality. That IR holds strictly implies workers obtains rents, even when the firm structures incentives optimally.

The role of limited liability can be seen by considering Equation 1.1 without imposing LL. Unlimited liability allows the firm to threaten workers with arbitrarily severe punishments. To dissuade $\sigma^i$, the firm need only set $W_{kj}$ negative enough for $r^i \in \mathcal{R}^i_{kj}$ such that $\overline{P}(\mathcal{R}^i_{kj}) < Pr(\mathcal{R}^i_{kj})$.

With limited liability $W_{kj} \geq 0$ and such threats are unavailable. The firm must dissuade shirking by providing workers with rents in equilibrium. As in Shapiro and Stiglitz (1984), threats take the form of losing these rents if caught deviating.

The incentive constraint can be simplified by noting the firm should never offer a positive transfer when a worker fails to match on any of his monitored tasks. Setting $W_{kj} > 0$ for $k \neq j$ is both directly and indirectly costly: not only are workers paid when reports in $\mathcal{R}^i_{kj}$ are realized, but $W_{kj} > 0$ makes Equation 1.1 harder to satisfy since $\overline{P}(\mathcal{R}^i_{kj}) \leq Pr(\mathcal{R}^i_{kj})$ whenever $k \neq j$.

**Lemma 2.** The optimal incentive contract sets $W^i(r^i) = 0$ $\forall r^i \in \mathcal{R}^i_{kj}$ whenever $k \neq j$.

Lemma 2 implies workers are required to produce matching output whenever multiple workers are assigned to a task. Each worker’s transfer is therefore a function of the reports of all workers. Even though workers produce output individually, the payment scheme treats them as if they are a team and punishes everyone when any one worker shirks.

From $\sum_{k=0}^{J} \overline{P}(\mathcal{R}^i_{kk}) = \overline{P}(\mathcal{R}^i) = 1$ and the no induced breach condition (B), Equation 1.1 can be rewritten in terms of the required transfer $W$:

$$W \geq \frac{c(J - \sum_{k=0}^{J} k \cdot \zeta(k|\sigma^i))}{1 - Pr(\mathcal{R}^i)}, \quad \forall \sigma^i \in \Sigma^i. \quad (1.2)$$
The incentive constraint is further simplified by viewing the firm’s zero
tolerance of mismatches as the firm requiring acceptable output on every assigned
task. Acceptable is then taken to include the vacuous case in which only one
worker was assigned the task.

Worker incentives are governed by the probability a worker believes his
report will be compared to the report of another agent. So while implementation
plan $Q^i = (q_1, ..., q_J)$ specifies hiring two workers for task $t$ with probability $q_t$,
a worker assigned $t$ believes with probability $1 - \frac{2q_t}{1+q_t}$ he is the only agent employed.
With probability $\frac{2q_t}{1+q_t}$ the worker is one of two agents assigned $t$.

Generic strategy $\sigma^i = (p^i_1, ..., p^i_J)$ specifies the probability $i$ exerts effort on
each task $t \in J^i$. The worker produces acceptable output on $t$ with probability
$\frac{1}{1+q_t} + q_t(1 - p^i_t + \pi)$. Since the firm requires acceptable output on all assigned tasks,
the incentive compatible wage satisfies

$$ W \geq \frac{c(J - \sum_{k=0}^{J} k \cdot \zeta(k|\sigma^i))}{1 - \left[\frac{1 - q_t + 2q_t(1 - p^i_t + \pi)}{1 + q_t}\right]^J}, \quad \forall \sigma^i \in \Sigma^i. \quad (1.3) $$

The payment required to induce effort can now be derived. Theorem 1
establishes several properties of an optimal contract. It shows that implementation
plans employing different monitoring probabilities make inefficient use of the firm’s
monitoring ability. It also shows that it is sufficient for the firm to dissuade $\sigma^i$.

**Theorem 1.** The optimal contract will take the form of an implementation plan
$Q^i = (q, ..., q)$ that specifies hiring a second agent with probability $q$ for each task
and a transfer

$$ W^i(r^i) = \begin{cases} J \cdot w^*(q, \pi, c, J) & \forall r^i \in \mathcal{R}^i \\ 0 & \forall r^i \not\in \mathcal{R}^i, \end{cases} $$

where the equivalent per-task wage is $w^*(q, \pi, c, J) = \frac{c}{1 - \left[\frac{1 - q + 2q\pi}{1 + q}\right]^J}$.

**Corollary 1** (Comparative Statics). $w^*(q, \pi, c, J)$ is decreasing in $q$ and $J$ and
increasing in $\pi$ and $c$.

Theorem 1 establishes a schedule of implementation plans and transfer
schemes capable of inducing equilibrium behavior. The $(Q^i, W^i)$ combinations
trade-off higher monitoring probabilities against lower wages.
As in the contractible effort benchmark, the firm retains the possibility of guessing the state blindly instead of hiring workers. Theorem 1 implies there exist effort costs for which the firm would undertake an employment contract were effort contractible but not when effort is non-contractible.

**Corollary 2.** For effort costs \( c \in ((1 - \pi)\left(1 - \frac{1 - \pi + 2q^*}{1 + q}ight)^J, 1 - \pi] \) the firm will not employ any workers despite it being efficient to do so were shirking not a concern.

Theorem 1 follows from the log-convexity of the probability of producing matching output when shirking, which implies the firm optimally chooses an implementation plan specifying a common monitoring probability for all tasks. This is because an implementation plan with common monitoring probability \( \bar{q} = \frac{1}{J} \sum_{k=1}^{J} q_k \) specifies a lower wage paid to the same expected number of workers as \( Q^* = (q_1, ..., q_J) \). Alternately, the firm can adopt a common monitoring probability \( \hat{q} \) where \( \left[\frac{1-\hat{q}+2\pi \hat{q}}{1+\hat{q}}\right]^J = \left[\frac{1-q_1+2\pi q_1}{1+q_1}\right] \times \cdots \times \left[\frac{1-q_J+2\pi q_J}{1+q_J}\right] \). This implementation plan pays the same wage to fewer expected workers.

Lemma 3 establishes that given any implementation plan \( Q^i = (q_1, ..., q_J) \), there exists another implementation plan \( \hat{Q}^i = (q, ..., q) \) that offers the same transfer scheme but hires fewer workers. Optimality is then a case of trading off the stronger disincentives (and lower wages) of more intensive monitoring against the greater number of expected workers.

### 1.5.2 Optimization

Theorem 1 begins by considering any implementation plan \( Q^i \) for a \( J \)-sized job. It shows that there exists another implementation plan specifying a constant monitoring probability across tasks that the firm prefers to \( Q^i \). Theorem 1 then derives the minimum transfer required to induce effort. Thus, Theorem 1 describes several properties the optimal organization of incentives must satisfy.

The monitoring probability \( q \) is now taken to be endogenously chosen by the firm. In addition to being a key component of the firm’s transfer scheme, the implementation plan also affects how many workers the firm expects to hire. The size of the job, \( J \), continues to be exogenously given for the time being.
The firm learns the state for each task in equilibrium. Since any \((Q^i, W^i)\) combination satisfying Theorem 1 provides the same information about the unknown states to the firm, optimality is determined by minimizing the firm’s expected wage bill.

With monitoring probability \(q\), \(1 + q\) workers are hired in expectation for each task and each is paid the equivalent of \(w^*(q, \pi, c, J)\) per task.\(^{21}\) The firm’s optimal monitoring probability then solves

\[
q^*(\pi, c, J) = \arg \max_q (1 + q) \cdot \frac{c}{1 - \left[\frac{1 - q + 2q\pi}{1 + q}\right]^J}. \quad (1.4)
\]

The optimal monitoring probability is implicitly determined to balance the cost of greater monitoring – more workers are hired in expectation – against the benefit of lower wages. While \(q^*(\pi, c, 1) = 1\), for \(J > 1\) \(q^*(\pi, c, J) < 1\) and the firm does not hire a monitor deterministically.

**Theorem 2** (Stochastic Monitoring). \(q^*(\pi, c, J) < 1\) if and only if \(J > 1\).

\(^{21}\)Implicit in writing the per-task wage bill as \((1 + q) \cdot w\) is that when two workers are hired for a task, both are paid the same wage. This is shown as a property of the optimum in Section 1.6.1.
The optimal monitoring probability responds intuitively to changes in the environment. (See Figure 1.4.) As the prior becomes more pronounced, shirking becomes more attractive and the firm must monitoring employees more frequently. The optimal monitoring probability does not vary with changes in the cost of effort; instead, the firm adjusts the transfer.

From the firm’s perspective, monitoring and bundling are substitutes and the firm trades the intensive margin for the extensive margin: each task is monitored less intensively in exchange for assigning more tasks.

**Theorem 3.** \( q^*(\pi, c, J) \to 0 \) as \( J \to \infty \).

Theorem 1 shows that for any job size the firm can approximate the action profile of the contractible effort benchmark. Theorem 2 shows that this is achieved at great cost: the firm can obtain higher expected payoffs by monitoring more often and not approximating the benchmark action profile. While Rahman (2012) considers when approximating the contractible effort benchmark action profile is feasible, Corollary 3 shows that this is optimal only in the limit.

**Corollary 3.** For any \((\pi, c)\), given \( \varepsilon > 0 \), \( \exists J_\varepsilon < \infty \) such that \( q^*(\pi, c, J) > \varepsilon \) for all \( J < J_\varepsilon \).
Remark 2 (Virtual Monitoring Only in Limit). When monitoring is chosen endogenously, virtual monitoring is adopted only in the limit as $J \to \infty$.

For fixed $J$, Corollary 1 shows $w^*(q, \pi, c, J)$ is increasing in $q$. One may suspect $w^*(q^*(\pi, c, J), \pi, c, J)$ to be increasing in $J$; after all, if the firm is rarely employing a monitor, the worker has a high probability of producing acceptable output and so requires a higher wage to eschew shirking. This fails to account for the direct effect on $w^*$ of increasing $J$, however: shirking is less attractive for large $J$ since the worker needs to produce acceptable output for more tasks. Along the optimal path as $J$ increases, the direct effect of increasing $J$ on $w^*(q^*(\pi, c, J), \pi, c, J)$ dominates the indirect effect of decreasing $q^*$. Not only are fewer workers being hired for each task, but each worker is paid a lower per-task wage. Taken together, the efficiency loss to the firm relative to the contractible effort benchmark vanishes asymptotically as $J \to \infty$. (See Figure 1.5.)

Result 1 (Asymptotic Efficiency). As $J \to \infty$,

$$(1 + q^*(\pi, c, J)) \cdot w^*(q^*(\pi, c, J), \pi, c, J) \to c.$$
working and shirking. When negative payments are possible, the wedge can be monetary punishments for poor (or unfortunate) outcomes. With limited liability, sticks are unavailable and a carrot must be used. The wedge becomes a reward that is withheld for poor performance.

Bundling presents the firm another option when dismissal and punishment are unavailable. Instead of a costly bonus on top of the worker’s earnings, the reward is simply the receipt of the wages the worker accumulated throughout the job. By bundling multiple tasks together into one job, the firm is able to withhold the worker’s earnings until the entire job is successfully completed. The disincentive to shirking would be dulled if the firm instead paid on a piece-rate basis.

Bundling plays a larger role here than simply scaling up the firm’s available rewards and punishments; bundling multiple tasks together reduces the per-task cost of monitoring. The firm’s problem is non-linear in the number of tasks it assigns to a worker because of the role played by bundling in the optimal monitoring technology.

Remark 3. The firm optimally bundles as many tasks together as possible: \( J^* = \overline{J} \).

1.6 Discussion

Bundling takes the form of assigning a worker multiple tasks and conditioning payment for any task on satisfactory performance on all tasks. An alternative interpretation of bundling is provided by considering multidimensional tasks.

Suppose an agent is assigned a chore consisting of multiple components. For example, a worker must complete a tax return with additional schedules for each source of non-wage income. The tasks of Section 1.4 are now individual tax forms and a job is the entire tax return. The main result of Section 1.5 is that the Internal Revenue Service should monitor individual tax forms and impose the harshest possible penalty upon uncovering any irregularities. There is evidence the IRS follows such a strategy. Tax returns claiming unusually large charitable deductions, Schedule C self-employment income or business expenses invite an
audit of the entire return (Barrett, 2011; CBSNews, 2010).

An entertaining example of the theory of Section 1.5 is provided by the rock band Van Halen. Like many musical acts, the band’s performance contract with event venues is a long, complicated document specifying hundreds of individual items. Within the 53-page rider is an obscure provision often taken as prima facie evidence of rock excess: a bowl of M&M’s is to be provided with all brown candies removed. As the band’s lead signer explained in his autobiography, the unusual request performed a monitoring function:

Van Halen was the first band to take huge productions into tertiary, third-level markets. We’d pull up with nine eighteen-wheeler trucks, full of gear, where the standard was three trucks, max. And there were many, many technical errors – whether it was the girders couldn’t support the weight, or the flooring would sink in, or the doors weren’t big enough to move the gear through.

The contract rider read like a version of the Chinese Yellow Pages because there was so much equipment, and so many human beings to make it function. So just as a little test, in the technical aspect of the rider, it would say “Article 148: There will be fifteen amperage voltage sockets at twenty-foot spaces, evenly, providing nineteen amperes ...” This kind of thing. And article number 126, in the middle of nowhere, was: “There will be no brown M&M’s in the backstage area, upon pain of forfeiture of the show, with full compensation.”

So, when I would walk backstage, if I saw a brown M&M in that bowl ... well, line-check the entire production (Roth, 1997, pp. 97-98).

Such an arrangement is not unusual. Managers hiring subordinates are frequently faced with the problem of judging a worker’s ability on the basis of limited information. As anyone with experience in such hiring decisions can attest, candidates are often removed from consideration for having spelling or grammar mistakes on their resumes. Poor performance on one component is taken as evidence of low quality on other, unrelated dimensions. For the same reason an investing website lists “Math Errors” and “Failure to Sign the Return” as the numbers 2 and 3 “red flags” prompting a tax audit (Investopedia, 2010).

Allen (2011, p. 43) suggests an inability to monitor individual tasks led to bundling before the industrial revolution: “Paying workers for specific tasks meant

\footnote{The number 1 “red flag” is “Overestimating Donated Amounts.”}
both parties needed to be able to separate these tasks from others. The inability to do this during the pre-modern era meant that all tasks were essentially bundled together, and as a result, the laborer became a servant under the general and universal supervision of the master.” The argument put forth in this paper is that bundling becomes even more attractive when individual measurement is possible since it permits the principal to strengthen the incentives provided to agents.

The incentive structure of Section 1.5 is similar to that employed by the British Admiralty. The navy struggled with cowardice and needed to arrange adequate rewards supported by an effective monitoring apparatus. Opportunities to amass wealth were made hostage to acceptable performance. “Any slipup discovered by the Admiralty meant [...] that half pay was given to those captains who made slight mistakes” (Allen, 2011, p. 123).

Monitoring was performed by the battle line and multiple record keeping. With the creation of Fighting Instructions in the 1600s, the navy codified a set of rules to enable easier identification of shirking at sea. The instructions specified the formation of a “line of battle” with all ships keeping in line with the chief. The tactical disadvantage of the battle line was outweighed by the ease of monitoring it offered: identifying when a captain was failing to engage in battle was simply a matter of spotting who was “out of line.”

With the creation of the battle line, the Admiralty invested in a costly monitoring technology to generate desirable behavior in battle. To induce desirable behavior outside of battle, the Admiralty required commissioned lieutenants and noncommissioned masters on board each ship to keep detailed journals of the captain’s performance. The journals were turned over to the Admiralty upon reaching shore and different accounts of the officers’ performance were compared.

Costly duplication – like requiring multiple overlapping accounts of a captain’s actions – is a feature common throughout the experimentation literature. There, additional signals are acquired for learning purposes. Each signal increases the firm’s knowledge about an unknown state and the firm acquires signals until the marginal benefit from better information is outweighed by the marginal cost of purchasing another signal.
There is no learning justification for hiring multiple workers in the firm’s problem in Section 1.4. Hiring multiple agents is for incentive reasons only. Any departure from the contractible effort benchmark is the cost of generating adequate incentives to induce effort. The theory thus provides a complementary rationale for costly duplication to that offered in the experimentation literature. In the example of the British Admiralty, both motivations were likely in play, though the historical record suggests monitoring was the larger concern.

### 1.6.1 Hierarchy

Alchian and Demsetz (1972) suggest the firm best aligns incentive by having specialized monitors. It may seem reasonable to suspect such a result here: a worker believing with certainty that two workers are assigned to a task knows his output will be verified and, thus, requires a lower wage in equilibrium. The model of Section 1.4 can be modified to examine this question.

It is assumed in Section 1.4 that workers know only the firm’s implementation plan for each assigned task. That is, workers know how many workers the firm expects to hire for each task, but a hired worker does not know whether he is the first or second agent the firm hires for the task. Each hired worker believes a task is assigned to two agents with probability \( \frac{2q}{1+q} \). This need not be the case. Instead, the firm can inform each worker it hires whether it is the first agent assigned to the task or the second agent. The firm creates specialized monitors by doing so.

Given any monitoring probability \( q \), \( 1 + q \) workers are hired in expectation on each task. The probability that a given worker is the first agent assigned to a task is \( \frac{1}{1+q} \); likewise, with probability \( \frac{q}{1+q} \), a randomly chosen worker is the second agent assigned to the task.

A worker informed that he is the first agent hired knows his output is monitored with probability \( q \). A worker told that he is the second agent assigned to a task knows that his output is necessarily checked by another agent. Thus, before informing a worker about whether he is the first or second agent hired, the worker’s ex ante belief about the probability his output is monitored is given by \( \frac{1}{1+q} q + \frac{q}{1+q} 1 = \frac{2q}{1+q} \). This implies that dedicating certain workers as specialized
monitors results in a mean-preserving spread of beliefs when \( q \) is known by all workers. Since Theorem 1 implies that the wage required to induce effort is convex in a worker’s belief that a task is assigned to two agents, the firm’s expected wage bill is minimized by treating workers symmetrically.

**Remark 4 (Hierarchy).** *The optimal organization of the firm does not include a hierarchy of monitors and subordinates.*

Contrary to the suggestion of Alchian and Demsetz (1972), creating a vertical hierarchy within the firm is not optimal.\(^{23}\) While it is true that subordinates – those who believe their work is often verified by a supervisor – receive lower wages within a hierarchical organization, supervisors believe their output is often unverified by the firm and therefore command a higher wage. In the context of the model of Section 1.4, supervisors are akin to the first worker employed on a task: only sometimes is their output compared to that of another agent. Subordinates are like the second worker assigned to a task. By definition first workers are more numerous than second workers.

Alchian and Demsetz (1972, p. 782) famously ask “Who will monitor the monitor?” The answer provided here is “the monitored”: the monitor of the monitor is the monitored. The dynamic of workers simultaneously monitoring each other is a key feature of Section 1.5.

### 1.6.2 Virtual monitoring

An incentive contract exhibits virtual monitoring if for any \( \varepsilon > 0 \) the firm induces a high-effort equilibrium while hiring a monitor with probability less than \( \varepsilon \). In the contractible effort benchmark (Remark 1), each worker exerts effort and the firm never employs a monitor. Interest in virtual monitoring arrangements stems from the action profile under virtual monitoring being arbitrarily close to the action profile under the contractible effort benchmark.

Rahman (2012) examines when virtual monitoring is feasible for the firm. From Theorem 1, virtual monitoring is always feasible but it requires unboundedly

\(^{23}\)Williamson (1967), Mirrlees (1976) and Calvo and Wellisz (1978) also discuss firm size.
large wage payments. While the firm can approximate the first-best action profile, it cannot approximate its payoffs in the first-best equilibrium. Since the required wage diverges to infinity as the monitoring probability shrinks, virtual monitoring is not a feature of the optimal incentive organization.

**Remark 5 (Virtual Monitoring).** There exists a transfer scheme capable of inducing equilibrium for any monitoring probability \( q > 0 \), but \( w^*(q, \pi, c, J) \to \infty \) as \( q \to 0 \).

### 1.6.3 Mediated contracts and the “Gold Standard”

In a mediated contract, a third-party, or mediator, provides agent-specific action recommendations to introduce correlation among agents’ action choices (Rahman and Obara, 2010). Rahman (2012) studies a multilateral principal-agent problem with costly effort where payments depend on the recommendation-contingent actions of the workers.

By privately recommending actions to workers, the firm asks agents a question for which it already knows the answer. In Rahman (2012), the question posed to a monitor is whether a subordinate exerted effort; the firm randomly instructs the subordinate to shirk and rewards the monitor only for correctly reporting the subordinate’s action.

Such mediated contracts are not literally available in our setting. However, the dynamic of posing workers with questions for which the firm already knows the answer can be recreated just the same. In fact, the use of questions with known answers is already commonplace in crowdsourcing, where it’s known as the “gold standard.”

The firm simply seeds each worker’s job with tasks for which the state has already been learned. Workers are paid if they perform satisfactorily on this subset of tasks.

Regardless of their availability, these mediated-style contracts are not optimal here. To see this, suppose that the firm has costlessly acquired a collection of tasks for which the state is privately known to the firm. In actuality, the firm

---

24 See www.crowdflower.com, for example.
would have had to hire workers to discover the state, thereby making this knowledge costly to acquire in the first place, but the suboptimality of an incentive organization built around known tasks can be shown even in this more demanding setting.

Workers are assigned $J$ tasks and the firm optimizes over how many known tasks to include within this set; let $n$ be the optimal number. Worker reports on their $J$-sized jobs are judged on the basis of their performance on the subset of known tasks. Payment is provided if and only if the worker performs satisfactorily on all $n$ tasks.

Assigned $J$ tasks, a worker must be compensated for exerting effort on all $J$ tasks. (Workers must also receive rents to dissuade them from shirking, as above.) Call this compensation $\tilde{w}$. The firm is paying each worker $\tilde{w}$ in equilibrium and is obtaining knowledge of the state on $J - n$ new tasks. So with $2J$ tasks assigned to two workers, the cost is $2\tilde{w}$ while the benefit is $2(J - n)$.

Alternatively, the firm could tell each worker that he will be monitored on $n$ of the $J$ tasks the firm assigns him. The disincentive to shirking is just the same, so two workers will cost the firm just as much as before, $2\tilde{w}$. But now the firm is learning the state for $J + (J - n) > 2(J - n)$ new tasks. The same wage bill is now spread over a greater number of valuable tasks: the effective per-task wage is reduced by using an incentive organization based around the results in Section 1.5. Mediated-style contracts are inefficient because they monitor workers independently; instead, the optimal incentive organization in Section 1.5 monitors workers simultaneously. This advantage is reduced as $J$ grows large.

**Remark 6.** Mediated-style contracts, like the “Gold Standard,” are dominated by the incentive organization of Section 1.5 that monitors multiple workers simultaneously.

### 1.6.4 Alternative models and robustness

The model of Section 1.4 considers a firm optimally inducing all workers to exert effort on each task. It may not be optimal for the firm to induce high-effort; at the most basic level, the firm participation constraint must still be satisfied.
For some constellations of parameters, high-effort will produce the best possible
payoff for the firm while for other constellations, the firm may want to induce less
effort. One way to view the results from Section 1.5 is to consider them as deriving
the firm’s optimal organization of incentives when the parameters of the model are
such that high-effort leads to the highest possible payoff for the firm.

Section 1.4 presents a benchmark model in which effort perfectly reveals a
task’s unknown state and all workers share a binary cost of effort. This stark setting
best isolates the incentive issues facing the firm. The firm hires multiple workers
purely for incentive reasons; there is no learning justification for the duplication of
tasks.

Alternatively, Section 1.4 could have presented a model with imperfect but
sufficiently precise signals so that in the contractible effort benchmark the firm
wants to hire exactly one worker. The interpretation of the results in Section
1.5 remains largely unchanged: incentive concerns push the firm away from the
contractible effort benchmark and towards hiring additional workers. The firm
now receives an additional learning benefit as well.

The main results of Section 1.5 derive the firm’s optimal incentive organiza-
tion. The firm induces workers to undertake costly effort by periodically monitoring
their output through the employment of additional workers. Incentive contracts
optimally require workers to complete as many tasks as permissible.

These results obtain outside of the benchmark setting of Section 1.4. No
matter if effort costs are a continuously increasing function, workers exhibit het-
erogenous costs of effort, or exerting effort imperfectly reveals the state, the effect
is that mismatches occur in equilibrium and the firm may want to introduce some
measure of forgiveness into its incentive structure.

Consider the case in which effort is not perfectly precise. In the benchmark
model of Section 1.4, workers necessarily produce satisfactory output in equilibrium
and, therefore, they obtain payment with certainty when working. This is not the
case when effort imperfectly reveals the state. Even with all workers exerting effort
on all tasks, workers sometimes report different states for the same task and the
incentive contracts from Section 1.5 are unable to induce effort. By displaying
leniency upon observing mismatches the firm is able to again convince workers to exert effort.

There are three levers the firm can employ to introduce leniency. It can offer a positive wage payment despite a worker occasionally failing to produce matching output. In addition to setting an interior match rate, the firm can assign fewer tasks to each worker or monitor each task less intensively. The effect of each lever is the same: a worker is able to produce matching output on fewer tasks and still obtain payment. This does not imply the firm is indifferent among the levers though. Indeed, the optimal incentive organization always bundles as many tasks together as possible.

Result 2. Let $\gamma \in (\pi, 1]$ be the precision of the signal obtained by exerting effort: $\gamma = \Pr(s_t = \omega_t | \omega_t)$. The optimal incentive organization bundles as many tasks together as permissible.

Result 2 establishes that the firm does not introduce leniency by endogenously constraining the size of jobs. So, how does the firm structure incentives? Compared to the benchmark model of Section 1.4, the firm monitors less frequently when mistakes happen in equilibrium. Instead of directly forgiving mismatches when they occur, forgiveness for mismatches is introduced by the firm detecting mismatches less often. Interestingly, as signals become less precise, the optimal monitoring probability falls. When signals are imprecise mismatches occur more frequently, making the worker’s constraints harder to satisfy. The firm relaxes these constraints by monitoring less frequently. Instead of forgiving more mismatches or assigning fewer tasks, the firm elects to catch mismatches less often.

1.6.5 Other extensions

Generalized payoffs

The firm’s payoff is 1 for matching the state and 0 for failing to match the state in Section 1.4. Allowing the firm’s payoffs for matching (and failing to match) the state to depend on the state only affects the firm’s participation constraint (FP). Let $\alpha_\omega \geq 0$ be the firm’s payoff from matching the state when
the state is \( \omega \) and \( \beta_\omega \leq 0 \) be the firm’s payoff from failing to match state \( \omega \). So \( \alpha_1 \) (\( \beta_1 \)) corresponds to permitting (prohibiting) harmless content to be shared and \( \alpha_0 \) (\( \beta_0 \)) represents removing (failing to remove) an objectionable item.

Firm optimization remains as before. The sole change is to when the firm prefers the proposed contract to guessing the state blindly. With payoffs \( \alpha_\omega \) and \( \beta_\omega \), the firm’s expected payoff from selecting \( A_t = 0 \) (\( A_t = 1 \)) without hiring any workers is \( (1 - \pi)\alpha_0 + \pi \beta_1 \). The proposed contract must exceed both of these values, so the firm’s participation constraint becomes

\[
\pi \alpha_1 + (1 - \pi) \alpha_0 - (1 + q) \left( 1 - \frac{c}{1 - \left[ \frac{1 - q + 2q\pi}{1 + q} \right]^J} \right) \geq \max\{ (1 - \pi)\alpha_0 + \pi \beta_1; \pi \alpha_1 + (1 - \pi) \beta_0 \},
\]

which is equivalent to

\[
c \leq \min\{ (\pi(\alpha_1 - \beta_1); (1 - \pi)(\alpha_0 - \beta_0)) \left( 1 - \left[ \frac{1 - q + 2q\pi}{1 + q} \right]^J \right) \}. \]

Corollary 2 adjusts accordingly.

**Generalized tasks**

Section 1.4 considers a binary state space \( \Omega = \{0, 1\} \) to simplify the exposition. It is often without loss of generality to presume tasks can be subdivided into binary choices: instead of asking workers to classify a product as a “white, long-sleeve blouse,” the firm can instead structure tasks so that workers separately categorize whether the item is “white,” “long-sleeve,” and a “blouse.”

Even without the possibility of such a subdivision process, the results presented in Section 1.5 continue to apply with more general state spaces. Let \( \Omega = \{\omega^1, \omega^2, \ldots\} \) and redefine \( \pi = \max_k Pr(\omega = \omega^k) \). That is, \( \pi \) is the probability of the most likely state. From the firm’s perspective, this is the only state that matters for inducing effort, and so the results in Section 1.5 carry-through unchanged.

In Section 1.4 the firm has access to a stream of identical tasks. In particular, \( \pi = Pr(\omega_t = 1) \forall t \). Suppose instead that for each task \( t \) the prior belief is \( \pi_t = Pr(\omega_t = 1) \) and that the firm hires an additional worker for \( t \) with probability \( q_t \). A worker’s probability of matching on task \( t \) when shirking is \( \frac{1 - q + 2q\pi}{1 + q_t} \).
The incentive compatible wage scheme is determined by modifying Theorem 1. Consider a monitoring technology \((q_1, \ldots, q_J)\) and

\[
\hat{t} \in \arg \max_{t=1,\ldots,J} \frac{1 - q_t + 2\pi_t q_t}{1 + q_t},
\]

so that the temptation to shirk is greatest on task \(\hat{t}\). If a worker shirks on \(\hat{t}\), he shirks on all tasks, so it is sufficient to discourage the worker’s strategy calling for no effort.

Now suppose \(\frac{1-q_t+2\pi_t q_t}{1+q_t} > \frac{1-q_t+2\pi_\hat{t} q_{\hat{t}}}{1+q_t}\) for some task \(t\). This cannot be optimal since the firm could lower the monitoring probability on task \(t\) – thereby increasing \(\frac{1-q_t+2\pi_t q_t}{1+q_t}\) – without affecting incentives. So \(\frac{1-q_t+2\pi_t q_t}{1+q_t} = \frac{1-q_{\hat{t}}+2\pi_{\hat{t}} q_{\hat{t}}}{1+q_{\hat{t}}}\) \(\forall t\). This defines the monitoring probability on task \(t\) in terms of parameters and the monitoring probability on any other task. For \(\pi_j > \pi_k\), \(q_j > q_k\): the firm monitors more intensively when the prior is more pronounced, which parallels the finding in Section 1.5.

The transfer, \(W\), to the worker for acceptable output must satisfy \((1 - \left(\frac{1-q_1+2\pi_1 q_1}{1+q_1}\right)^J) \cdot W \geq Jc\). The optimal incentive organization is determined by the system

\[
\min_{(q_1, \ldots, q_J)} \left(1 + \frac{\sum_{k=1}^J q_k}{J}\right), \frac{c}{1 - \left[\frac{1-q_1+2\pi_1 q_1}{1+q_1}\right]^J}
\]

subject to

\[q_t = \left(\frac{1 - \pi_t}{1 - \pi_1} \frac{1 + q_1}{q_1} - 1\right)^{-1}, \quad \text{for } t = 1, \ldots, J\]

\[q_t \in (0, 1], \quad \text{for } t = 1, \ldots, J.\]

The objective function states that the firm is minimizing the expected per-task wage bill. The constraints allow the firm to treat this multivariate optimization problem as a univariate problem.

The solution to this problem calls for monitoring the task with the most pronounced prior more intensively than in the optimal incentive organization from Section 1.5 with \(\pi = \max_t \pi_t\). In Section 1.5, monitoring one task more intensively implied monitoring all tasks more intensively. Here, monitoring the task more
intensively allows the firm to decrease the intensity with which it monitors other tasks.

1.7 Conclusion

New information technology permits firms and workers to interact through spot labor markets. Compared to conventional labor markets, spot markets offer significant advantages for a firm. A flexible and scalable workforce is available to start work immediately and no preexisting relationship with a worker is presumed nor is the promise of a continuing relationship required.

The minimal interaction between the firm and its employees raises new challenges. The firm must provide adequate supervision to ensure workers are acting faithfully on its behalf. Workers are compensated for their effort, but the exertion of effort is costly and unobservable. Furthermore, the quality of a worker’s output cannot be verified directly.

With traditional reputation mechanisms inapplicable and the threat of large penalties for poor performance unavailable, the firm creates incentives for effort by periodically hiring additional workers to duplicate some of the tasks it has already assigned. Wages are then made contingent upon satisfactory performance on all tasks.

The firm’s monitoring technology is endogenously determined to balance the costs and benefits of supervision. Employing monitors more frequently keeps wages down but leads to greater duplication of assignments. In the optimal organization of incentives, the firm bundles multiple tasks together for each worker. Monitoring and bundling are strategic substitutes: as the firm assigns more tasks to each worker, it monitors each task less frequently.

I especially thank J. Aislinn Bohren for great collaboration as coauthor on this chapter. The chapter benefits greatly from discussion with Nageeb Ali, Cécile Aubert, Gordon Dahl, Roger Gordon, Kevin Lingerfelt, Mark Machina, David Miller, David Rahman, Branislav Slantchev, Joel Sobel, Ross Starr and Joel Watson.
The preceding chapter was coauthored with J. Aislinn Bohren.
Chapter 2

Lobbying for Influence with Strategic Lawmakers

Abstract

In *Citizens United v. Federal Election Commission* (2010), the Supreme Court ruled that corporations can legally use their general treasury funds for political donations. The decision overturned decades of campaign finance law and generated significant fear that rich corporations would be able to use their wealth to effectively buy preferential treatment from politicians. We study a discriminatory all-pay contest model with a cap in which lobbyists’ valuations are endogenously determined by a lawmaker’s policy proposal. We show that as the contribution limit is relaxed, (i) total contributions accruing to the politician always increase, which is in contrast to existing results, (ii) the policy proposed by the lawmaker may become more centrist, and (iii) the corporation’s likelihood of buying the lawmaker’s vote does not increase even when the politician is right-leaning.

2.1 Introduction

On January 21, 2010, the U.S. Supreme Court overturned century-old restrictions on corporate political spending. Prior to the ruling in *Citizens United*
v. Federal Election Commission, corporations and labor unions had to use political action committees funded by voluntary employee contributions for politicking. Contributions were strictly limited – $5,000 per individual – as were expenditures. Though they could defray administrative costs, corporations and unions could not contribute to the PAC directly.

Citizens United removed these restrictions, legalizing the deployment of general treasury funds for political ends. The decision releases vast new sums into the political theatre: during the 2007-2008 election cycle federal PACs spent $1.2 billion, only slightly below the $1.5 billion spent by FEC-registered political parties. During the same period the Fortune 100 companies had revenues of $13.1 trillion and profits of $605 billion.¹

Many commentators were incensed by the ruling. The decision was called “reckless,” “disastrous,” and a “strike[] at the heart of the democracy” (Dionne, 2010; Editorial, 2010a,c). Editorial boards warned the faith citizens harbor in their democracy is at risk (Editorial, 2010d,b). The Court’s verdict, one wrote, “pave[s] the way for corporations to use their vast treasuries to overwhelm elections and intimidate elected officials into doing their bidding” (Editorial, 2010c). The dissenting opinion warned of undermining “the integrity, competitiveness, and democratic responsiveness of the electoral process” 558 U.S. 83 (2010) (opinion of J. Stevens).

Economic theory suggests that the ill-effects predicted by the decision’s critics may be overstated. At its heart, the lobbying process looks much like a discriminatory all-pay auction. Competing interests attempt to persuade the lawmaker to award a prize to their side. Both sides lose their investment while only the winning side claims the prize. Pre-Citizens United, the interests were constrained in how much they could spend persuading the lawmaker. The Court’s ruling relaxed these restraints. The ruling did not, however, remove all financial constraints faced by the special interest groups.

We present a formal model to study the effects of financial constraints on lobbying expenditures. The lawmaker assumes a central role in our model. In his

---

role as policymaker, he designs the prize over which the lobbyists are competing. As such, the lobbyists’ valuations are endogenous. In his role as fundraiser, he determines the mechanism by which the prize is awarded. In particular, we allow the lawmaker to discriminate against certain bidders so that the highest bidder does not always win the prize.

Formally, the lawmaker first proposes a prize to allocate. The lobbyists’ valuations of the prize depend on what the lawmaker proposes. Next, the lawmaker stipulates the terms of the competition between the lobbyists. This can be as simple as awarding the prize to the lobbyist offering the larger contribution, but this need not be the case. The lawmaker confers an advantage upon one lobbyist so that she is capable of winning the prize without offering the larger contribution.

Previous studies of financially-constrained lobbying settings direct their attention away from lawmakers (Che and Gale, 1998; Pastine and Pastine, 2009). The lawmaker selling rights to shape policy plays no strategic role within the models. He is the purveyor of a good – his vote in the most extreme settings – but he has no control over how much the policy is worth or how the rights are awarded. These studies examine how lobbying expenditures change as the maximum contributable amount changes. In other words, the behavior of the lawmaker is held fixed while the environment in which he operates varies. In our model, when the environment changes – for example, when the Supreme Court alters the restrictions governing political donations – the politician’s behavior changes too. He changes the prize and the terms of its allocation. Throughout, we assume, the politician behaves strategically.

The intuition for why a lawmaker may not want to award the prize to the highest bidder is similar to the intuition for why a seller may want to eliminate potential bidders from competition and why a cap on bids may actually raise total expected bids in all-pay auction settings: when one bidder has a high valuation relative to his competition, he is likely to win the item. The other bidders know this, and in equilibrium, they do not bid aggressively. This in turn enables the high value bidder to bid less aggressively as well. Eliminating the high value bidder from the auction or constraining the maximum amount he can spend levels
the playing field by reducing his advantage, thereby spurring higher bids from the other bidders. Discriminating against the high value bidder works the same way in our model, but it is more effective at raising the seller’s revenue. By requiring strong bidders to significantly outbid their competitors, the seller is able to sap some of the surplus the high value bidder would have accrued.

Another contribution of this paper is to redirect attention from total lobbying expenditures as the measure of interest following a change in the regulatory environment. Caps were imposed upon the amounts and means by which donors could contribute to politicians not because of the deleterious effects of money itself, but, rather, because of the possibility of money distorting policy. To the extent that previous studies have focused exclusively on total contributions without regard to the effects on enacted policy, these studies have missed the point.\(^2\) Money is instrumental in our approach. The politician’s behavior is determined by his objective of maximizing the donations he receives. His behavior at the optimum determines policy, which is the object of interest here.

Contrary to the concerns voiced above, the effect on political competition of *Citizens United* is far from clear. First, while *Citizens United* relaxes restraints on contribution limits, it does not eliminate them. Even if a corporation can now spend many times more money than before, if the competition can still only spend relatively small amounts, the corporation’s deep pockets are not of much use. In fact, they may become a liability.\(^3\) Second, one of the most effective ways for the optimizing lawmaker to maximize contributions is to ensure healthy competition. It is often not in the politician’s interest to allow one special interest group to gain too large of a valuation advantage. When both special interest groups value the prize equally, the lobbying effort of each will be intense. For example, for some increases in the effective contribution limit, policy remains centrist and there is no effect on the policy from the relaxation of regulatory contribution limits. Lobbying

---

\(^2\)Of course, when policy is held fixed despite a change in environment, this criticism cannot be remedied.

\(^3\)This concern is not academic: the Court noted, “some corporations have affirmatively urged Congress to place limits on their electioneering communications. These corporations fear that officeholders will shake them down ... [and] that they will have to spend increasing sums on elections. ... [Removing the limits] can impose a kind of implicit tax” 558 U.S. 78 (2010) (opinion of J. Stevens).
expenditures, on the other hand, do rise when contribution limits are relaxed. Thus, any study looking only at total lobbying contributions miscasts the effect of the regulatory change. Finally, previous studies show that contributions may actually be larger with a cap than without, or when the politician refuses to accept donations from certain groups. These results depend on the politician’s inability to act strategically and have not been borne empirically. When the politician-as-bill-writer assumes a more central role, limiting donations to PACs or refusing to accept donations from certain groups can no longer increase contributions. The strategic politician has better tools at his disposal to increase contributions.

In the next section we briefly discuss the central institutional background motivating our model. We then introduce our model of the lobbying process as a discriminatory all-pay auction with endogenous bidder valuations. A simplified example, analysis of the more general setting, an overview of the most relevant literature, and a conclusion follow. All proofs are contained in the Appendix.

2.2 Background

By the late 1800s distrust of campaign funding practices was growing. In his successful 1896 presidential bid, William McKinley financed the bulk of his campaign with “regular assessments on corporations of consequence throughout the country” in exchange for support of the business agenda (Thayer, 1974, pp. 49-50). State legislatures began enacting campaign finance legislation and Theodore Roosevelt made “clean government” a major theme of his 1904 presidential candidacy. Congress followed suit banning direct corporate donations to federal candidates with the Tillman Act (1907). Sentiment was such that the Senate felt compelled to even offer a motivation for the act, instead saying “the evils of the use of [corporate] money in connection with political elections are so generally recognized that the committee deems it unnecessary to make any argument in favor of the general purpose of this measure” (Senate Report No. 3056, 59th Congress, 1st

---

5The same restrictions were extended to labor unions in the 1940s.
Session, 2 (1906)).

Two Congressional acts and a string of Supreme Court cases further rearranged campaign finance to its pre-
_Citizens United_ state.\(^7\)

Fear of the inherently corrupting role that corporate treasury funds pose to the political system has changed little.\(^8\) The contributions of political action committees are capped at $5,000 per federal candidate with individual contributions restricted to $2,400. PACs are not limited in their ability to spend independently of a candidate’s campaign organization, though voluntary contributions to the political committee are constrained by a $5,000 per individual limit. The purpose of these spending restrictions is not to limit money in politics _per se_, but to prevent the contributions from buying political favors (or the appearance thereof).

### 2.3 A model of lobbying with endogenous valuations

A lawmaker writes a bill that he considers making into law. The bill is a policy proposal impacting two special interest groups. The interests – one a corporation and the other a labor union – have competing preferences. They lobby the lawmaker hoping to influence the policy by making political contributions.

Formally, we model the lobbying game as a discriminatory all-pay auction with private values and simultaneous bids. There are three risk-neutral players, a lawmaker and two interests _L_ and _R_. The lawmaker is writing a bill that proposes a policy \(x \in \mathcal{X} = [x, \bar{x}]\). The policy \(x\) includes some allocation of benefits to be awarded to one of the interest groups. The lawmaker is selling the decision of which interest group to privilege and by how much. _L_ and _R_ have preferences over the

---

\(^6\)The act was ineffective in removing corporate contributions from political campaigns as executives simply took raises and contributed as individuals (Malbin, 1984, p. 244).


\(^8\)See, for example, Supp. Brief for Appellee, _Citizens United v. Federal Election Commission_. 

policy the lawmaker proposes.\footnote{\textit{We assume the lawmaker is decisive in that whichever allocation he supports ultimately becomes law, perhaps because the lawmaker is the median member of the legislative body.}} Denote $L$’s ($R$’s) money-denominated valuation of winning the spoils included in policy $x$ as $v_L(x)$ ($v_R(x)$). Let $v_L(\cdot)$ be decreasing and $v_R(\cdot)$ increasing in $x$. To fix ideas, think of $L$ as a labor union preferring “left-wing” policies while $R$ is a corporation preferring policies more “right-wing.”

To make the analysis interesting, assume there exists a centrist policy $x^c$ such that $v_L(x^c) = v_R(x^c)$. The setting described above can be thought of as one in which a lawmaker is awarding some government largesse. The interest groups are competing to acquire as large a share of the largesse as possible conditional on currying the politician’s favor. A ‘low’ or ‘left’ policy $x$ corresponds to giving the majority of the spoils to group $L$, while a higher value of $x$ awards a larger share to $R$. The value functions $v_i(\cdot)$ are known to all. Without loss of generality assume $v_i(x) > 0 \forall x$.

The interests compete to influence the lawmaker’s policy proposal by contributing amounts $b_L$ and $b_R$. The contributions are irrevocable and are forfeited to the politician once made. They cannot be made conditional on the outcome of the policy.\footnote{This setting resembles menu auction models of resource allocation (Bernheim and Whinston, 1986).} Finally, the interest groups have the option of not bidding at all, for which they attain utility zero.\footnote{View this as a reduced-form version of a model in which the lawmaker chooses both the size and distribution of a pie. A special interest exiting the lobbying process receives no share of the spoils.} As there are no pure-strategy equilibria in the standard all-pay auction, we represent interest $i$’s strategy by the distribution function of bids made in equilibrium $F_i(b) \equiv F(b_i(v_i(x))) \in [0, 1]$.

The special interests are not bidding over some exogenous prize placed before a lawmaker to distribute; the lawmaker himself designs the prize. Let $u(b_L, b_R)$ denote the lawmaker’s utility from contributions $b_L$ and $b_R$ with $\frac{\partial}{\partial b} u(b_L, b_R) > 0$. Since contribution behavior will be in mixed-strategies, the lawmaker seeks to maximize the expected contributions he receives where $W(x) \equiv E[u(b_L, b_R)|F_L, F_R, x]$ is the campaign contributions he expects to receive from writing bill $x$.\footnote{With a risk neutral seller, $W(x) = E[b_L + b_R|F_L, F_R, x]$.} The lawmaker influences the interests’ donations by crafting different bills.
The strategic politician not only devises the policy the bidders are competing over, but he also designs the rules of the competition. Two bids of equal size do not need to be treated equally. We suppose the politician can provide an advantage to a special interest group. The benefit, which may depend on the valuations \( v_i(x) \), takes the form of a level of discrimination \( \gamma(v_L(x), v_R(x)) \geq 0 \) provided in favor of one interest group. The discriminatory benefit will be abbreviated \( \gamma(x) \) in what follows. If the labor union receives benefit \( \gamma \) under the lawmaker’s auction, the corporation wins the politician’s vote only when \( b_R \geq b_L + \gamma \). In effect, the lawmaker tells the corporation, “I have a long relationship with the labor union and it’s a cause I feel strongly about. In order to win my vote, you are going to have to really bowl me over.” Of course, the lawmaker does not actually need to have any allegiance one way or the other: he uses the terms of the lobbying game to increase his expected receipts.

On some level, the lawmaker is selling the policy process outright to the winning bidder; if the contributions are judged to be equal, each bidder wins with equal probability. Selling one’s vote is extreme, but as Justice Stevens notes in his dissent to *Citizens United*, “the difference between selling a vote and selling access is a matter of degree, not kind” 558 U.S. 57 (2010) (opinion of J. Stevens). Whether access, an allocation, or a vote is for sale changes little in the interpretation of our model. Finally, to amplify the concerns of those opposed to the *Citizens United* ruling, we suppose the lawmaker is right-leaning so that if \( W(x') = W(x'') \) he selects the proposal more advantageous to \( R \) when resolving his indifference.

Requiring special interests to contribute to lawmakers through political action committees constrains donations. A union or corporation can only contribute as much as it can raise through voluntary employee contributions. Let \( c_L \) and \( c_R \) denote these amounts. The *Citizens United* ruling allows general treasury funds to be used for political donations. Contributions are now limited to \( c'_i > c_i \).

---

13An alternative formulation of the model could have allowed the prize \( x \) to be an arbitrary function of the realized bids \( b_i \). As bidding strategies are non-trivially mixed, the lobbyists would not know the actual value of the prize over which they are competing to win when making their contributions. Although less general, we feel the setup presented in this paper is more in line with the actual lobbying process.

14Campos and Giovannoni (2007) examine firms in 25 transition economies and find evidence that lobbying and corruption are substitute goods.
2.4 Example

Before proceeding to the general analysis, briefly consider a simplified example. Suppose the lawmaker is considering the division of a defense procurement contract. Two interest groups, \( L \) and \( R \), each want to be awarded as much of the contract as possible. The politician can select an allocation \( x \in [.2,.9] \). Allocation \( x \) provides \( L \) with spoils worth \( v_L(x) \). \( R \)'s valuation of the allocation is defined similarly. As the groups are in competition over the contract, let \( v_L(x) = 1 - x \) and \( v_R(x) = x \). Finally, assume there is a cap \( c \) on lobbying contributions.

This example is special in assuming the valuation functions \( v_i(\cdot) \) are linear. Linearity leads the comparison of maximum revenue across different auction designs to be particularly simple. Since the marginal values \( v'_i(x) \) of an increase in \( x \) are constant, if a movement towards a more extreme policy delivers greater revenue, even further movements in policy are still more desirable. That is, with linear value functions, revenue is maximized either with a centrist policy \( (x = .5) \) or with an extreme policy; every policy proposal in between generates lower expected revenue.

Define \( v_1(x) \) to be the larger of the two valuations when the policy is \( x \). When the lawmaker runs a traditional non-discriminatory all-pay auction, revenue is maximized by choosing policy \( x = .5 \) so that \( v = .5 = v_1(x) = v_2(x) \). For modest cap levels \( c \geq .25 \), expected lobbying contributions are \(.5\); for more stringent cap levels, \( c < .25 \), expected contributions \( W(x) = 2c \). The lawmaker wants both groups to value the policy, which leads the optimal policy proposal to be centrist.

When the cap is set at \( c \leq .5 \), the politician can do no better than this centrist policy with no discrimination. That is, for \( c \leq v \), for any \( \gamma > 0 \), the maximum attainable revenue is less than that generated by \( \gamma = 0 \) and \( v_1 = v_2 = v \). As the cap is relaxed, optimal lawmaker behavior is to select an extreme policy, \( x = .9 \) in this case, and give the lower valuation group an advantage \( \gamma = c - .1 \). Finally, when the cap no longer binds at all, the lawmaker continues proposing an extreme policy while offering \( \gamma = v_1 - v_2 = .8 \). While the expected lobbying contributions rise as the cap is relaxed, the policy proposal initially remains centrist. Even from this simple example it is easy to see that looking only at expected contributions...
leads to a distorted picture of the effect of a change in the regulatory environment.

2.5 Analysis

The special interest groups want to buy the lawmaker’s vote but they are limited in how much they can spend. We examine the effect of *Citizens United* by comparing equilibrium behavior pre- and post-ruling. We begin by analyzing equilibrium behavior when the lawmaker cannot discriminate among the bidders. So $\gamma = 0$ and the vote is awarded to whichever special interest contributes the largest lobbying donation. We later build on the results here to allow the amount of discrimination to be chosen endogenously by the optimizing lawmaker.

Forced to lobby with PACs funded entirely by voluntary employee contributions, the interests can donate up to $c_L$ and $c_R$. Define $c = \min\{c_L, c_R\}$ as the effective bidding cap created by campaign finance law. It is clear that even the richer interest’s donations will never exceed $c$.

2.5.1 Endogenous valuations but no discrimination

Given $x$, $v_L(x)$ and $v_R(x)$ are primitive. Che and Gale (1998) consider a static game with no discrimination where valuations $v_1 > v_2 > 0$ are given exogenously and bids are constrained by a common cap $c > 0$. We first extend their results to the case when valuations are endogenously determined by the legislation written by a strategic lawmaker. Accordingly, the expected revenue the lawmaker obtains from selling his vote is now endogenous. In equilibrium, the lawmaker writes a bill promising maximal revenue.

The intuition for our results can be seen by considering a cap $c > \frac{v_2}{2}$. The expected revenue in this case is fixed at $[1 + \frac{v_2}{v_1}]\frac{c}{2}$ with exogenous valuations. For the strategic politician, an increase in $v_2$ generates greater expected revenue, even when the increase in $v_2$ is obtainable only by decreasing $v_1$ significantly. The lawmaker wants to stiffen the competition over his vote by eliminating some of the gap between $v_1$ and $v_2$. If the lawmaker is able to change the lobbyists’ valuations by changing the policy he proposes, the strategic lawmaker with endogenous bill-
writing ability eliminates the gap in valuations.

The key driving force in our analysis is the relationship between the spending cap \(c\) and the valuation \(v\) when \(x\) satisfies \(v_L(x) = v_R(x) = v\). Call a spending cap stringent if \(c < \frac{v}{2}\) and modest otherwise.

**Theorem 4.** In the lobbying game with endogenous bill writing and no discrimination, equilibrium expected revenue is \(W(x) = v\) for \(c > \frac{v}{2}\) and \(W(x) = 2c\) for \(c \leq \frac{v}{2}\), where \(v\) is the valuation of policy \(x\) such that \(v_L(x) = v_R(x) \equiv v\).

For \(c > \frac{v}{2}\), \(x\) satisfies \(v_L(x) = v_R(x)\). For \(c \leq \frac{v}{2}\), the policy proposal is \(x \in \{x' : W(x') = 2c\}\). The corporation successfully buys the lawmaker’s vote with probability \(\frac{1}{2}\) in all equilibria.

**Proof.** See Appendix. \(\square\)

There are multiple policy proposals capable of attaining expected revenue \(W(x) = 2c\) when \(c \leq \frac{v}{2}\). Supposing the lawmaker resolves his indifference over competing policies offering the same expected revenue by choosing the policy most advantageous to the corporation, in equilibrium the lawmaker optimally proposes \(x = \arg\max_{x'}\{x' : v_L(x') \geq 2c\}\) and the equilibrium policy proposal is unique.\(^{15}\) With indifference resolved in favor of the corporation, \(x = \arg\max_{x'}\{x' : v_L(x') = 2c\}\) is selected.

When lobbyists’ valuations are asymmetric, as is the case generically when valuations are given exogenously, there exists underdissipation of rents (Baye et al., 1993). The politician tenders his vote to the lobbyist with the highest bid but the auction’s expected revenue is strictly below the lower valuation \(v_2\). This remains true even when campaign donations are unrestricted.

When valuations are determined endogenously by lawmaker behavior, the politician proposes policy \(x\) for which \(v_L(x) = v_R(x) = v\) and expects revenue \(W(x) = v\) from selling his vote. For modest spending limits (\(c \geq \frac{v}{2}\)) the auction generates expected revenue equal to the lower valuation of the contestants. The politician’s strategic behavior ensures full dissipation of rents.

\(^{15}\)This resolution of indifference is most in the spirit of those concerned about the Citizens United ruling.
The *Citizens United* ruling nullified campaign finance law requiring labor unions and corporations to contribute to political campaigns solely through employee-funded political action committees; general treasury funds were made available for lobbying purposes. The decision increases the amounts the special interests can spend from $c_L$ and $c_R$ to $c'_L > c_L$ and $c'_R > c_R$ respectively.

Equilibrium behavior with a modest spending cap is the same as in the absence of any contribution limits. Expected revenue for a given policy $x'$, $W(x') = \left[ 1 + \frac{v_2(x')}{v_1(x')} \right] \frac{v_2(x')}{2}$, is uniquely maximized by selecting $x$ such that $v_L(x) = v_R(x) = v$. In equilibrium the contestants randomize uniformly over $[0, v]$, each expecting to pay $\frac{v}{2}$ and successfully buy the lawmaker’s vote half of the time.

When both special interest groups are able to create well-funded PACs, the strategic bill writer proposes the policy for which the interest groups have the same valuation. Contrast this with the case of a stringent cap. When $c < \frac{v}{2}$ there are many policies capable of generating maximal revenue. If, as we suppose above, the politician selects from these the policy providing the greatest surplus to his pet interest, the relaxation of stringent campaign finance restrictions has a moderating effect on the policy process. From society’s point of view, lobbying contributions increase post-*Citizens United* but not without benefit: the policy proposal becomes more centrist. Never does the relaxation of spending restrictions increase the corporation’s likelihood of buying the lawmaker’s vote. Contrary to the fears of the ruling’s opposition, the labor union prefers the equilibrium policy proposal post-*Citizens United* to the policy proposal pre-*Citizens United*. Importantly, this holds regardless of which special interest has the greater ability to pay. As campaign finance law limited political contributions as an instrument to prevent policy from being corrupted by extreme interests, it is far from clear that the sky is falling post-*Citizens United* when the effect may be to moderate the policy proposed by the lawmaker.

---

16Recall that a spending cap is called modest if $c \geq \frac{v}{2}$ and stringent otherwise.
2.5.2 Endogenous valuations and endogenous discrimination

We now allow the lawmaker to treat equal bids from the lobbyists differently. We begin by establishing that the maximum advantage a lawmaker provides an interest group is limited. Consider policy proposal $x$ and suppose group $L$ is provided with an advantage $\gamma(x) > 0$. If $\gamma(x) > v_R(x)$, group $R$ must bid more than its valuation in order to win the auction. Since $R$ does better by remaining out of the lobbying process entirely, $L$ is able to win the prize with a bid of zero. The expected revenue generated by such an auction design is zero and so the strategic lawmaker will never set $\gamma(x) > v_1(x)$. For similar reasons the advantage $\gamma(\cdot) \geq c$ will never be employed.

Pastine and Pastine (2009) consider an exogenous valuation all-pay auction embellished with an exogenous benefit provided to one of the two bidders. We first extend their results when lobbyists are unconstrained in how much they can contribute to allow the benefit to be endogenously determined by a lawmaker seeking to maximize expected contributions.

**Proposition 1.** In the absence of a cap on contributions, for any valuations $v_1(x)$ and $v_2(x) \leq v_1(x)$, expected contributions are maximized by providing the lower valuation group with advantage $\gamma(x) = v_1(x) - v_2(x)$. With $\gamma(x)$ set optimally, the expected contributions from a policy proposal $x$ are $\tilde{W}(x) = v_1(x) - \frac{v_2(x)}{2} + \frac{(v_2(x))^2}{2v_1(x)}$.

**Proof.** See Appendix. \qed

In the absence of a cap on contributions, Proposition 1 shows that for any given policy, expected revenue is maximized by completely offsetting the valuation advantage of the stronger lobbyist. The reasoning here is similar to that informing the optimal auction design literature and third-degree price discrimination behavior of a monopolist. In an optimal auction, a bidder with a higher valuation is required to make a higher bid in order to win the item (Myerson, 1981); the additional hurdle imposed upon high valuation bidders dissuades them from underbidding relative to their true valuation. In a setting with multiple differentiated markets, the price for the monopolist’s good is set higher in markets
Figure 2.1: Bidding strategies with $\gamma$ set optimally when there is no cap.

with lower price elasticity of demand (Bulow and Roberts, 1989). With $\gamma(x)$ set optimally, the high valuation bidder is less able to intimidate the low valuation bidder.

An interior optimal policy proposal requires that $v_2(x)$ not be too much smaller than $v_1(x)$. This is because when valuations differ significantly, the overwhelming majority of the lawmaker’s lobbying receipts are expected to come from the high valuation lobbyist. Further decreases in $v_2$ do not much affect expected receipts while further increases in $v_1$ offset with additional advantages for the lower valuation lobbyist cause the high valuation lobbyist to bid more aggressively. See Figure ??, which shows the equilibrium bidding strategies for the lobbyists when $\gamma(x)$ is chosen optimally.

Whenever $v_1(x) > 2v_2(x)$, a more extreme policy proposal produces higher expected lobbying contributions. This is precisely what is meant by $v_2(x)$ being too small relative to $v_1(x)$ for $x$ to be an equilibrium policy proposal.

Remark 7. Absent a cap on lobbying contributions, if $v_1(x) > 2v_2(x)$ and $x \notin \{\bar{x}, \overline{x}\}$, then the policy proposal $x$ cannot be an equilibrium.

Though useful for generating greater expected contributions, providing a discriminatory advantage to the low valuation lobbyist is an imperfect substitute for that lobbyist having a higher valuation. To see this, consider the rev-
enue produced when valuations are \( v_1 \) and \( v_2 \) versus that generated with valuations \( v_1 \) and \( v_2' < v_2 \) and \( \gamma = v_2 - v_2' \). In the second case advantage \( \gamma \) is provided to the low valuation lobbyist to exactly offset the decrease in her valuation; \( v_1 \) remains the same in both cases, which would happen if \( v_1(\cdot) \) was not strictly monotonic. Abusing notation slightly, the expected contributions under the first case are \( W(v_1, v_2) = \frac{v_2}{2} + \frac{v_2' v_2}{2v_1} \). The second case has expected contributions \( W(v_1, v_2') = \frac{v_2'}{2} + v_2 - v_2' + \frac{(v_2')^2}{2v_1} \). Expected revenue in the first case will be greater whenever \( v_2 + v_2' > v_1 \). For small changes in \( v_2 \), this is \( 2v_2 > v_1 \). Holding \( v_1 \) fixed, when \( v_2 \) is large relative to \( v_1 \), the revenue effect of a decrease in valuation \( v_2 \) cannot be fully compensated by an increase in the advantage given to the lobbyist. Here already we see the attractiveness of centrist policies.

For \( \gamma(x) \) optimally chosen, an increase in \( v_1(x) \) coupled with a decrease in \( v_2(x) \) and offset by the requisite change in \( \gamma(x) \) causes the high valuation bidder’s new bid strategy to first-order stochastically dominate his old strategy. When \( v_2(x) \) is small relative to \( v_1(x) \), this tradeoff is worthwhile since, in expectation, the lawmaker is receiving a contribution from the high valuation bidder more than twice as large as that from the low valuation bidder. (See Figure ??.)

Now consider a cap on donations that is binding \((c \leq v_1(x))\). A given cap \( c \) is binding for some policy proposals but not for others. Analysis of equilibrium behavior now entails comparing the optimal policy and advantage conditional on the cap being binding versus optimal behavior conditional on the cap being non-binding.

As before, we proceed by first extending Pastine and Pastine (2009) to allow the strategic lawmaker to endogenously determine the rules governing the auction. Recall that Pastine and Pastine (2009) consider a setting with exogenous valuations and exogenous advantage. In the context of their model, a cap \( c \) is either binding or it’s not; in our setting whether a cap is binding is determined by the behavior of the lawmaker. We next show that conditional on the cap being binding, the lawmaker optimally sets \( \gamma(x) = c - v_2(x) \).

**Proposition 2.** **Conditional on a contribution cap c being binding, for a given policy proposal x, the lawmaker optimally sets \( \gamma(x) = c - v_2(x) \).** The expected con-
Figure 2.2: Comparing $\gamma$ set optimally versus an increase in $\gamma$, with binding cap.

Contributions from policy $x$ with $\gamma(x)$ chosen optimally are bounded above by

$$\tilde{W}(x) = c + \frac{v_2(x)}{2} - \frac{c-v_2(x)}{v_1(x)} + \frac{(v_2(x))^2}{2v_1(x)}$$

Proof. See Appendix.

The proposition above establishes that the optimal advantage to give to
the low valuation bidder is $\gamma(x) = c - v_2(x)$. The lawmaker wants to bring the
low valuation lobbyist’s maximum effective bid up to the cap. Interestingly, the
lawmaker does not want to eliminate the full gap in valuations with a bidding
advantage. Formally, $\gamma = c - v_2(x)$ produces greater expected contributions than
$\gamma' > c - v_2(x)$ even when $v_2(x) + \gamma < v_1(x)$. This is because the increase in
$\gamma$ actually lowers the expected contributions from the high value lobbyist while
leaving the low valuation lobbyist’s strategy unchanged. Figure 2.2 shows exactly
this. Optimally, $\gamma(x) = c - v_2(x)$. With a binding cap the optimizing lawmaker
wants to allow the high valuation lobbyist to maintain some of his advantage. This
is not the case when the cap is not binding.

The optimal policy proposal depends on the ratio of marginal valuations.\(^{17}\)

However, if the lawmaker is already providing a large advantage to the low val-
uation bidder to make him competitive, the lawmaker should just go ahead and

\(^{17}\)In order for a policy proposal $x$ to be optimal, it must be that $\frac{v'_1(x)}{v'_2(x)} = \Lambda(x, c)$ where $\Lambda(x, c) = \frac{(v_1(x))^2 - 2v_1(x) \cdot c + 2v_1(x)v_2(x) \cdot (2v_2(x) \cdot c - (v_2(x))^2)}{2v_1(x)}$.\(\)
Figure 2.3: Bidding strategies with $\gamma$ set optimally when there is a binding cap.

propose an even more beneficial policy for the high valuation lobbyist coupled with an even greater competitive disadvantage. As Figure ?? shows, an increase in $v_1$ causes a FOSD shift in the high valuation lobbyist’s bidding strategy. The revenue effect on the low valuation lobbyist is unclear: the ensuing decrease in $v_2$ may increase expected contributions from the low valuation lobbyist.

Clear conclusions are difficult in this general setting. Even when a more extreme policy produces greater expected contributions than a less extreme policy, it does not necessarily follow that a maximally extreme policy proposal is the best strategy for the lawmaker. The optimality conditions stated above are local and the partial revenue ordering $W(x'') > W(x')$ for $x'' > x'$ does not imply $W(x') > W(x)$ for $x' > x$. More extreme policies may be better than less extreme policies, but both may be worse than a centrist policy.

As the cap on lobbying expenditures is raised, the politician is better able to take advantage of a lobbyist with a high valuation. As to be expected, expected contributions rise as the cap is relaxed. But the effect on policy is less pronounced: a change in the regulatory environment that leads to considerably larger expected lobbying donations may have no effect on policy.

In fact, one reason to be optimistic following the change in the regulatory environment caused by *Citizens United* is the enduring allure of designing a policy

---

18When $v_1(x) < 2(c - v_2(x)) = 2\gamma(x)$, the policy should be made more extreme.
that both lobbyists value equally. The power of symmetric valuations is so strong that policies that both lobbyists value less can yield greater expected revenues. In the context of the model above, it may well be that \( W(x) > W(x') \) even though \( v_i(x) < v_i(x') \) for both lobbyists.

**Remark 8.** Even when \( v_1(x) < v_1(x') \) and \( v_2(x) < v_2(x') \), it may be that \( W(x) > W(x') \) if the valuation gap \( v_1(x) - v_2(x) < v_1(x') - v_2(x') \). This remains a possibility whether or not a cap on contributions is present.

The lawmaker’s enemy is not low valuations of the prize but, rather, unequal competition in the awarding of the prize.

The strategic lawmaker also takes advantage of the high valuation lobbyist in the other sense of the term: the expressed fear of lobbying contributions being used by politicians to shake-down monied interests is real. The lobbying game can be used as an implicit tax. By increasing the corporation’s valuation for the policy and then offsetting the increase with a competitive handicap, the lawmaker is able to siphon more resources from the corporation.

A final point deserves emphasis. *Citizens United* may have changed the regulatory landscape of the lobbying environment by removing a statutory cap on lobbying contributions, but it did not remove all financial constraints. If, as the ruling’s opponents fear, the corporation now mans a massive war chest for lobbying while the labor union remains cash-poor, the effect on both policy and contributions is minimal. Both special interests are effectively constrained by the poorer group. While the corporation is now allowed to use its general treasury funds, in equilibrium it will not. The amount lobbying contributions increase and the degree to which policy becomes more coopted is governed by the resources of the poorer group, not the wealthier group.

### 2.6 Related literature

The present paper is related to prior work on both all-pay auctions and models of influence buying.
The all-pay auction without a bidding cap has been analyzed in several prior studies (Moulin, 1986; Hillman and Riley, 1989; Baye et al., 1993, 1996; Siegel, 2009). Laffont and Robert (1996) show the optimality of the all-pay auction with a reserve price when bidders’ financial constraints are commonly known. Economic theory suggests the spending caps may actually increase donations (Che and Gale, 1998). Gavious et al. (2002) relate the revenue implications of bid caps to the shape of bidders’ marginal cost functions. Other variations include alternating offer protocols (Yildirim, 2005), having a committee award the prize (Amegashie, 2003), making the size of the reward depend on bids (Kaplan et al., 2002), and endowing sellers with policy preferences (Pastine and Pastine, 2009). The treatment of lobbying as an all-pay auction dates back to Tullock (1980). Some studies permit lobbyists to break the law and contribute more than the cap (Kaplan and Wettstein, 2006; Che and Gale, 2006).

Lobbying is often modeled as a game of strategic information transmission (Austen-Smith, 1993; Lohmann, 1995; Ambrus et al., 2010; Ivanov, 2009). Dekel et al. (2009) consider an alternating-offer bargaining game where lobbyists compete to buy influence by purchasing votes. (Groseclose and Snyder (1996) and Banks (2000) also consider a sequential though not alternating bargaining model.) Potters et al. (1997) motivate campaign donations as an indirect route to buy influence by helping sway undecided voters. Potters and Sloof (1996) survey the empirical data on the ways interest groups affect political behavior while Becker (1983) studies political redistribution caused by competing pressure groups.

### 2.7 Conclusion

In Citizens United v. Federal Election Commission, the Supreme Court ruled that corporations can legally use their general treasury funds for political donations. The decision overturned decades of campaign finance law and generated significant fear that rich corporations would use their wealth to effectively buy preferential treatment from politicians. Using a stylized model that captures many of the salient features of the institutional setting, we have shown that this
fear is not necessarily borne out once lawmakers are admitted as strategic agents. Unlike previous studies, we have focused on the effect of the regulatory change on policy and not contributions. Even when lobbying contributions rise, the effect on policy is unclear. There are several reasons to be optimistic about the regulatory landscape post-*Citizens United*. First, the ruling did not remove all financial constraints facing lobbyists. In a lobbying setting where one group remains financially-constrained while the other gains access to large amounts of potential contributions, the wealth advantage of the second group is of little use in equilibrium. Second, moving policy to the extremes causes one group to value highly the policy while the other holds it in low regard. The decrease in competitive forces can be only partially offset by discriminating between lobbyists. Indeed, the enduring allure of a centrist policy remains even absent limits on campaign contributions.

I thank Chulyoung Kim for his role as coauthor on this chapter and the UCSD theory group, especially Nageeb Ali, David Miller, Joel Sobel and Joel Watson for helpful discussions. Gordon Dahl, Gary Jacobson and Sebastian Saiegh also deserve thanks for improving the chapter.

*The preceding chapter was coauthored with Chulyoung Kim.*
Appendix A

Appendix to Chapter 1

A.1 Crowdsourcing

This section provides a brief overview of the crowdsourcing labor spot market. Crowdsourcing is the process of delegating work to an undefined group of people (a crowd) through an open call online. Of the dozens of work exchanges where firms can hire workers, Amazon’s Mechanical Turk (AMT) is the most prominent. Created in-house in 2005 to find duplicates among the company’s product webpages, the service rapidly expanded and by 2007 comprised a pool of more than 100,000 workers in over 100 countries completing various types of tasks, such as transcribing podcasts, rating and tagging images, and writing/rewriting sentences. There are now more than 150,000 jobs available at any time (Caulfield, 2011).

The paid crowdsourcing market has grown considerably since AMT’s founding. oDesk, a competing platform, has 2.3 million registered workers and posted half a million jobs in the second quarter of 2012 (oDesk, 2012). Almost 8 million hours of work were performed in that quarter alone and worker earnings on oDesk tallied $250 million in 2011 (Vanham, 2012).

A paid survey conducted on AMT in February 2010 revealed workers from 68 different countries; the United States is most prevalent at 45% followed by India at 34%. Young workers are overrepresented, even when compared to the general population of Internet users. Self-reported education levels are also greater than those of the general populations.
Most workers spend a day or less working on AMT, completing 20-100 jobs and earning $20 or less per week. There exists a high-end of the income distribution with workers earning more than $1,000 per month. More than 20% of Indian workers report AMT as their primary source of income (10% of American workers), with an additional 35% (60% for American workers) using AMT as a secondary source of income. The primary motivation for working on AMT is to earn cash while spending free time fruitfully (60% of American workers and 70% of Indian workers).\footnote{See, also, Mason and Watts (2009) and Suri and Watts (2011) on demographics and Ross et al. (2010) and Ipeirotis (2010) on earnings.}

Field experiments carried out on AMT suggest workers respond to economic incentives in a predictable fashion (Mason and Watts, 2009; Horton and Chilton, 2010; Horton et al., ress; Paolacci et al., 2010). Explicitly informing workers that the accuracy of their responses was being measured and used to determine whether payment for their work would be provided had no discernible effect on either the quality or quantity of output: participants appear to treat their pay as necessarily performance dependent. Several studies have shown that the performance of subject-matter experts can be achieved via crowdsourcing (Snow et al., 2008; Marge et al., 2010; Urbano et al., 2010; Alonso and Mizzaro, 2009; Paolacci et al., 2010).\footnote{Horton et al. (ress) and Suri and Watts (2011) recreate laboratory experiments on AMT and find no significant differences between the settings.}

Guidelines for creating contracts on AMT are scant and consist of little more than “be clear.” The definition of acceptable output and the conditions for payment are not normally presented on the main job posting page. Figure A.1 provides an example from an actual job posting on AMT.

Hiring multiple workers is probably the most common method of quality assurance (Mason and Suri, 2012). (The job in Figure A.1 states “we [the firm] verify ALL answers.”) Obtaining multiple responses is cost-efficient for many tasks on AMT. It also tends to perform well (Snow et al., 2008). Another option, known as the “Gold Standard,” is built around the idea of including tasks within each job for which the firm already knows the correct answer. A worker’s performance can
Figure A.1: Example job posting on Amazon’s Mechanical Turk.

then be judged on the subset of these tasks.\textsuperscript{3}

\textsuperscript{3}Little et al. (2010) find that for tasks with clearly unreasonable answers, firms can effectively vet the responses of workers by employing additional agents for judgement. The additional agents verify whether the first worker provided a coherent response to the question.
A.2 Proofs

Proof of Remark 1

Remark 1. When effort is contractible, $Q^* = (0, \ldots, 0)$ and

$$W^*(r^i) = \begin{cases} J \cdot c & \forall r^i \in \overline{R}^i \text{ and } \sigma^i = \overline{\sigma}^i \\ 0 & \forall r^i \in \overline{R}^i \text{ and } \sigma^i \neq \overline{\sigma}^i. \end{cases}$$

Proof. IR is satisfied with a per-task wage equal to the worker’s cost of effort. No monitors need to be employed due to the lack of incentive considerations.

Proof of Lemma 1

Lemma 1. For any equilibrium of the firm’s problem, there exists another equilibrium providing the same expected payoffs to the firm and all agents in which (a) the firm adopts an implementation plan $Q^i = (q_1, \ldots, q_J)$ with $q_t = \sum_{n=2}^{\infty} f_t(n)$ and $f_t(n) = 0$ for all $n > 2$, and (b) workers use truthful recommendation strategies $\rho = \overline{\rho}$.

Proof. In equilibrium, two workers are sufficient to induce high effort on every task. Hiring additional workers on a task changes neither the incentives facing workers nor the probability the firm takes the correct action. The only effect is that the firm’s wage bill increases.

The firm correctly interprets recommendations in equilibrium. For any equilibrium with $\rho \neq \overline{\rho}$, there exists another equilibrium where recommendation strategies are $\overline{\rho}$ yielding identical expected payoffs for all actors.

Proof of Lemma 2

Lemma 2. The optimal incentive organization sets $W^i(r^i) = 0 \ \forall r^i \in \overline{R}^i_{kj}$ whenever $k \neq j$. 
Proof. Equation 1.1 holds

\[ W_{00}[\overline{Pr}(R_{00}^i) - Pr(R_{00}^i)] + W_{10}[\overline{Pr}(R_{10}^i) - Pr(R_{10}^i)] + \ldots + W_{JJ}[\overline{Pr}(R_{JJ}^i) - Pr(R_{JJ}^i)] \geq c[J - \sum_{k=0}^{J} k \cdot \zeta(k|\sigma^i)]. \]

For any \( R_{kj}^i \) with \( k \neq j \), \( \overline{Pr}(R_{kj}^i) = 0 \). (If all workers are exerting effort, then a worker necessarily produces matching output whenever monitored.) So \( \overline{Pr}(R_{kj}^i) \leq Pr(R_{kj}^i) \) and setting \( W^i(r^i) > 0 \) for \( r^i \in R_{kj}^i \) can only make Equation 1.1 harder to satisfy.

Proof of Theorem 1

The proof of Theorem 1 is divided into two lemmas. We begin by noting that is is necessary and sufficient that the firm discourage strategies \( \sigma^i = (p_1^i, \ldots, p_J^i) \) where \( p_t^i \in \{0, 1\} \). By discouraging deterministic effort strategies the firm also discourages strategies in which workers mix in choosing to exert effort. The first lemma shows that for any implementation plan \( Q^i = (q_1^i, ..., q_J^i) \) with some \( q_j^i \neq q_k^i \), there exists an implementation plan such that \( q_t^i = q \) for \( t = 1, ..., J \) providing the same disincentives to shirking at lower expected cost. The second lemma shows that it is sufficient for the firm to discourage \( \sigma^i \). The wage required to induce effort follows.

Lemma 3. For any implementation plan \( Q = (q_1, ..., q_J) \) with \( q_j \neq q_k \) for some \( j, k \), there exists another implementation plan \( \hat{Q} = (q, ..., q) \) providing the same disincentives to shirking at lower expected cost.

Proof. The incentive constraint requires

\[ \overline{Pr}(\overline{R}^i) \cdot W - Jc \geq Pr(\overline{R}^i) \cdot W - nc \]

for all strategies \( \sigma^i \) that specify shirking on exactly \( n = 0, ..., J \) tasks. Since \( \overline{Pr}(\overline{R}^i) = 1 \), the required transfer is governed by \( Pr(\overline{R}^i) \), which itself depends on the firm’s implementation plan.
Let \( Q = (q_1, ..., q_J) \) be given and consider strategy \( \sigma^i \) in which \( i \) shirks on \( n \) tasks and the firm’s monitoring probability on each such task is \( q_k, k = 1, ..., n \). Then \( P_r(\mathcal{R}^i) = \left[ \frac{1 - q_1 + 2q_1 \pi}{1 + q_1} \right] \cdot \cdots \cdot \left[ \frac{1 - q_n + 2q_n \pi}{1 + q_n} \right] \).

For any collection of monitoring probabilities \( (q_1, ..., q_J) \), the firm can set a uniform monitoring probability \( q \) such that \( P_r(\mathcal{R}^i) \) is the same. In other words, there exists \( q \) such that
\[
\left[ \frac{1 - q + 2q \pi}{1 + q} \right]^n = \left[ \frac{1 - q_1 + 2q_1 \pi}{1 + q_1} \right] \cdot \cdots \cdot \left[ \frac{1 - q_n + 2q_n \pi}{1 + q_n} \right].
\]

Formally, \( \left[ \frac{1 - q + 2q \pi}{1 + q} \right] \) is the geometric mean of \( \left[ \frac{1 - q_1 + 2q_1 \pi}{1 + q_1} \right] \cdot \cdots \cdot \left[ \frac{1 - q_n + 2q_n \pi}{1 + q_n} \right] \).

The firm cares about the expected wage bill. For each task \( t \), \( 1 + q_t \) workers are being paid in expectation, so the firm expects to pay \( J + \sum_{k=1}^{J} q_k \) workers in total. Suppose under implementation plan \( Q = (q_1, ..., q_J) \), \( q_1 > q_2 \), so that
\[
\frac{1 - q_1 + 2q_1 \pi}{1 + q_1} < \frac{1 - q_2 + 2q_2 \pi}{1 + q_2}
\]
and
\[
\ln \left( \frac{1 - q_1 + 2q_1 \pi}{1 + q_1} \right) < \ln \left( \frac{1 - q_2 + 2q_2 \pi}{1 + q_2} \right).
\]

Consider a geometric-mean-preserving scrunch of \( \left[ \frac{1 - q_1 + 2q_1 \pi}{1 + q_1} \right] \cdot \cdots \cdot \left[ \frac{1 - q_n + 2q_n \pi}{1 + q_n} \right] \) obtained by decreasing \( q_1 \) and simultaneously increasing \( q_2 \). By definition, the scrunch preserves the average of the logs \( \frac{1}{n} \sum_{j=1}^{n} \ln \left( \frac{1 - q_j + 2q_j \pi}{1 + q_j} \right) \) as well as \( P_r(\mathcal{R}^i) \).

Since \( \ln \left( \frac{1 - q + 2q \pi}{1 + q} \right) \) is decreasing and concave in \( q_t \), the geometric-mean-preserving scrunch in which \( q_1 \) is decreased and \( q_2 \) is increased permits \( q_1 \) to fall by more than \( q_2 \) rises. Thus, the sum \( \sum_{k=1}^{J} q_k \) decreases. The implementation plan resulting from the scrunch provides the same transfer and discourages shirking exactly as effectively as before. But since it has reduced the expected number of workers hired, it results in a lower expected wage bill. The firm’s optimal implementation plan must then be of the form \( \hat{Q} = (q, ..., q) \) specifying a constant monitoring probability \( q \in [0, 1] \).

\( \square \)

**Lemma 4.** It is necessary and sufficient that the transfer dissuades shirking on all tasks. In other words, the incentive compatible wage is determined by \( \sigma^i \).

**Proof.** Let \( \sigma^i \) call for shirking on exactly \( J - n \) tasks (equivalent to exerting effort on \( n \) tasks). Implementation plan \( Q = (q, ..., q) \) implies \( P_r(\mathcal{R}^i) = \left[ \frac{1 - q + 2q \pi}{1 + q} \right]^{J-n} \).
The incentive compatibility constraint can now be written as
\[ W \geq \frac{c(J - n)}{1 - \left(\frac{1 - q + 2q\pi}{1 + q}\right)^{J-n}}. \]

Define \( \nu = \frac{1-q+2q\pi}{1+q} \) and \( x = J - n \), the constraint can be rewritten as \( W \geq c \frac{x}{1-\nu^x} \).

The right-hand side of this expression is increasing in \( x \) since the sign of \( \frac{\partial}{\partial x} \left( \frac{x}{1-\nu^x} \right) \) is given by \( 1 - \nu^x + \nu^x \ln(\nu^x) \), which is positive. (The expression \( 1 - \nu^x + \nu^x \ln(\nu^x) \) is decreasing in \( \nu \), but it is still positive as \( \nu^x \to 1 \) from below.)

The transfer must be large enough to dissuade shirking on all tasks, so
\[ W \geq \frac{Jc}{1 - \left(\frac{1 - q + 2q\pi}{1 + q}\right)^{J-n}}. \]

**Theorem 1.** The optimal contract will take the form of an implementation plan \( Q^i = (q, ..., q) \) that specifies hiring a second agent with probability \( q \) for each task and a transfer
\[ W^i(r^i) = \begin{cases} J \cdot w^*(q, \pi, c, J) & \forall r^i \in R^i \\ 0 & \forall r^i \not\in R^i, \end{cases} \]
where the equivalent per-task wage is \( w^*(q, \pi, c, J) = \frac{c}{1 - \left(\frac{1 - q + 2q\pi}{1 + q}\right)^J} \).

**Proof.** The firm will select the lowest transfer capable of inducing effort on all tasks. Lemma 3 and Lemma 4 establish properties of the optimal contract. For any implementation plan \( Q^i = (q, ..., q) \), the transfer scheme specifies \( W^i(r^i) = \frac{Jc}{1 - \left(\frac{1 - q + 2q\pi}{1 + q}\right)^{J-n}} \) for reports \( r^i \in R^i \) and \( W^i(r^i) = 0 \) for \( r^i \not\in R^i \). The lump-sum transfer is equivalent to a per-task wage of \( w(q, \pi, c, J) = \frac{c}{1 - \left(\frac{1 - q + 2q\pi}{1 + q}\right)^J} \).

**Proof of Corollary 1**
**Corollary 1.** \( w^*(q, \pi, c, J) \) is decreasing in \( q \) and \( J \) and increasing in \( \pi \) and \( c \).

**Proof.** The claims follow immediately from consideration of
\[ w^*(q, \pi, c, J) = \frac{c}{1 - \left(\frac{1 - q + 2q\pi}{1 + q}\right)^J}. \]
Proof of Corollary 2

Corollary 2. For effort costs \( c \in \left( (1 - \pi)^{1 - \left[ \frac{1 - q + 2q\pi}{1 + q} \right]^J}, 1 - \pi \right) \) the firm will not employ any workers despite it being efficient to do so were shirking not a concern.

Proof. The firm is willing to pay the equivalent per-task wage \( w^*(q, \pi, c, J) \) instead of hiring no worker only if

\[
1 - (1 + q) \frac{c}{1 - \left[ \frac{1 - q + 2q\pi}{1 + q} \right]^J} \geq \pi.
\]

The left-hand side gives the firm’s expected per-task payoffs from the proposed contract while the right-hand side is the firm’s expected per-task payoffs from guessing the state. This condition is equivalent to

\[
c \leq (1 - \pi) \frac{1 - \left[ \frac{1 - q + 2q\pi}{1 + q} \right]^J}{1 + q} < 1 - \pi.
\]

The relevant condition when effort is contractible is \( c \leq 1 - \pi \). Thus, for effort costs \( c \in \left( (1 - \pi)^{1 - \left[ \frac{1 - q + 2q\pi}{1 + q} \right]^J}, 1 - \pi \right) \) the firm will not employ any workers despite hiring workers when effort is contractible.

Proof of Theorem 2

Theorem 2. \( q^*(\pi, c, J) < 1 \) if and only if \( J > 1 \).

Proof. The firm’s expected payoff per task is

\[
1 - (1 + q) \left( \frac{c}{1 - \left[ \frac{1 - q + 2q\pi}{1 + q} \right]^J} \right).
\]

Let \( (\pi, c, J) \) be given and maximize this payoff with respect to \( q \). To simplify notation, let \( \nu = \frac{1 - q + 2q\pi}{1 + q} \) be the probability of producing acceptable output on any task when shirking. Note that \( \nu \in [\pi, 1] \) is decreasing over \( q = 0 \) to \( q = 1 \). The first-order condition stipulates

\[
c \cdot \left[ -2J(\pi - 1)\nu' + \nu' + (2\pi - 1)q(\nu' - 1) - 1 \right] = 0,
\]

where

\[
\nu' = \frac{\left( \frac{1 - q + 2q\pi}{1 + q} \right)'}{\left( \frac{1 - q + 2q\pi}{1 + q} \right)}.
\]
which requires

\[-2J(\pi - 1)\nu^J + \nu^J + (2\pi - 1)q(\nu^J - 1) - 1 = 0. \tag{A.1}\]

When $J = 1$, $\nu = \pi$ satisfies Equation A.1, so $q^*(\pi, c, 1) = 1$.

Suppose $J > 1$. When $q = 0$, the left-hand side of Equation A.1 is positive. It is negative when $q = 1$. To see this, note that the left-hand side reduces to $-2J\pi^J + 2\pi^J + 2\pi^{J+1} - 2\pi$ at $q = 1$. This expression is less than zero for $-J\pi^J + J\pi^{J-1} + \pi^J < 1$. The sum $-J\pi^J + J\pi^{J-1} + \pi^J$ is strictly increasing in $\pi$ and equals unity at $\pi = 1$, so it is less than unity for $\pi < 1$. Thus, the left-hand side of Equation A.1 is negative when $q = 1$ for any $\pi < 1$. Since Equation A.1 is continuous in $q$, it is satisfied with equality at some interior $q^*$ by application of the intermediate value theorem. The second-order condition assures $q^*$ identifies the unique maximum of the firm’s optimization problem.

\[\square\]

**Proof of Theorem 3**

**Theorem 3.** $q^*(\pi, c, J) \to 0$ as $J \to \infty$.

**Proof.** Let $(\pi, c)$ be given. The dependence of the monitoring probability on these parameters will often be suppressed. Denote $\nu = \frac{1-q+2q\pi}{1+q}$. Rearranging Equation A.1 in the proof of Theorem 2, the optimal monitoring probability $q$ must satisfy

\[\frac{J\nu^J}{1-\nu^J} - \frac{1-q+2q\pi}{2(1-\pi)} = 0. \tag{A.2}\]

Denote the left-hand side of Equation A.2 as $f(J, q(J))$ and treat $J$ as a continuous variable. (The firm’s problem is well-defined for $J \in \mathbb{R}_+$ and all relevant objects are monotonic in $J$.) The partial derivative of $f$ with respect to $J$ is $\nu^J\left(-\nu^J + J\ln(\nu) + 1\right)\left(\nu^J - 1\right)^2$. The sign of this derivative is governed by $1 - \nu^J + J\ln(\nu) \equiv 1 - x + \ln(x)$ where $x = \nu^J$. The sum $1 - x + \ln(x)$ is increasing over $x \in (0, 1]$, so it is increasing in $\nu^J$. Since $\lim_{x \to 1} 1 - x + \ln(x) = 0$, $1 - \nu^J + \ln(\nu^J)$ is negative. Thus, the partial derivative of $f$ with respect to $J$, $f_J$, is less than 0. The partial derivative of $f$ with respect to $q$, $f_q$, is also negative. Implicit differentiation holds $q_J = -\frac{f_J}{f_q}$, which shows the optimal monitoring probability is decreasing in the job size.
Since \( q^*(J) \equiv q^*(\pi, c, J) \) is decreasing in \( J \) and bounded below,

\[
\hat{q} = \lim_{J \to \infty} q^*(J)
\]

exists. To see that the limiting monitoring probability is zero, suppose otherwise that \( \hat{q} > 0 \). Define \( \nu^*(J) = \frac{1 - q^*(J) + 2q^*(J)}{1 + q^*(J)} \). Since \( \hat{q} \) exists, \( \hat{\nu} = \lim_{J \to \infty} \nu^*(J) \) exists.

Equation A.1 implies \(-2J(\pi - 1)\nu^*(J)^{J} + \nu^*(J)^{J} + (2\pi - 1)q^*(J)(\nu^*(J)^{J} - 1) = 1\) must hold at each \( J \). This requires \( \lim_{J \to \infty} J\nu^*(J)^{J} \) to converge to a finite value, implying \( \lim_{J \to \infty} \nu^*(J)^{J} = 0 \).

Then, for \( J \) sufficiently large, \( 2J(1 - \pi)\nu^*(J)^{J} - (2\pi - 1)q^*(J) \approx 1 \) and so

\[
\lim_{J \to \infty} J\nu^*(J)^{J} = \lim_{J \to \infty} \frac{1 + q^*(J)}{2(1 - \pi)} \nu^*(J)
\]

and

\[
\lim_{J \to \infty} J\nu^*(J)^{J-1} = \lim_{J \to \infty} \frac{1 + q^*(J)}{2(1 - \pi)}.
\]

Let \( \varepsilon > 0 \) be given. Since \( \nu^*(J) \) is convergent, for \( J \) large \((J - 1)\nu^*(J - 1)^{J-1} > (J - 1)\nu^*(J - 1)^{J-1} - \frac{\varepsilon}{2}\). Likewise, since \( J\nu^*(J)^{J} \) is convergent, it is Cauchy convergent and there exists \( J \) sufficiently large so that \((J - 1)\nu^*(J - 1)^{J-1} - J\nu^*(J)^{J} < \frac{\varepsilon}{2}\). Thus

\[
\frac{\varepsilon}{2} > (J - 1)\nu^*(J - 1)^{J-1} - J\nu^*(J)^{J} \\
> (J - 1)\nu^*(J)^{J-1} - \frac{\varepsilon}{2} - J\nu^*(J)^{J} \\
= \frac{1 + q^*(J)}{2(1 - \pi)} - \nu^*(J)^{J-1} - J\nu^*(J)^{J} - \frac{\varepsilon}{2} \\
= \frac{1 + q^*(J)}{2(1 - \pi)} - \nu^*(J)^{J-1} - \frac{1 + q^*(J)}{2(1 - \pi)}\nu^*(J) - \frac{\varepsilon}{2} \\
= \frac{1 + q^*(J)}{2(1 - \pi)}[1 - \nu^*(J)] - \nu^*(J)^{J-1} - \frac{\varepsilon}{2}.
\]

So for \( J \) sufficiently large, \( \varepsilon > \frac{1 + q^*(J)}{2(1 - \pi)}[1 - \nu^*(J)] - \nu^*(J)^{J-1} \). But since \( \nu^*(J)^{J} \to 0 \), this means \( 1 = \lim_{J \to \infty} \nu^*(J) = \lim_{J \to \infty} \frac{1 - q^*(J) + 2q^*(J)}{1 + q^*(J)} \). Recalling that \( \pi < 1 \) was given, this requires \( q^*(J) \to 0 \). This contradicts the maintained hypothesis that \( \hat{q} > 0 \) and the limiting monitoring probability is zero.

Therefore \( q^*(\pi, c, J) \to 0 \) as \( J \to \infty \).

\[ \square \]

Proof of Corollary 3
Corollary 3. For any \((\pi, c)\), given \(\varepsilon > 0\), \(\exists J_\varepsilon < \infty\) such that \(q^*(\pi, c, J) > \varepsilon\) for all \(J < J_\varepsilon\).

Proof. The result is immediate from noting the firm’s per-task wage bill as \(q \to 0\) holding fixed \(J\) grows without bound: \(\lim_{q \to 0} (1 + q) \cdot \frac{c}{1 - \frac{2q}{1+q}} \) does not exist. \(\Box\)

Proof of Result 1

Result 1. As \(J \to \infty\), \((1 + q^*(\pi, c, J)) \cdot w^*(q^*(\pi, c, J), \pi, c, J) \to c\).

Proof. The proof of Theorem 3 shows \(\left[\frac{1 - q^*(\pi, c, J) + 2q^*(\pi, c, J)}{1 + q^*(\pi, c, J)}\right]^J \to 0\) as \(J \to \infty\). Thus, \(w^*(q^*(\pi, c, J), \pi, c, J) \to c\) as \(J \to \infty\). Since \(q^*(\pi, c, J) \to 0\) as \(J \to \infty\), \((1 + q^*(\pi, c, J)) \cdot w^*(q^*(\pi, c, J), \pi, c, J) \to c\) as \(J \to \infty\). \(\Box\)

Result 3. The transfer scheme in the firm’s optimal contract without \(B\) is

\[
W^i(r^i) = \begin{cases} J \cdot w^*(q, \pi, c, J) & \forall r^i \in \mathcal{R}_{iJ}^i \\ 0 & \forall r^i \not\in \mathcal{R}_{iJ}^i, \end{cases}
\]

where the equivalent per-task wage is \(w^*(q, \pi, c, J) = \frac{c}{1 - \frac{2q}{1+q}}\).

Proof. Lemma 2 continues to apply for the same reasons as before. The general incentive constraint remains

\[
W_{00}\left[\overline{Pr}(\mathcal{R}_{00}^i) - Pr(\mathcal{R}_{00}^i)\right] + ... + W_{JJ}\left[\overline{Pr}(\mathcal{R}_{JJ}^i) - Pr(\mathcal{R}_{JJ}^i)\right] \geq c\left[J - \sum_{k=0}^{J} k \cdot \zeta(k|\sigma^i)\right].
\]

Since \(\overline{Pr}(\mathcal{R}_{jj}^i) - Pr(\mathcal{R}_{jj}^i)\) is largest at \(j = J\), inserting a wedge between \(\overline{\sigma}^i\) and \(\sigma^i\) is most effective here. A wage of \(W_{JJ}\) payable only for reports \(r^i \in \mathcal{R}_{JJ}^i\) is capable of inducing effort, where \(W_{JJ} = \frac{c\left[J - \sum_{k=0}^{J} k \cdot \zeta(k|\sigma^i)\right]}{Pr(\mathcal{R}_{jj}^i) - Pr(\mathcal{R}_{jj}^i)}\). Paying transfer \(W\) for other reports allows \(W_{JJ}\) to be lowered, but the expected wage bill increases nonetheless. \(\Box\)
Appendix B

Appendix to Chapter 2

Proof of Theorem 4

Theorem 4: In the lobbying game with endogenous bill writing and no discrimination, equilibrium expected revenue is $W(x) = v$ for $c > \frac{v}{2}$ and $W(x) = 2c$ for $c \leq \frac{v}{2}$, where $v$ is the valuation of policy $x$ such that $v_L(x) = v_R(x) \equiv v$.

For $c > \frac{v}{2}$, $x$ satisfies $v_L(x) = v_R(x)$. For $c \leq \frac{v}{2}$, the policy proposal is $x \in \{x' : W(x') = 2c\}$. The corporation successfully buys the lawmaker’s vote with probability $\frac{1}{2}$ in all equilibria.

Proof. This proofs builds heavily from the analysis of Che and Gale (1998). In particular, their lemmas 1, 2, and 3 continue to restrict equilibrium behavior when valuations are endogenous since the lemmas do not depend on $v_1 > v_2$. We solve for equilibrium bidding strategies and expected revenue in three cases: (1) $c \geq v$; (2) $c < \frac{v}{2}$, and; (3) $c \in [\frac{v}{2}, v]$ where $v$ is defined as the valuation of policy $x$ satisfying $v_L(x) = v_R(x)$. For expositional simplicity, the proof considers $v_i$ strictly monotonic; when valuation functions are weakly monotonic, the equilibrium policy proposal is the right-most policy satisfying the conditions below.

Suppose $c \geq v$ and consider a policy proposal $x'$ such that $v_R(x') > v_L(x')$. Since $v_i(\cdot)$ is monotonic, it must be that $v_R(x') > v > v_L(x')$. Since $c > \frac{v_L(x')}{2}$, expected revenue is given by $W(x') = \left[1 + \frac{v_L(x')}{v_R(x')}\right] \frac{v_L(x')}{2}$ (Che and Gale, 1998). For small changes, $W$ is increasing in $v_L$ and decreasing in $v_R$. This means a small increase in $x'$ generates greater expected revenue, so $x'$ such that $v_R(x') > v_L(x')$
cannot be an equilibrium policy proposal. The lawmaker proposes \( x \) such that \( v_R(x) = v_L(x) \) in equilibrium.

We now solve for equilibrium bidding behavior. Equilibrium bidding behavior must be in mixed-strategies. Suppose, to the contrary, that bids \((b_L, b_R)\) constitute a pure-strategy equilibrium. Let \( b_R > b_L \). If \( b_R < v \), \( L \) would benefit from deviating to \( \hat{b}_L = b_R + \varepsilon \). If \( b_R \geq v \), \( L \)'s best response is to bid zero. \( R \) would then benefit from deviating to \( \hat{b}_R = \varepsilon \). Thus, equilibrium bidding strategies cannot be pure.

Moreover, the strategies cannot place strictly positive mass on any positive bids. If \( R \) were placing positive mass on some bid \( b > 0 \), then \( L \) would benefit from increasing its bids in the interval \([b - \varepsilon, b]\) to \( b + \varepsilon \); there must be an interval of bids below \( b \) that \( L \) does not use. Then \( R \) would benefit from decreasing its bid from \( b \) to \( b - \varepsilon \). The distribution of bids is therefore continuous on its support. For similar reasons the interests groups must both have infimum bids of zero and supremum bids equal to \( v \). Finally, at most one group bids zero with positive probability; if both \( L \) and \( R \) bid zero with positive probability, \( R \) could increase its payoffs by moving its bid mass from zero to \( \varepsilon \).

Denote the equilibrium bidding distributions as \( F_L(\cdot) = F_R(\cdot) \equiv F(\cdot) \). Neither party bids zero with positive probability and the equilibrium distributions of bids are continuous on \([0, v]\). Since ties occur with zero probability, the payoffs to \( R \) from bidding \( b \) are \( vF_L(b) - b = vF(b) - b \). Suppose the groups obtain surplus \( \gamma > 0 \) from the auction. Since the bidders are randomizing over multiple bid amounts, the expected surplus must be \( \gamma \) for each bid. Then \( vF(b) - b = \gamma \), so \( F(b) = \frac{\gamma + b}{v} \). If \( \gamma \geq v \) and a contestant bids \( \varepsilon \), they expect to win spoils worth \( v - \varepsilon \) with probability \( \frac{\gamma + \varepsilon}{v} \). The expected surplus is worth strictly less than \( \gamma \), it must be that \( \gamma < v \). Now consider a bid \( v - \varepsilon \). The expected surplus is \( \frac{\gamma + v - \varepsilon}{v} \varepsilon < \frac{2v - \varepsilon}{v} \varepsilon < v\varepsilon \). For \( \varepsilon \) small enough, this surplus is less than \( \gamma \) and hence \( \gamma = 0 \). Thus, equilibrium bidding strategies satisfy \( F(b) = \frac{b}{v} \) when \( c > v \).

The expected bid of each contestant is \( \int_0^v \frac{b}{v} db = \frac{v}{2} \) and \( W = v \) in equilibrium.

Suppose \( c < \frac{v}{2} \) and consider policy \( x' \) such that \( v_L(x') = v_R(x') = v \). Fol-
lowing Che and Gale (1998), when $c \leq \frac{v}{2}$ both contestants bid $c$ with probability one. To see this, first note (for the same reasons as above) that equilibrium strategies cannot prescribe strictly positive mass for any bids between zero and the cap. (The argument used above does not apply for mass on $c$ in this case.) $L$ and $R$ cannot both bid zero with positive probability either for then it would be profitable to move the mass at zero to $\varepsilon$. Now suppose bidder $i$’s infimum bid is $b \in (0, c)$. Bidder $j$ must not be bidding in $(0, b)$, so $i$ could profitably move mass from $(b, b + \delta)$ to an arbitrarily small bid $\varepsilon$, contradicting the hypothesis that $i$’s infimum bid is $b$ in equilibrium. Now suppose $i$’s infimum bid is zero. Since $c < \frac{v}{2}$, bidding $c$ has expected payoff of at least $\frac{1}{2}v - c > 0$. In order for $i$ to be indifferent over bidding $c$ and placing a bid near zero, $j$ must have an atom at zero; if $j$’s equilibrium strategy was atomless at zero, then $i$’s expected payoff from bidding near enough to zero would be below the expected payoff from bidding $c$. Since $j$’s infimum must also be zero, $i$ must place an atom at zero for the same reason. However, with both contestants donating zero with strictly positive probability, it is profitable to move the atom to some arbitrarily small positive bid. Thus it cannot be the case that the infimum bid of either $L$ or $R$ is zero: both contestants must contribute $c$ to the lawmaker and $W(x') = 2c$.

Consider an alternative proposal so that $v_L(x) < v_R(x)$. The above argument holds for $v_i(x) \geq 2c$ with expected revenues remaining $W(x) = 2c$. Though revenue remains the same, the proposals offer different expected surpluses to the interest groups. Since the corporation attains the greatest surplus from proposal $x$ such that $v_L(x) = 2c$, this is the policy the lawmaker proposes. Each interest group wins the lawmaker’s vote half of the time but the corporation earns more from buying the vote than the labor union.

When $c \in [\frac{v}{2}, v)$, revenue is uniquely maximized by proposing $x$ such that $v_L(x) = v_R(x) = v$. To see this, first consider a policy $x'$ such that $2c \geq v_i(x')$ but $v_L(x') \neq v_R(x')$. Revenue is given by $W(x') = \left[ 1 + \frac{v_2(x')}{v_1(x')} \right] \frac{v_2(x')}{2}$, which is maximized when $v_L(x) = v_R(x) = v$ since $2c \geq v$.

With $c \leq v$, $c$ can no longer be the infimum bid: suppose $i$’s infimum is $c$. Then $j$ must be bidding either zero or $c$ since intermediate bids lose for sure.
Donating $c$ has expected payoff $\frac{1}{2}v - c < 0$ since $c > \frac{v}{2}$. If, on the other hand, $j$ was bidding zero exactly, then $c$ would no longer be an equilibrium bid for $i$. Thus, $j$ must be randomizing between zero and $c$. But since $\frac{1}{2}v - c < 0$, this cannot be optimal either: both contestants must have infimum contributions of zero.

Suppose $i$’s equilibrium strategies is atomless at zero. (We already know both contestants cannot have atoms at zero.) If $i$’s strategy is also atomless at $c$, $j$ would be unwilling to bid arbitrarily near zero since a bid of $c$ has expected payoff $v - c > 0$. So $i$ must place an atom at $c$ in order for $j$ to have infimum zero. Since $i$ has an atom at $c$, $j$ must place no mass just below $c$. If $j$’s strategy is atomless at $c$, then it would be profitable for $i$ to move the mass from $c$ to just below $c$. This cannot be an equilibrium either as $j$ would then have a profitable deviation to $c$. In equilibrium both contestants must contribute the maximum amount with strictly positive probability.

Since both contestants donate $c$ with strictly positive probability there must be an interval just below $c$ that neither contestant bids with nonzero probability. Denote this interval as $(\tilde{b}, c)$. The contestants must bid $b \in (0, \tilde{b}]$ with positive probability: it is clear that if one bidder has an interval of zero density in $(0, \tilde{b}]$ the other bidder must as well. But if both bidders have a (common) interval of zero density, then it would be profitable to move mass from just above the interval to the bottom of it. Repeating this exercise both contestants end up with zero density on $(0, \tilde{b}]$, which implies zero density is placed on $(0, c)$: the contestants must be randomizing between zero and $c$. Since there is no mixed-strategy equilibrium calling for mixtures over only zero and $c$ (each contestant would have to expect to win the vote with probability greater than one-half in order to bid $c$), it must be that bids $b \in (0, \tilde{b}]$ are made with nonzero probability.

The bidders are therefore indifferent over bids $b \in (0, \tilde{b}] \cup \{c\}$. Equilibrium requires the expected payoff from bidding $b \in (0, \tilde{b})$, $vF_i(b) - b$, to equal the expected payoff from bidding $c$, $v(F_i(\tilde{b}) + \frac{1 - F_i(\tilde{b})}{2}) - c$, for contestant $j$. In particular, $vF_i(\tilde{b}) - \tilde{b} = v(F_i(\tilde{b}) + \frac{1 - F_i(\tilde{b})}{2}) - c$. Thus, $F_i(\tilde{b})$ can be expressed as a function of $c, \tilde{b}$ and $v$. This, in turn, implies $F_i(b)$ can be expressed a function of $c, \tilde{b}, v$ and $b$ only. It follows that $F_i(b) = F_j(b) \equiv F(b)$ for all $b$. The equilibrium bid distribution $F(b)$
satisfies $F(b) = \frac{1}{2} + \frac{F(b)}{2} - \frac{c-b}{v}$. Since it cannot be that both contestants have atoms at zero, $F(0) = 0$. Then $0 = F(0) = \frac{1}{2} + \frac{F(\bar{b})}{2} - \frac{c-0}{v}$ and $F(\bar{b}) = \frac{2c}{v} - 1$. Plugging $F(\bar{b})$ back into the expression for $F(b)$, we have $F(b) = \frac{b}{v}$. Because $vF(b) - b = 0$ and the contestant is indifferent between bidding $b$ and $\bar{b}$, $0 = v(F(\bar{b}) + \frac{1-F(\bar{b})}{2}) - \bar{b}$, which implies $\bar{b} = 2c - v$. The equilibrium bid function satisfies

$$F(b) = \begin{cases} \frac{b}{v}, & \text{for } [0, 2c-v) \\ \frac{2c}{v} - 1, & \text{for } [2c-v, c) \\ 1, & \text{at } c \end{cases}$$

Each bidder bids $\frac{v}{2}$ in expectation. Therefore $W = v$ and the corporation successfully buys the lawmaker’s vote one-half of the time in the equilibrium.

Proof of Proposition 1

Proposition 1: In the absence of a cap on contributions, for any valuations $v_1(x)$ and $v_2(x) \leq v_1(x)$, expected contributions are maximized by providing the lower valuation group with advantage $\gamma(x) = v_1(x) - v_2(x)$. With $\gamma(x)$ set optimally, the expected contributions from a policy proposal $x$ are $\tilde{W}(x) = v_1(x) - \frac{v_2(x)}{2} + \frac{(v_2(x))^2}{2v_1(x)}$.

Proof. Pastine and Pastine (2009) calculate expected revenue from the all-pay auction when $v_1$, $v_2$, and $\gamma > 0$ are giving exogenously. Since advantage $\gamma > v_1$ leads to zero contributions, we may restrict attention to $\gamma \leq v_1$. For significant advantages, $\gamma \geq v_1 - v_2$, expected revenue is given by $W = \frac{v_2^2 - \gamma^2}{2v_2} + \frac{(v_1-\gamma)^2}{2v_1}$. As $\frac{\partial W}{\partial \gamma} < 0$ since $v_1 > \gamma$, expected contributions are maximized by setting $\gamma = v_1 - v_2$.

For smaller advantages, $\gamma \leq v_1 - v_2$, expected revenue is given by $W = \frac{v_2^2 + 2\gamma}{2} + \frac{v_2^2}{2v_1}$. This expression is again maximized at $\gamma = v_1 - v_2$.

Proof of Proposition 2

Proposition 2: Conditional on a contribution cap $c$ being binding, for a given policy proposal $x$, the lawmaker optimally sets $\gamma(x) = c - v_2(x)$. The expected contributions from policy $x$ with $\gamma(x)$ chosen optimally are bounded above by $\tilde{W}(x) = c + \frac{v_2(x)}{2} - \frac{c-v_2(x)}{v_1(x)} + \frac{(v_2(x))^2}{2v_1(x)}$. 

Proof. With cap $c$, valuations $v_1$ and $v_2 < v_1$, and advantage $\gamma > 0$ given exogenously, Pastine and Pastine (2009) establish that expected revenue is $W = \frac{c^2 - \gamma^2}{2v_2} + \frac{c - \gamma}{2v_1} (2v_1 - c - \gamma)$.

For the same reasons stated earlier, we can restrict attention to advantages $\gamma(x) < \min\{c, v_1(x)\}$. Now if $c > v_2(x) + \gamma(x)$, the cap is not binding. So let $\gamma(x) > c - v_2(x)$. Then, conditional on the cap being binding, the optimal advantage to give to the low valuation bidder is $\gamma(x) = c - v_2(x)$. Expected contributions with $\gamma(x)$ so chosen approach from below $\tilde{W}(x) = c + \frac{v_2(x)}{2} - \frac{c-v_2(x)}{v_1(x)} + \frac{(v_2(x))^2}{2v_1(x)}$. \qed
Bibliography


Caulfield, B. (2011). Amazon mechanical turk will hook you up with porn jobs that pay pennies. *Forbes*.


oDesk (2012). odesk at a glance.


