

Essays on Corruption and Enforcement

Alexander Robert Henke

A dissertation

submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2016

Reading Committee:

Fahad Khalil, Chair

Jacques Lawarrée, Chair

Quan Wen

Program Authorized to Offer Degree:

Economics

©Copyright 2016

Alexander Robert Henke

University of Washington

Abstract

Essays on Corruption and Enforcement

Alexander Henke

Chair of the Supervisory Committee:

Professor Fahad Khalil

Professor Jacques Lawarrée

Department of Economics

This dissertation examines the role and limits of enforcement in environments where the enforcing party faces issues related to corruption and misreporting. In the first chapter, we construct a principal – agent – auditor model with adverse selection in which the principal optimally allows bribery to occur due to the potential for extortion. This result mirrors the moral hazard model of Khalil, Lawarrée and Yun (2010). I introduce a probability that the agent is “honest” in that she cannot lie or engage in corruption. Because the principal cannot distinguish who is honest and who is not a priori, the existence of honest agents creates an additional dimension of adverse selection. Honest agents cannot reduce their expected penalties through bribery, and strategic agents can pretend to be honest, so the principal must either allow additional rent for all dishonest agents or shut down honest, low-income agents. This would avoid the new adverse selection issue, at the cost of revenue. In this way, honesty hurts the principal.

In the second chapter, I examine the optimality of commitment to an auditing scheme. Some auditors forgo committing to audit when they seemingly have the capability of doing so, such as the IRS. One explanation is that auditors tend to receive subjective information that may be too late to affect production and transfers but could affect auditing decisions, and contracting upon this information creates additional incentive issues which prevent the auditors from effectively committing to future actions. I construct a principal – agent model with costly auditing occurring after production and transfers, and I find that the principal may optimally forgo commitment to auditing when he observes a strong private signal of the agent’s type prior to auditing.

In the third chapter, we examine how the death penalty affects the incidence of reported child rape for different types of offenders. In 2008, the Supreme Court ruled capital punishment for child rape unconstitutional. We use this natural experiment to find evidence that the death penalty reduces reported child rapes for strangers and acquaintances. This evidence suggests that capital punishment deters potential offenders from committing child rape.

TABLE OF CONTENTS

	Page
List of Figures	iv
List of Tables	v
Chapter 1 Honest Agents in a Corrupt Equilibrium	1
1.1 Introduction	1
1.2 The Setup	6
1.3 Benchmark Contracts.....	9
1.3.1 Incorruptible Auditor.....	10
1.3.2 Corruptible Auditor.....	11
1.3.2.1 Bounty Hunter Contract.....	12
1.3.2.2 Least-cost Corruption-Proof Contract.....	13
1.3.2.3 Optimal Contract Allowing Corruption.....	15
1.4 Optimal Contracts when Agents may be Honest	17
1.4.1 LCCP Revisited.....	19
1.4.2 Allowing Corruption.....	22
1.4.2.1 Feasible Corruption Cases.....	22
1.4.2.2 BA contract.....	22
Proposition 1.1.....	27
1.4.2.3 Shut-down as an option.....	27
1.4.2.4 Shut-down contract.....	27
1.4.3 Comparing Contracts.....	28
1.4.3.1 Shut down of H1 is optimal for small q	28
Lemma 1.1.....	30
Proposition 1.2: Honesty hurts the principal.....	30
1.5 Conclusion	31
Chapter 2 Commitment to Auditing with a Private Signal.....	34
2.1 Introduction	34
2.2 The Model.....	37

2.2.1	Benchmark Contracts.....	39
2.2.1.1	Commitment to auditing, no signal.....	39
2.2.1.2	Commitment to auditing with a public signal.....	39
2.2.1.3	Auditing without commitment, with a private signal.....	40
	Lemma 2.1.....	41
2.2.2	Auditing with commitment and a private signal	42
	Lemma 2.2.....	43
	Lemma 2.3.....	43
2.2.2.1	Commitment to auditing with noncompliance.....	43
	Proposition 2.1: Forgoing commitment to auditing is optimal.....	45
2.3	Extensions	45
2.3.1	Productive effort, transfer-dependent penalties.....	45
2.3.2	The model.....	45
2.3.3	Benchmark contracts.....	45
2.3.3.1	Second best.....	45
2.3.3.2	Commitment to auditing, no signal.....	46
2.3.4	Auditing without commitment, with a private signal.....	46
2.3.5	Commitment to auditing with a private signal.....	47
2.4	Conclusion	49
Chapter 3 The Impact of Capital Punishment on Child Rape Offenses		55
3.1	Introduction	55
3.2	Background.....	56
3.2.1	Evolution of Capital Punishment for Child Rape	56
3.2.2	Theories and Empirical Studies of Capital Punishment	56
3.3	The Decision to Offend.....	59
3.4	Identification Strategy and Hypotheses	62
3.5	Data.....	63
3.6	Empirical Results.....	64

3.6.1 Empirical Framework	64
3.6.2 Estimates of the Deterrent Effect of the Capital Punishment on Child Rape.....	65
3.6.3 Alternative Specifications and Robustness Checks	65
3.7 Conclusion	66
Bibliography	76
Appendix.....	80
Appendix 1A.....	80
Appendix 1B.....	90
Appendix 1C.....	97
Appendix 1D.....	121
Appendix 2A.....	131
Appendix 2B.....	134
Appendix 2C.....	137

LIST OF FIGURES

Figure Number	Page
1.1 Optimal Contracts, Numerical Example.....	33
3.1 Histogram of Average Child Rape Rate by County-Year.....	69

LIST OF TABLES

Table Number	Page
3.1 Death Penalty Statutes for the Rape of a Child (before Kennedy v. Louisiana).....	68
3.2 Summary Statistics for Reported Child Rape	69
3.3 States and Counties in Sample.....	70
3.4 The Deterrent Effect on Child Rape	71
3.5 The Deterrent Effect on Child Rape including Interaction Term	71
3.6 The Deterrent Effect on Child Rape with Different Arrest Measure.....	72
3.7 The Deterrent Effect on Child Rape with Different Arrest Measure and Interaction Term ..	72
3.8 Robustness checks.....	73
3.9 Marginal Effect of Death Penalty by Arrest Rate in Table 5.....	74
3.10 Marginal Effect of Death Penalty by Arrest Rate in Table 7.....	75

DEDICATION

To my wife, Dr. Lin-chi Hsu.

Chapter 1

Honest Agents in a Corrupt Equilibrium

1.1 Introduction

Hiring an auditor to inspect an agent can provide significant advantages in incentives, but the possibility of corruption poses varied challenges. Empirical evidence suggests that workers tend to migrate away from corrupt countries¹, strongly suggesting that corruption makes them worse off, all else equal. Bribery payments may especially hurt poor migrants (Dincer and Gunalp 2012) and reduce resources devoted to social welfare programs (Gupta, Davoodi, and Alonso-Terme 2002). These results suggest that access to the benefits of a corrupt society, and the costs of engaging in it, are heterogeneous.

Auditing based schemes used to elicit truthful reporting are sensitive to the ability of the auditor to alter or hide evidence. North Korean refugee survey data suggests that where corruption is easy, relevant, and hard to detect, it is prevalent (Kim 2010). Additionally, the optimality of such schemes depends crucially on how often corruption would actually occur without any countermeasures. Consider two organizations: One where both auditors and agents are incapable of engaging in side contracts, and one where both were completely corruptible. In the incorruptible organization, the principal need not impose costly restrictions to deter bribery. In the corrupt organization, the principal may either incur significant cost deterring all corruption or possibly even allow corruption to occur².

¹ See (Dimant, Krieger, and Meierrieks 2013), (Cooray and Schneider 2016), and (De Has 2007) for recent evidence

² Kofman and Lawarrée (1993; 1996), Che (1995), Mookherjee and Png (1995), Strausz (1997), Olsen and Torsvik (1998), Lambert-Mogiliansky (1998), Khalil and Lawarrée (2006), Acemoglu and Verdier (2000), and Auriol (2006) for some examples on optimally allowing corruption. Additionally, this paper bases its information structure, and the ensuing rationale for allowing corruption, on the work of Khalil, Lawarrée and Yun (2010).

In our construction, “honest” agents can neither lie nor engage in corrupt behavior; their only decision is whether to accept or reject the principal’s contract. “Strategic” agents are unrestricted in their behavior. A corruption-free organization contains honest agents, and a corrupt organization contains strategic agents. Clearly, the principal would prefer the honest organization to the strategic one. In spite of this, increasing the proportion of honest of agents in a primarily-corrupt agent pool hurts the principal, because the principal can’t effectively distinguish between honest and strategic agents when offering a contract.

The literature tends to view prosocial behavior such as incorruptibility as weakly beneficial to the principal. Mittendorf (2008) demonstrates that ethical behavior has spillover effects on unethical agents in an environment where relative performance evaluation is optimal but agent collusion is a potential threat. Kofman and Lawarrée (1996) find that the principal may benefit from allowing some collusion, if an auditor is incorruptible with enough likelihood. In their model, incorruptibility never hurts the principal’s profits, and eventually improves them. Importantly, the agent’s information about the possibility of collusion is symmetric with the principal, which means she is unsure of the possibility of corruption when signing the contract. So while the principal cannot screen between incorruptible and corruptible auditors, the agent will accept a contract which offers a payoff smaller than her outside option upon the discovery that she cannot engage in bribery, as long as her expected payoff is equal to her outside option *ex ante*.

If instead all auditors are corruptible, but some agents are honest, the agent clearly knows whether she can engage in bribery prior to signing of the contract. This creates an additional dimension of asymmetric information which is difficult for the principal to screen, as the only contractible signal related to honesty in any way is the auditor’s report. If the principal compensates the honest agent for her inability to extract rent via bribery, strategic agents will claim additional rent by potentially mimicking the honest agent. If only a small percentage of agents are honest, the principal will shut them down to reduce

the strategic agents' rent; this means that the principal's profits decrease as honesty, and hence shutdown, becomes increasingly common.

Multidimensional screening has a broad literature in pricing and auctions³, but its treatment in the case of prosocial behavior is relatively new. Bénabou and Tirole (2006) develop a model with multidimensional screening and prosocial behavior, focusing on reputational, extrinsic, and intrinsic motivations to perform a good action. Severinov and Deneckere (2006) develop a “password mechanism” to screen agents with message space restrictions, taking advantage of the agent's ability to repeatedly report a willingness to pay. A key difference with our main model is the assumption that message space is not limited to one dimension of type; honest agents can claim to be honest, but so can strategic agents.

This model uses the information structure developed by Khalil, Lawarrée and Yun (2010), hereafter referred to as KLY, to provide the basis for allowing corruption in equilibrium. In their moral hazard model, a supervisor inspects the agent's behavior after production has occurred. The supervisor obtains either hard evidence⁴ of the agent's effort, or finds nothing conclusive. If the supervisor wishes to alter his report, he can unilaterally hide evidence, or he can negotiate an illegal side contract with the agent and collaboratively alter the signal to any value. If the principal wishes to deter all corruption, he must deter both bribery – i.e. the supervisor-agent coalition altering the signal to improve both of their payoffs – and extortion – i.e. the supervisor demanding a payment simply to report what he found, using a credible threat to hide evidence. This creates non-separabilities in the constraints that deter corruption (Tirole, 1992) which make it optimal to allow bribery for an accurate-enough supervisor. Extortion, however, is never optimal, as it punishes the agent for good behavior.

³ Rochet and Choné (1998), Figalli et al (2011), and Manelli and Vincent (2007) represent a more technical portion of the literature.

⁴ KLY provides a thorough discussion of hard versus soft evidence.

We use the information framework developed by KLY in a simple adverse selection model and find similar results when all agents are strategic. We then include the probability that the agent is honest and develop a key contribution: Honesty hurts the principal by inducing shutdown of the low-productivity honest type in the optimal contract. When the principal faces a primarily corrupt agent pool, policies to develop honesty in some agents, or to cut their ties with corrupt auditors, may backfire by forcing these honest agents to shut down their projects.

Yun (2012) creates a regulator-inspector-firm model based on Mookherjee and Png (1995) to examine optimal contracting when falsifying reports and side-contracting are possible but costly. He finds that forms of corruption can exist in equilibrium, including extortion and framing, as even a falsified report is informative when falsification is costly. Additionally, increasing the cost of this corrupt behavior, including both information distortion and side-contracting, can harm social welfare. Yun emphasizes the resources wasted in engaging in corrupt behavior; even if corruption is reduced by making these behaviors more costly, more resources end up being devoted to corrupt behavior overall. He concludes that a fight against corruption ideally increases the cost of corruption high enough that it is deterred entirely. This paper's insights diverge from Yun's in the following ways: Limitations on corruption can be harmful even when they entail no direct waste of resources. Even a cost of corruption high enough to deter all corruption, i.e. honesty, can be harmful if it is heterogeneous and creates a new dimension of asymmetric information. If a policy increases the population of incorruptible agents, i.e. increasing the probability the agent is honest, as opposed to increasing the cost of corruption for everyone, and the auditing technology is relatively strong, corruption will not be deterred in equilibrium.

Ahlin and Bose (2007) develop a model where a productive agent must gain approval from a partially-honest bureaucracy to proceed with a project of varying productivity. Honest bureaucrats examine and approve worthwhile projects, while corrupt bureaucrats demand a bribe of some value for approval. If rejected, the agents can try again

in a second round, at the cost of a time discount. As the proportion of honest bureaucrats increases, a more productive agent's willingness to pay a bribe in the first round decreases, as she is more likely to encounter an honest bureaucrat and be approved without cost if she waits. However, unproductive agents remain willing to pay a high bribe, as they will only ever be approved by corrupt bureaucrats. For the bureaucrat to obtain a bribe from any agent he encounters, he must offer the lower bribe demand. At a certain threshold of honesty, corrupt bureaucrats will demand a large bribe that only unproductive agents will accept, creating costly delay and shutdown in some extensions for productive agents. In their model, bureaucratic honesty is non-monotonic in social welfare because it alters the agent's endogenous outside option and causes her essentially to pay a search cost for a better bureaucrat. If no bureaucrat is honest, the productive agent will simply pay the high bribe and complete her project.

Ahlin and Bose's work creates complementary insights while remaining highly distinct in construction and interpretation. While honest bureaucrats have some use in an informative capacity, corrupt bureaucrats are entirely useless; their entire function is to demand harassment bribes, which would be interpreted as extortion payments in the corruptible auditing literature. This divergence exists, in part, due to Ahlin and Bose's focus on corruption in allocation as opposed to corruption in inspection. A corrupt bureaucrat allows the agent to perform a task, while a corrupt auditor can help alter information and hence change incentives post-production. Importantly, the problem with honesty in Ahlin and Bose's model is not the existence of honest bureaucrats in general, but rather the anticipation of honest bureaucrats in a second round of interaction due to the agent's dynamic outside option.

The rest of the paper is organized as follows: Section 1.2 describes the model setup. Section 1.3 describes benchmark contracts and a replication of KLY's main results without honesty. Section 1.4 describes optimal contracts under various parameter conditions, and the main results. Section 1.5 concludes.

1.2 The Setup

A principal (it) contracts with an agent (she) to form a productive relationship. The agent incurs an investment $c > 0$ and privately receives θ as productive income. The principal then “taxes” a portion of the income t from the agent. One interpretation is that the principal is an investor, the agent is an entrepreneur, θ is the income generated by the entrepreneur’s productive project, c is the entrepreneur’s cost of initial investment, and t represents the investor’s share of the project’s gains.⁵ Limited liability implies that $t \leq \theta$. The agent can reject the relationship and collect an outside option normalized to 0.

We have a standard adverse selection framework where the productivity parameter θ can take two values, θ_1 (low type) or θ_2 (high type), known privately to the agent before signing the contract. Note that $\theta_2 - \theta_1 = \Delta\theta > 0$. It is common knowledge that the probability that $\theta = \theta_1$ is f_1 , and the probability that $\theta = \theta_2$ is $f_2 = 1 - f_1$. In the first best case, when the principal observes θ , it simply extracts all the agent’s income less the investment cost, $t_i = \theta_i - c$, for $i = 1, 2$. When the principal does not observe θ , there is no screening possible and it must offer a pooling contract, setting $t_1 = t_2 = \theta_1 - c$, with the high type obtaining an information rent, denoted by $u_2 = \Delta\theta$. We refer to this as the second best contract.⁶ Next, we introduce auditing. In order to focus on the issues of corruption of the auditor, we model auditing as the only screening device for the principal.

The principal can hire a costless, corruptible, risk neutral auditor (he) to collect a signal of the agent’s type. Following KLY (2010), we assume that the signal shows the

⁵ Alternative interpretations of c include: A reduced form cost of effort; a fixed non-pecuniary penalty reimbursed by the principal in some circumstances; an outside option, making it an opportunity cost of contracting with the principal instead of an accounting cost. An alternative model that eliminates c could generate similar results as long as it allowed the principal to enact some form of punishment when the agent is neither caught nor exonerated.

⁶ As is typically the case in this kind of a setting, a low type does not have an incentive to claim to be a high type.

agent's correct type (1 or 2) with probability p and shows no information (\emptyset) with probability $1 - p$. The auditor can freely hide information, i.e. change any signal into \emptyset , but he requires the agent's help to change the signal to 1 or 2.⁷ The auditor then submits his potentially-modified report to the principal. Given the combined reports i of the agent and j of the auditor, the transfer from the agent to the principal is denoted by $t_{i,j}$ and the payment from the principal to the auditor by $w_{i,j}$. For example, $t_{1,\emptyset}$ represents the transfer from the agent to the principal when the agent reports $\theta = \theta_1$, and the auditor reports no evidence. The auditor also has limited liability such that $w \geq 0$, and his reservation utility is zero.

A key distinction between bribery and extortion is whether the corrupt behavior benefits (the case of bribery) or hurts the agent (the case of extortion). We define bribery and extortion as follows:

Definition 1. *Bribery* occurs when one party accepts a payment in return for misreporting information in favor of the other party.

Definition 2. *Extortion* occurs when the supervisor obtains a payment from the agent by threatening to misreport evidence that was favorable to the agent.

Definition 3. *Framing* occurs if the attempt at extortion fails and the supervisor misreports information that was favorable to the agent.

We use the generic term corruption to describe bribery, extortion and framing. We assume that with probability $q > 0$ the agent is *honest*, in that she reveals her type truthfully and does not engage in a side contract with the auditor. With probability $(1 - q) > 0$, the agent is *strategic* in that she can misrepresent her type as well as engage

⁷ If the auditor could change the information by herself, his report would be useless in this model.

in a side-contract (unless given the right incentives). Therefore, our agent can be one of four different types depending on her productivity (θ) and her honesty: $\{H1, S1, H2, S2\}$ honest low and high productivity and strategic low and high productivity, respectively.

The timing of the game is as follows:

1. Nature determines the productivity type and honesty type. They are privately known to the agent.
2. The principal offers a take-it-or-leave-it contract to the agent representing a menu of transfers $\{t_{ij}, w_{ij}\}$, where i represents agent's report and j the auditor's report about the agent's productivity type. The agent and auditor accept or reject the contract.
3. The agent reports her type, i.e. the productivity parameter and her honesty.
4. The auditor observes a signal about the agent's productivity type.
5. The auditor and agent potentially enter into an illicit side-contract, make transfers, and alter the auditor's report. We assume that the auditor and agent bargains under symmetric information, i.e., the agent knows the auditor's signal and the auditor knows the agent's honesty.⁸
6. After receiving the auditor's report, the principal collects t and pays w .

We make the following parameter assumptions:

$$(A1): \frac{f_1}{1-f_1} > \frac{\Delta\theta}{\theta_1 - c}$$

$$(A2): p \leq 1 - \frac{c}{\Delta\theta}$$

⁸ In this model, whenever the auditor lacks hard information, he is useless. This also applies to the case where the auditor is allowed to report whether the agent is honest or strategic. Given a lack of hard information, the auditor can threaten to report whatever is worse for the agent and extract concessions whether the agent lied or told the truth.

$$(A3): p > \frac{f_1}{1 - f_1}$$

(A1) states that the principal does not find it profitable to shut down the low type in the absence of auditing. This assumption is primarily to confirm that our result of shutting down honest low types is not simply because the principal prefers to shut down all low types; it also helps simplify further the analysis by ensuring that partial shutdown only occurs when the probability of honesty is small.

(A2) ensures that the high type's Incentive Compatibility constraints will always bind. This condition is stronger than necessary to make sure that the principal will not achieve the first best when deterring all corruption, but its imposition simplifies the analysis.

(A3) implies that, when $q = 0$, allowing bribery is superior to deterring all corruption.

1.3 Benchmark Auditing Contracts

We first examine the case where the agent is strategic with probability 1 ($q = 0$). This is the adverse selection analog of KLY, which is a moral hazard model. An adverse selection model is more apt in presenting our key insights, which we do by showing the benefit of shutting down some types (honest low types). However, the results of this section with $q = 0$ closely mirror those of KLY, with some minor differences that are detailed in Appendix 1A.

1.3.1. Incorruptible auditor with $q = 0$

If the auditor is incorruptible, he requires no transfer to report the original signal. Auditing only takes place after the agent reported low productivity. We assume limited liability on the agent so that the maximum tax the principal can inflict upon a report of no evidence by the auditor is to collect the entire income reported by the agent. For instance, the principal is not allowed to collect more than θ_1 unless it has proof from the auditor of a larger amount of realized income. When the auditor makes a report of high productivity about an agent who claimed to be a low type, the principal can collect the entire verified income of the project, i.e., $t_{1,2} = \theta_2$. Beside the limited liability constraints – labeled *LLC* – the principal must impose the usual *IR* and *IC* constraints. The maximization problem is

$$\max_{t_{i,j}} f_1(pt_{1,1} + (1-p)t_{1,\emptyset}) + (1-f_1)(t_2) \text{ s. t.}$$

$$IR_1: \theta_1 - c - pt_{1,1} - (1-p)t_{1,\emptyset} \geq 0$$

$$IR_2: \theta_2 - c - t_2 \geq 0$$

$$IC_2: \theta_2 - c - t_2 \geq \theta_2 - c - pt_{1,2} - (1-p)t_{1,\emptyset}$$

$$LLC: t_{1,\emptyset} \leq \theta_1, t_{1,2} \leq \theta_2, t_{1,1} \leq \theta_1, t_2 \leq \theta_2$$

By (A1) and (A2), the principal does not find it profitable to shut down the low type, and IR_2 is slack.⁹ The optimal contract is:

$$t_{1,\emptyset} = \theta_1$$

⁹This contract offers less rent than the second best, and hence has less reason for shutdown; additionally, notice that a binding (A2) parameter constraint sets $u_2 = 0$.

$$t_{1,2} = \theta_2$$

$$t_{1,1} = \theta_1 - \frac{c}{p}$$

$$t_2 = \theta_2 - \Delta\theta(1 - p)$$

$$u_2 = \Delta\theta(1 - p) - c,$$

where u_2 is the rent of the high productivity agent.

The low type receives no rent and is compensated for c when the auditor confirms her report. The rent of the high type decreases with the quality of auditing (p) and with c , but the high type's transfer does not change with c . So in short, c helps incentives, but hurts overall profits as the principal must compensate the agent for her cost c .

1.3.2 Corruptible auditor with $q = 0$

Consider now the case where the auditor's report is manipulable. If the principal wants to deter bribery and extortion, the optimal contract must satisfy additional constraints. To deter bribery, the contract must satisfy Coalition Incentive Compatibility (*CIC*) constraints that remove the agent's incentive to bribe the auditor to alter the signal. Defining $T_{1,j} = t_{1,j} - w_{1,j}, j \in \{1,2, \emptyset\}$, the *CIC* are:

$$CIC_{1,\emptyset}: T_{1,1} \geq T_{1,\emptyset}$$

$$CIC_{\emptyset,1}: T_{1,\emptyset} \geq T_{1,1}$$

$$CIC_{1,2}: T_{1,1} \geq T_{1,2}$$

$$CIC_{2,1}: T_{1,2} \geq T_{1,1}$$

$$CIC_{\emptyset,2}: T_{1,\emptyset} \geq T_{1,2}$$

$$CIC_{2,\emptyset}: T_{1,2} \geq T_{1,\emptyset}.$$

$CIC_{1,\emptyset}$ and $CIC_{\emptyset,1}$ jointly imply that $T_{1,1} = T_{1,\emptyset}$. $CIC_{1,2}$ and $CIC_{2,1}$ jointly imply that $T_{1,1} = T_{1,2}$. Thus the *CIC* conditions reduce to:

$$(CIC) \quad T_{1,1} = T_{1,2} = T_{1,\emptyset}$$

To deter extortion, the contract must satisfy No Extortion (*NE*) constraints that remove the auditor's incentive to extort the agent by hiding evidence:

$$NE_2: w_{1,2} \geq w_{1,\emptyset}$$

$$NE_1: w_{1,1} \geq w_{1,\emptyset}$$

Note that these constraints also deter framing.

1.3.2.1 Bounty hunter contract ($q = 0$)

We begin our analysis by briefly looking at the well-studied case, where the auditor can engage in bribery but not in extortion or framing. This would be the case if the auditor cannot hide evidence unilaterally. In this case, the NE_i constraints are not relevant. Then, as is well known (see e.g. Tirole (1986) or Kofman-Lawarree (1993)), the principal would turn the auditor into a bounty hunter and deter bribery. The agent has an incentive to bribe the auditor after a null signal and after a contradictory signal following her report of θ_1 . The principal optimally deters these incentives by setting the reward equal to the stake of collusion: $w_{1,\emptyset} = t_{1,\emptyset} - t_{1,1} = \frac{c}{p}$, and $w_{1,2} = t_{1,2} - t_{1,\emptyset} = \theta_2 - \theta_1 + \frac{c}{p}$. Compared to the incorruptible auditor case, the high type agent's rent remains unchanged at $u_2 = \Delta\theta(1 - p) - c$, but the principal has to pay the auditor a reward $w_{1,\emptyset} = \frac{c}{p}$ in equilibrium,

which is the cost of deterring bribery. Notice that the bounty hunter contract violates NE_1 , since it sets $w_{1,\emptyset} = \frac{c}{p} > w_{1,1} = 0$. Therefore, if extortion or framing is possible, the constraints NE cannot be ignored.

1.3.2.2 Least-Cost Corruption Proof contract ($q = 0$)

Consider next the case where the principal deters both bribery and extortion. This contract must satisfy both CIC and NE constraints. We call the optimal contract that deters both bribery and extortion the Least-Cost Corruption Proof ($LCCP$) contract. As we explain below, it is not necessarily optimal to deter both bribery and extortion but the ($LCCP$) contract is a useful benchmark.

The principal maximizes

$$f_1 \left(p(t_{1,1} - w_{1,1}) + (1 - p)(t_{1,\emptyset} - w_{1,\emptyset}) \right) + (1 - f_1)t_2$$

Subject to the following constraints:

$$IR_1: \theta_1 - c - pt_{1,1} - (1 - p)t_{1,\emptyset} \geq 0$$

$$IR_2: \theta_2 - c - t_2 \geq 0$$

$$IC_2: \theta_2 - c - t_2 \geq \theta_2 - c - pt_{1,2} - (1 - p)t_{1,\emptyset},$$

and the $CICs$, NEs , and $LLCs$.

The *LCCP* contract is:¹⁰

$$t_2 = \theta_2 - \Delta\theta(1-p) - (1-p)c$$

$$t_{1,1} = t_{1,\emptyset} = t_1 = \theta_1 - c$$

$$w_{1,1} = w_{1,\emptyset} = w_1 = 0$$

$$t_{1,2} = \theta_2$$

$$w_{1,2} = \Delta\theta + c$$

$$u_2 = \Delta\theta(1-p) - pc.$$

To deter extortion, the principal must pay the auditor a zero wage whether he reports \emptyset or 1 ($w_{1,1} = w_{1,\emptyset} = 0$), but then the *CIC* imply that the transfers after the two signals must also be the same, $t_{1,1} = t_{1,\emptyset}$. In other words, the threat of extortion restricts the ability of the principal to use the auditor's information. The principal can no longer differentiate between the reports (1,1) and (1, \emptyset). The low productivity agent still receives no rent, and the rent of the high productivity agent is higher than without extortion ($u_2 = \Delta\theta(1-p) - pc$). It is only if the auditor reports (2) that his wage is positive but this wage is off the equilibrium path, so the auditor receives no rent in equilibrium. However,

¹⁰ Note that the *LCCP* solution is valid even for a higher p that allowed by (A2). The constraint that prevents the first best from being possible is $p < \frac{\Delta\theta}{\Delta\theta+c} = 1 - \frac{c}{\Delta\theta+c}$, which is strictly weaker than (A2), as $1 - \frac{c}{\Delta\theta+c} > 1 - \frac{c}{\Delta\theta}$.

detering both bribery and extortion is not necessarily optimal in this setting. We show next that it may be optimal to allow bribery, but extortion is never tolerated.

1.3.2.3 Optimal Contract Allowing Corruption ($q = 0$)

Before analyzing a possible optimal contract allowing corruption, we need to clarify how the agent and the auditor negotiate the bribe. First, we assume symmetric information during the bribe negotiation, i.e., the agent and auditor know the agent's report and the auditor's signal. We use the Nash bargaining solution to determine the equilibrium bribe. The agent's bargaining power is represented by λ , with $0 < \lambda < 1$. Given the agent's report of type- i , the net transfer from the auditor-agent coalition to the principal given an auditor's report of k is then $T_{i,k} = t_{i,k} - w_{i,k}$. Hence, the net gain for the coalition of altering the auditor's report from signal j to k is $(t_{i,j} - w_{i,j}) - (t_{i,k} - w_{i,k})$. If the coalition decides to alter the report from j to k , the agent pays $t_{i,k}$ to the principal and $(1 - \lambda)(t_{i,j} - t_{i,k}) - \lambda(w_{i,k} - w_{i,j})$ to the auditor as a bribe. We denote the total transfer paid by the agent as $t'_{i,j} = t_{i,j} + \lambda((t_{i,k} - w_{i,k}) - (t_{i,j} - w_{i,j})) = (1 - \lambda)t_{i,j} + \lambda(t_{i,k} - (w_{i,k} - w_{i,j}))$.

The agent-auditor coalition will choose the signal that entails the minimum net transfer to the principal, described as $T_M = \min(t_{1,1} - w_{1,1}; t_{1,\emptyset} - w_{1,\emptyset}; t_{1,2} - w_{1,2})$. If the unadulterated signal is already the most profitable signal for the coalition, $t'_{i,j} = t_{i,j}$, then no side contract is necessary. We prove in Appendix 1A that extortion will be deterred and the only form of bribery that may be allowed is for the coalition to report 1 after the auditor receives a null signal.

The principal's maximization problem is:

$$\max_{t_{i,j}, w_{i,j}} f_1(T_M) + (1 - f_1)(t_2)$$

Subject to:

$$IR_1: \theta_1 - c - p \left(t_{1,1} + \lambda(T_M - t_{1,1} + w_{1,1}) \right) - (1-p) \left(t_{1,\emptyset} + \lambda(T_M - t_{1,\emptyset} + w_{1,\emptyset}) \right) \geq 0$$

$$IR_2: \theta_2 - c - t_2 \geq 0$$

$$\begin{aligned} IC_2: \theta_2 - c - t_2 & \\ & \geq \theta_2 - c - p \left(t_{1,2} + \lambda(T_M - t_{1,2} + w_{1,2}) \right) \\ & \quad - (1-p) \left(t_{1,\emptyset} + \lambda(T_M - t_{1,\emptyset} + w_{1,\emptyset}) \right) \end{aligned}$$

$$NE_1: w_{1,1} \geq w_{1,\emptyset},$$

And non-negativity constraints, and where $T_M = \min(t_{1,1} - w_{1,1}; t_{1,\emptyset} - w_{1,\emptyset}; t_{1,2} - w_{1,2})$.

The derivation of the contract is given in Appendix 1A. In particular, it is shown that the coalition will report 1 when the signal shows no evidence. The optimal contract is as follows:

$$w_{1,1} = w_{1,\emptyset} = 0$$

$$t_{1,2} = \theta_2$$

$$w_{1,2} = \theta_2 - \theta_1 + \frac{c}{p + (1-p)\lambda}$$

$$t_{1,1} = \theta_1 - \frac{c}{p + (1-p)\lambda}$$

$$t_{1,\emptyset} = \theta_1$$

$$t_2 = p\theta_2 + (1-p)\theta_1 - \frac{(1-p)\lambda}{p + (1-p)\lambda} c$$

$$u_2 = (1-p)\Delta\theta - \frac{p}{p + (1-p)\lambda} c.$$

This contract is superior to the *LCCP* contract when $p > \frac{f_1}{1-f_1}$, which is assumed in (A3). When allowing bribery, the principal can once again differentiate between the reports (1,1) and (1, \emptyset) by making $t_{1,1} < t_{1,\emptyset}$, triggering bribery. The trade-off presented by allowing bribery is clearest when considering different values of λ . If $\lambda = 0$, the agent has no bargaining power. When engaging in a collusive agreement with the auditor, she must provide the entire gains from collusion as a bribe. This is equivalent to the bounty hunter case. Alternatively, if the agent has all the bargaining power, i.e. $\lambda = 1$, then the agent retains all the gains from collusion, and there is effectively no bribe. The contract resembles the *LCCP* contract. Notice that λ affects the magnitude of difference between allowing and deterring bribery, but not the optimality.

1.4 Optimal Contracts when Agents may be Honest ($q > 0$)

The introduction of honest (H) and strategic (S) types creates a new dimension of asymmetric information. There are then four types, $\{H1, H2, S1, S2\}$. For example, S2 represents the strategic high type, whereas H1 is the honest low type. Transfers and constraints now refer to the agent's type report $\{S2, H2, S1, H1\}$ and the auditor's cost report of $j \in \{1, 2, \emptyset\}$. For example, $t_{S1, \emptyset}$ refers to the transfer when an agent reports S1 and the auditor reports no evidence.

The *IR* constraints are now written as IR_i , where $i \in \{H1, H2, S1, S2\}$. *IC* constraints are presented as $IC_{i \rightarrow i'}$ where $i, i' \in \{H1, H2, S1, S2\}$.

Note that auditing will only occur if $H1$ or $S1$ is reported and the transfers after a high productivity report are given by t_{H2} and t_{S2} .

Recall that $t'_{i,j}$ signifies the total transfers paid by the agent in the event that the agent reports a particular type i and the auditor *finds*, but does not necessarily report, a signal j . In the case of strategic agents, the existence and nature of collusive side-transfers depend on the implementation of *CIC* constraints and the minimum agent-auditor coalition transfer for a particular type report, which is now defined as $T_{M,i}$, $i \in \{H1, S1, H2, S2\}$, but otherwise retains its characteristics as described in the previous section.

In the case of honest agents, there is no misreporting or collusive side agreement. However, framing is possible since an honest type will refuse to engage in a side payment. Defining $T_{i,j,F}$ as the net realized transfer granted to the principal after framing has potentially occurred given some agent report i and auditor signal j , $t_{i,j,F}$ as the transfer from the agent to the principal given no collaboration and potential framing, and $t'_{i,j}$ as the actual net transfer paid by a strategic agent given some agent report i and auditor signal j .

Therefore, we define:

$$T_{H1,j} = t_{H1,j} - w_{H1,j}$$

$$T_{S1,j} = t_{S1,j} - w_{S1,j}$$

$$T_{H1,j,F} = t_{H1,j,F} - w_{H1,j,F}$$

$$T_{S1,j,F} = t_{S1,j,F} - w_{S1,j,F}$$

$$T_{H1,M} = \min(T_{H1,1}, T_{H1,2}, T_{H1,\emptyset})$$

$$T_{S1,M} = \min(T_{S1,1}, T_{S1,2}, T_{S1,\emptyset}).$$

In the case where extortion and framing do not occur,

$$t'_{H1,j} = t_{H1,j} + \lambda(T_{H1,M} - T_{H1,j})$$

$$t'_{S1,j} = t_{S1,j} + \lambda(T_{S1,M} - T_{S1,j}).$$

In the case where extortion and framing occur,

$$t'_{H1,j} = t_{H1,\emptyset} + \lambda(T_{H1,M} - T_{H1,\emptyset})$$

$$t'_{S1,j} = t_{S1,\emptyset} + \lambda(T_{S1,M} - T_{S1,\emptyset}).$$

We can state the principal's objective function:

$$f_1 \left(q(pT_{H1,1,F} + (1-p)T_{H1,\emptyset}) + (1-q)(T_{S1,M}) \right) + (1-f_1)(qt_{H2} + (1-q)t_{S2})$$

And the potential constraints are

$$IR_{H1}: \theta_1 - c - pt_{H1,1,F} - (1-p)t_{H1,\emptyset} \geq 0$$

$$IR_{H2}: \theta_2 - c - t_{H2} \geq 0$$

$$IR_{S1}: \theta_1 - c - pt'_{S1,1} - (1-p)t'_{S1,\emptyset} \geq 0$$

$$IR_{S2}: \theta_2 - c - t_{S2} \geq 0$$

$$IC_{S1 \rightarrow H1}: \theta_1 - c - pt'_{S1,1} - (1-p)t'_{S1,\emptyset} \geq \theta_1 - c - pt'_{H1,1} - (1-p)t'_{H1,\emptyset}$$

$$IC_{S2 \rightarrow H1}: \theta_2 - c - t_{S2} \geq \theta_2 - c - pt'_{H1,2} - (1-p)t'_{H1,\emptyset}$$

$$IC_{S_2 \rightarrow S_1}: \theta_2 - c - t_{S_2} \geq \theta_2 - c - pt'_{S_1,2} - (1-p)t'_{S_1,\emptyset}$$

$$IC_{S_2 \rightarrow H_2}: \theta_2 - c - t_{S_2} \geq \theta_2 - c - t_{H_2}$$

$$CIC_{H_1,j,j'}: T_{H_1,j} \geq T_{H_1,j'}, j, j' \in \{1,2,\emptyset\}$$

$$CIC_{S_1,j,j'}: T_{S_1,j} \geq T_{H_1,j'}, j, j' \in \{1,2,\emptyset\}$$

$$NE_{H_1,j}: w_{H_1,j} \geq w_{H_1,\emptyset}$$

$$NE_{S_1,j}: w_{S_1,j} \geq w_{S_1,\emptyset}$$

$$NonNeg_{H_1,j}: w_{H_1,j} \geq 0$$

$$NonNeg_{S_1,j}: w_{S_1,j} \geq 0$$

$$LLC_{H_1,2}: t_{H_1,2} \leq \theta_2$$

$$LLC_{S_1,2}: t_{S_1,2} \leq \theta_2$$

$$LLC_{H_1,\emptyset}: t_{H_1,\emptyset} \leq \theta_1$$

$$LLC_{S_1,\emptyset}: t_{S_1,\emptyset} \leq \theta_1$$

$$LLC_{H_1,1}: t_{H_1,1} \leq \theta_1$$

$$LLC_{S_1,1}: t_{S_1,1} \leq \theta_1$$

$$LLC_{H_2}: t_{H_2} \leq \theta_2$$

$$LLC_{S_2}: t_{S_2} \leq \theta_2.$$

Note that IR_{S1} can be ignored, as it is implied by the IR_{H1} and IC_{S1-H1} constraints. Similarly, IR_2 can be ignored, as it is implied by the IR_{H2} and IC_{S2-H2} constraints.

Also, any constraints preventing low types from emulating high types are ignored and it can be verified later that they are nonbinding in the optimal contract.

1.4.1 LCCP Revisited ($q > 0$)

Details are in Appendix 1B. As expected, neither H1 nor H2 receive any rent. Remarkably, the LCCP contracts for the strategic types are not affected by the introduction of honest types in to the model. When the principal deters all corruption, the honest agent's inability to bribe is rendered moot, and S1 and H1 are pooled. Hence, the LCCP contract for a strategic type is immune to any problems presented by the existence of honest agents. Whereas when all agents were strategic, (A3) meant that the LCCP contract was strictly suboptimal, the introduction of honest agents means that the LCCP contract may be optimal even when $p > \frac{f_1}{1-f_1}$.

The LCCP contract:

$$t_{H1,1} = t_{H1,\emptyset} = \theta_1 - c$$

$$t_{S1,1} = t_{S1,\emptyset} = \theta_1 - c$$

$$t_{H1,2} = t_{S1,2} = \theta_2$$

$$w_{H1,2} = w_{S1,2} = \Delta\theta + c$$

$$t_{H2} = \theta_2 - c$$

$$t_{S2} = \theta_2 - c - (\Delta\theta(1-p) - pc)$$

$$u_2 = (\Delta\theta(1 - p) - pc)$$

1.4.2 Allowing Corruption ($q > 0$)

We now explore the possibilities for corruption by not imposing the CIC or NE constraints. Without being able to rely on collusion-proofness to help present the problem, we have to take some initial steps to rule out some corruption cases in order to simplify the feasible set of contracts. Details are in Appendix 1C, but we outline next that if bribery is optimal, it will only occur after a signal of no information (\emptyset). To show this, we start by arguing that bribery will not occur after a signal of 2, and that extortion and framing will not occur in equilibrium.

1.4.2.1 Feasible corruption cases

Note first that an honest type will not get rent in equilibrium since they are non-strategic, which means $t_{H2} = \theta_2 - c$, and $pt_{H1,1,F} + (1 - p)t_{H1,\emptyset} = \theta_1 - c$. The proof is (C3) in Appendix 1C. Even though the expected transfer ($t_{H1,1,F} + (1 - p)t_{H1,\emptyset}$) is constant, the distribution of the two transfers ($t_{H1,1,F}, t_{H1,\emptyset}$) matter and will be discussed later. For example, bribery can be deterred by removing the stake of bribery and setting them equal to each other.

Characterizing the feasible cases of corruption requires tedious steps as there are many cases to consider. In Appendix 1C, in (C1), we prove that without loss of generality, we can assume there is no bribery when the auditor finds a signal of 2. Technically, we prove in (C1) that the coalition cannot gain by reporting or moving away from a signal of 2 through bribery. The main intuition for the result is this event is off the equilibrium path since the agent will be induced to report truthfully and auditing only takes place after the agent reports 1. Thus offering a large enough reward $w_{i,2}$ to deter alteration of the signal is not costly to the principal. Thus we find that it is optimal to set

$$w_{i,2} = t_{i,2} - \min\{T_{i,1}, T_{i,\emptyset}\} \text{ for } i \in \{H1, S1\}.$$

This also allows the principal to extract the maximal transfer after a contradictory signal of 2, which is proved in (C2). This also means that

$$t'_{i,2} = t_{i,2} = \theta_2.$$

Next we argue that extortion and framing will be deterred in equilibrium. Thus, the presence of honest agents (or whether there is hidden information or moral hazard) does not affect the notion that extortion works against efficient incentives. Suppose that, to prevent bribery, the principal sets $w_{i,1} < w_{i,\emptyset}$, for $i \in \{H1, S1\}$. This will lead to extortion or framing since the auditor will have a credible threat of suppressing a signal of 1. In that case, since the auditor will extort/frame after a signal of 1, a reward to the agent after a re-affirming signal of 1 ($t_{i,1} < t_{i,\emptyset}$) becomes ineffective. This implies that the principal cannot lose by setting $t_{i,1} = t_{i,\emptyset}$, but then there is not point offering a reward to the auditor either ($w_{i,1} = w_{i,\emptyset} = 0$). The principal will optimally deter extortion/framing.

Finally, we outline the argument that if there is bribery, it must be that the coalition alters a null signal to a signal 1, which is also the only form of corruption that can occur in equilibrium. Again, the proof is in Appendix 1C as part of (C8). Since we have already ruled out extortion/framing and bribery after a signal of 2, we are mainly arguing here that it is not optimal to allow bribery after a signal of 1.

The intuition for this result is based on the fact that it is efficient to provide incentives in the form of a lower tax after a signal of 1 as it is more informative than a null signal, i.e, $t_{i,1} < t_{i,\emptyset}$. But then, for the coalition to find it optimal to alter a signal of 1 to null, it would have to be that the reward to the auditor for reporting a null signal is very high: $t_{i,1} - w_{i,1} > t_{i,\emptyset} - w_{i,\emptyset}$, or $w_{i,\emptyset} - w_{i,1} > t_{i,1} - t_{i,\emptyset} > 0$, which would be suboptimal as it would induce extortion.

The auditor's wages on the equilibrium path will be set to 0 ($w_{i,\emptyset} = w_{i,1}=0$). There is no reason to pay the auditor after a signal 1 if the agent is rewarded ($t_{i,1} < t_{i,\emptyset}$), which implies that the auditor cannot be paid after a null signal either as that would only encourage extortion. Thus we have argued that

$$w_{i,1} = w_{i,\emptyset} = 0; \text{ and } t_{i,1} \leq t_{i,\emptyset}.$$

Having reduced the possible cases of corruption, we are now ready to present the principal's problem that will help us identify the optimal contract.

1.4.2.2 BA contract: Optimal contract when all types are included (no shut-down)

When the principal chooses the contract, it may find it optimal to shut down some types. In this sub-section, we characterize the optimal contract when the principal does not have the option to shut down any type. We will consider this possibility of shut down in the next sub-section. Collecting the results from the previous sub-section, we can present the principal's problem.

The maximization problem (BA contract or no shut down):

$$\begin{aligned} \max_{t,w} f_1 & (q(pt_{H1,1} + (1-p)t_{H1,\emptyset}) + (1-q)t_{S1,1}) \\ & + (1-f_1)(q(\theta_2 - c) + (1-q)t_{S2}) \text{ s. t.} \end{aligned}$$

$$IR_{H1}: \theta_1 - c - pt_{H1,1} - (1-p)t_{H1,\emptyset} = 0$$

$$IC_{S1 \rightarrow H1}: (p + (1-p)\lambda)t_{S1,1} + (1-p)(1-\lambda)t_{S1,\emptyset} \leq (p + (1-p)\lambda)t_{H1,1} + (1-p)(1-\lambda)t_{H1,\emptyset},$$

$$IC_{S2 \rightarrow S1}: t_{S2} \leq p\theta_2 + (1-p)(\lambda t_{S1,1} + (1-\lambda)t_{S1,\emptyset}),$$

$$IC_{S2 \rightarrow H1}: t_{S2} \leq p\theta_2 + (1-p)(\lambda t_{H1,1} + (1-\lambda)t_{H1,\emptyset}),$$

$$LLC_{S1,\emptyset}: \theta_1 \geq t_{S1,\emptyset},$$

$$LLC_{H1,\emptyset}: \theta_1 \geq t_{H1,\emptyset},$$

$$CIC_{H1,\emptyset}: t_{H1,1} \leq t_{H1,\emptyset},$$

$$CIC_{S1,\emptyset}: t_{S1,1} \leq t_{S1,\emptyset}.$$

It can be verified later that the ignored constraints are satisfied by the optimal contract. The optimal contract that induces bribery is denoted as the BA contract.

BA contract:

$$t_{S1,\emptyset} = t_{H1,\emptyset} = \theta_1$$

$$t_{S1,1} = t_{H1,1} = \theta_1 - \frac{c}{p}$$

$$t_{S2} = \theta_2 - c - \left(\Delta\theta(1-p) - \frac{p - (1-p)\lambda}{p} c \right)$$

$$u_2 = \Delta\theta(1-p) - pc - \frac{1-p}{p} (p-\lambda)c.$$

Note that we have written the rent u_2 in way that facilitates comparison with the LCCP rent. It can be seen that the LCCP rent is larger if $p > \lambda$.

Given the analysis so far, we know that the key option for the principal is whether to allow or deter bribery after a null signal. Some useful remarks:

- *Since S2 can claim to be either H1 or S1, it turns out that the principal has to offer the same transfers to H1 and S1. Otherwise, only one of the IC_{S2} are binding and the principal can improve its payoff by manipulating the transfers in the slack IC_{S2}.*

-
- Hence, if CIC constraints are violated for one low type, they will be violated for both
 - When $t_{H1,1} = t_{H1,\emptyset}$, we must also have $t_{S1,1} = t_{S1,\emptyset}$. Given IR_{H1} neither receive rent.
 - If $t_{H1,1} < t_{H1,\emptyset}$, S1 gets rent because they can bribe and pay the lower transfer.
 - The benefit of having $t_{H1,1} < t_{H1,\emptyset}$ is that S2's rent decreases if $p > \lambda$. The overall net benefit depends on the ratio $\frac{f_1}{1-f_1}$.

Under LCCP, the agent is not penalized after a null signal. By inducing bribery, the principal brings back a penalty after a null signal in the form of a bribe payment to the auditor. Thus, the bribe helps improve incentives. However, given truth-telling, on the equilibrium path this bribe is a cost imposed on the agent, for which she has to be compensated in the IR. Off the equilibrium path, the bribe acts as an incentive to tell the truth. The higher is p , the more efficient is auditing and the lower the cost of bribery for the agent on the equilibrium path.

If there were only strategic agents, as in section 3 ($q = 0$), the condition for optimality of inducing bribery is $p > \frac{f_1}{1-f_1}$. Note that λ is missing here.¹¹ The reason is the same as in KLY. While λ determines the size of the bribe, it is only p and f_1 which determine the relative cost (on the equilibrium path) and benefit (off the equilibrium path) of this bribe.¹² This depends on the linearity of the utility function of the agent. Thus, it is the frequency (coefficients) that determine the optimality of allowing bribery rather than the size of the bribe.

In contrast, when $q > 0$, the bargaining power λ also represents the relative 'advantage' of a strategic type over an honest type when there is bribery. The larger is λ , the smaller

¹¹ Note that q is also missing, because neither H1 nor H2 get rent and are irrelevant in determining the efficiency of auditing in lowering the rent.

¹² When λ is high, the agent pays a smaller bribe and the incentive effect is smaller.

the bribe and larger the rent. In other words, there is an additional cost imposed by the size of the rent for S1, and indirectly for S2. Even though the transfers of H1 and S1 are pooled, S1 obtains a rent since she can pay a bribe to reduce her payment after null signal. A similar relation holds for S2.

Thus, the difference $p - \lambda$ represents the benefit of inducing bribery, and we show that it is optimal to do so when $p - \lambda > \frac{f_1}{1-f_1}$.

Proposition 1.1: *If no type is shut down, BA dominates LCCP if and only if $p - \lambda > \frac{f_1}{1-f_1}$.*

Proof: In Appendix 1C.

1.4.2.3 Shut-down an option: if shut down, it will be H1

Up to now, we have not allowed the principal to choose one of his choice variables, which is the possibility of shutting down types. We analyze this choice and show that if there is shut down it must be H1. By shutting down H1, the principal can extract rent of S1, which in turn allows it to extract rent from S2 as well. But, shutting down S1 without shutting down H1 offers no benefit as S2 can still earn rent by claiming to be H1. Since not one wants to pretend to be S2, there is no gain from shutting down S2. The proofs are in Appendix 1C.

Here we only consider shut down when bribery is induced. If bribery is deterred, the LCCP is the optimal contract without any type shut down.

1.4.2.4 Shut-down contract (SD)

Details in Appendix 1D. In this case, bribery is allowed, and H1 is shut down. Since H1 is shut down, and the optimal transfer for H2 is simply $t_{H2} = \theta_2 - c$, the principal's problem is very similar to the problem when $q = 0$, and so is the solution:

$$t_{S1,\emptyset} = \theta_1$$

$$t_{S1,1} = \theta_1 - \frac{c}{p + (1-p)\lambda}$$

$$\theta_2 p + (1-p)(\lambda t_{S1,1} + (1-\lambda)t_{S1,\emptyset}) = t_{S2}$$

$$t_{S2} = \theta_2 p + (1-p) \left(\theta_1 - c \left(\frac{\lambda}{p + (1-p)\lambda} \right) \right) = \theta_2 p + \theta_1 (1-p) - \frac{(1-p)\lambda}{p + (1-p)\lambda} c$$

$$t_{S2} = \theta_2 - c - \left(\Delta\theta(1-p) - \frac{p}{p + (1-p)\lambda} c \right)$$

$$\begin{aligned} \Pi^{SD} &= f_1(1-q) \left(\theta_1 - c - \frac{(1-p)(1-\lambda)}{p + (1-p)\lambda} c \right) \\ &\quad + (1-f_1) \left(\theta_2 - c - (1-q) \left(\Delta\theta(1-p) - \frac{p}{p + (1-p)\lambda} c \right) \right) \end{aligned}$$

1.4.3 Comparing Contracts: Honesty hurts the principal for small q

The principal has two main options: Allow corruption or not, and shut down H1 or not. Shut down of H1 allows the principal to extract all rent from S1, which also reduces the rent of S2. However, the principal loses the surplus from H1, which we show next to be optimal only for small q .

1.4.3.1 Shut down of H1 is optimal for small q

Details in Appendix 1D. We can show this by comparing the profits of the three potentially optimal contracts.

Least Cost Corruption Proof (LCCP):

$$\Pi^{LCCP} = f_1(\theta_1 - c) + (1 - f_1)(\theta_2 - c - (1 - q)(\Delta\theta(1 - p) - pc))$$

Bribery Allowed (BA):

$$\begin{aligned} \Pi^{BA} = f_1 & \left(\theta_1 - c - (1 - q) \left(\frac{1 - p}{p} c \right) \right) \\ & + (1 - f_1) \left(\theta_2 - c - (1 - q) \left(\Delta\theta(1 - p) - \frac{p - (1 - p)\lambda}{p} c \right) \right) \end{aligned}$$

Bribery Allowed, Shut Down H1 (SD):

$$\begin{aligned} \Pi^{SD} = f_1(1 - q) & \left(\theta_1 - c - \frac{(1 - p)(1 - \lambda)}{p + (1 - p)\lambda} c \right) \\ & + (1 - f_1) \left(\theta_2 - c - (1 - q) \left(\Delta\theta(1 - p) - \frac{p}{p + (1 - p)\lambda} c \right) \right) \end{aligned}$$

Define the cut-off probability p^* by setting $\Pi^{BA} - \Pi^{LCCP}$ and solving for p :

$$p^* \equiv \lambda + \frac{f_1}{1 - f_1}$$

When $p > p^*$, $\Pi^{BA} > \Pi^{LCCP}$. When $p < p^*$, $\Pi^{LCCP} > \Pi^{BA}$.

This inequality can also determine which profit condition is tighter when comparing profits to SD.

For the purposes of proving that honesty hurts the principal, we will first show that the SD contract is optimal for a range of q , given (A3) holds.

(i) Comparing SD and LCCP contracts, $\Pi^{SD} - \Pi^{LCCP}$ must be positive:

$$q < \frac{((1 - f_1)p - f_1)(1 - p)(1 - \lambda)c}{((1 - f_1)p - f_1)(1 - p)(1 - \lambda)c + f_1(p + (1 - p)\lambda)(\theta_1 - c)} = q_1$$

By (A3), we know that $(1 - f_1)p > f_1$, which guarantees both the numerator and the denominator are positive, and so $0 < q_1 < 1$.

Note that q_1 is increasing in p . That is to say, while the LCCP and SD contracts both benefit from an increase in p , the SD contract benefits more. The cost of shutting down H1 depends only on $f_1 q(\theta_1 - c)$, but the general cost of allowing bribery is highly sensitive to p . The lower p gets, the less often bribery occurs, and the more efficient the SD contract becomes in that respect.

(ii) Comparing SD and BA contracts, $\Pi^{SD} - \Pi^{BA}$ must be positive:

$$q < \frac{(1-p)\lambda(f_1 + (1-f_1)(1-p)\lambda)c}{(1-p)\lambda(f_1 + (1-f_1)(1-p)\lambda)c + p(p + (1-p)\lambda)f_1(\theta_1 - c)} = q_2$$

Clearly, $0 < q_2 < 1$.

Note that q_2 is decreasing in p . This is because the BA contract is even more sensitive to changes in p than the SD contract, due to the inclusion of H1. Rent to both strategic types are based on the fact that, $(1-p)$ of the time, H1 gets punished and requires compensation for that punishment.

From these two contract comparisons, we now have the following lemma:

Lemma 1.1

Define $\underline{q} = \min\{q_1, q_2\}$. We know that $0 < \underline{q} < 1$, because both q_1 and q_2 are between 0 and 1. When $0 \leq q < \underline{q}$, the SD contract is the optimal contract.

See Figure 1.1 for a graphical example of the optimal contract space.

Proposition 1.2: Honesty hurts the principal

The principal's profit decreases in q when $0 < q < \underline{q}$.

Per Lemma 1.1, the optimal contract is the SD contract when $0 < q < \underline{q}$, where \underline{q} is defined in the previous section and exists between 0 and 1. Given the principal offers the SD contract, assumptions (A1) and (A3) are sufficient to show that the derivative of the principal's profit with respect to q is negative.

1.5 Conclusion

Policies that combat corruption by increasing the relative population of honest agents in a mostly-corrupt population can be counterproductive. This result is tied to the issue that honesty creates an additional dimension of adverse selection, which then creates a broad new set of restrictive constraints on the principal's optimal contract. Khalil, Lawarrée and Yun (2010) proposed a novel way for corruption to be optimal, but in the presence of honesty, allowing corruption in equilibrium creates additional rents for agents with the will to exploit it. The principal must choose between offering these rents and shutting down some honest agents. This policy of shutting down honest agents at a loss helps explain how a corrupt equilibrium is maintained in a larger society.

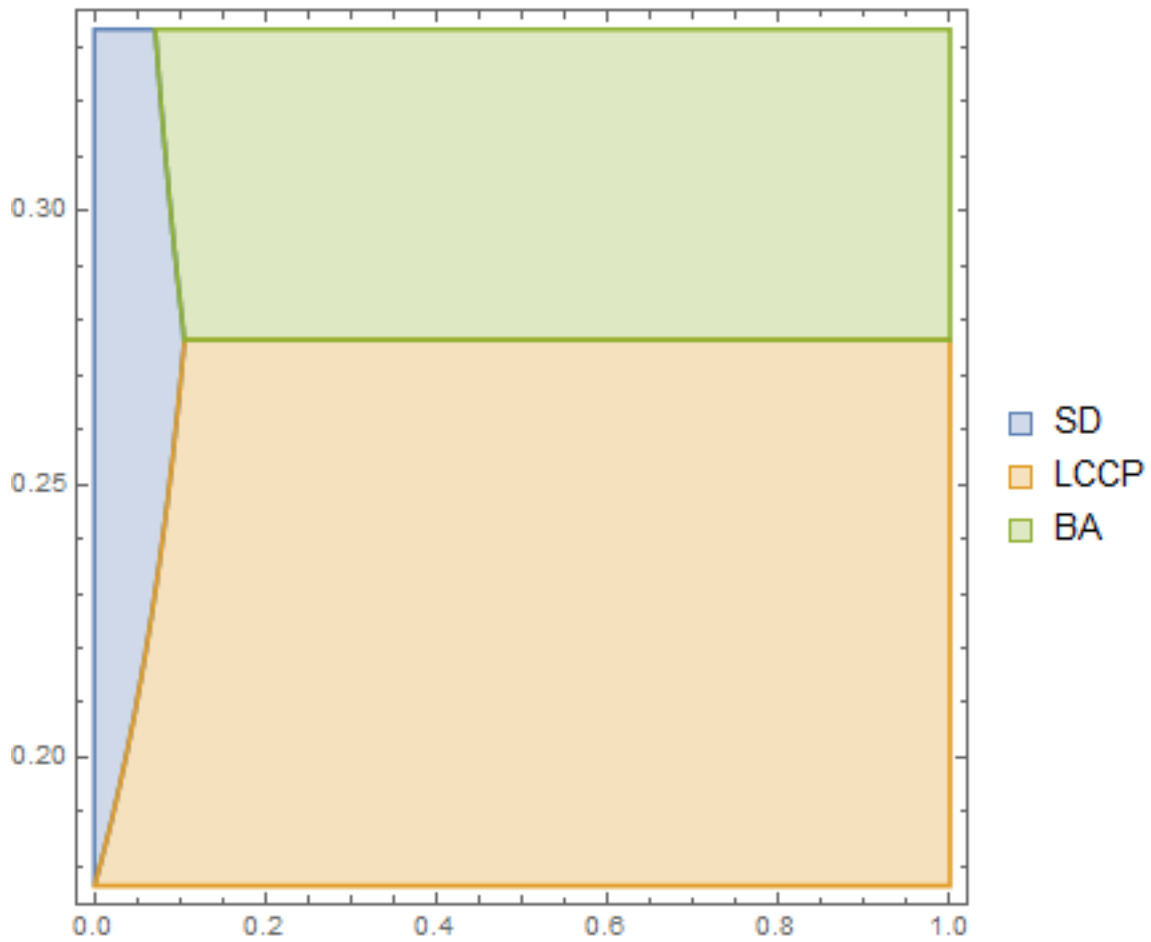
To be clear, the principal is not a benevolent social planner in this model, but rather a profit-maximizing investor. In theory, this investor could act on behalf of the government, but the objectives of this actor are self-serving. Notice how the principal's own self-serving motives lead to shutdown, which acts as a homogenizing force within the organization. A group that has some tendency to move towards honesty may remain corrupt, as much of the honest population goes to work elsewhere.

The contributions in this paper provide sharp contrast to insights developed by similar research. In Kofman and Lawarrée (1996), potential incorruptibility complements lax enforcement of corruption, whereas here, honesty in the agent can make lax enforcement more costly. Yun (2012) advocates for increasing the cost of corruption past a threshold that deters all corruption lest the costs lead to wasted resources, but this paper

shows even an infinite cost of corruption, with zero cost when implemented, can be harmful to the principal, as long as there is heterogeneity.

This paper emphasizes the role of heterogeneous ethics in a society, and how power and productivity change the agent's and the principal's fates. When corruption is prevalent honest types struggle to blend in, and strategic types capitalize. When honesty is prevalent, powerful corrupt agents obtain more rents than anyone else by leveraging their willingness to engage in illegal activity.

Figure 1.1: Optimal Contracts, Numerical Example



x-axis: p

y-axis: q

$\theta_1 = 10$

$\theta_2 = 11.5$

$c = 1$

$f_1 = 0.15$

$\lambda = 0.1$

Chapter 2

Commitment to Auditing with a Private Signal

2.1 Introduction

The literature on auditing without commitment¹³ argues that the ability to commit to an auditing scheme provides a strong benefit. The intuition is clear: Commitment removes issues of credibility for the principal, who, having deterred bad behavior, may wish to save on costs by being lax in enforcement *ex post*. One feature of commitment is transparency in auditing, meaning that the principal openly specifies aspects of the audit such as its frequency. Yet, many auditors and regulators lack explicit, contractual rules on when to audit that are shared with the potentially-audited party, the tax authority is not visibly committing to auditing in the sense considered in the literature. Even if there exist internal auditing rules, commitment to an auditing scheme requires a transparent, enforceable statement of those auditing rules.

Considerable effort in the literature has been expended to explain how regulators who appear to lack commitment to an auditing scheme may in fact imply it. Melumad and Mookherjee (1989) propose a method for the IRS to emulate full commitment by providing multiple external signals. What this ultimately shows is that an auditing agency such as the IRS *can* commit to an auditing scheme, but it is under no obligation to do so, and may in fact not. Furthermore, a standard feature of commitment to auditing is full compliance. While there are more compliers than a naïve theory of incentives may predict (Andreoni et al (1998)), there are still legitimate cases of tax fraud.

¹³ See Bester and Strausz (2001), Khalil (1997), and Khalil and Lawarrée (2006) for examples of a lack of commitment harming the principal's profits.

There is a technological reason¹⁴ to avoid revealing the specifics of a monitoring scheme or algorithm: The agent may exploit the scheme and find it easier to avoid detection. Farlee (2010) discusses the trade-off between disclosure and secrecy in a moral hazard setting with monitoring, focusing on a trade-off between risk and efficiency. Khalil and Lawarrée (2001) show that if there are many observable signals which may indicate wrongdoing, and the regulator finds it too costly to observe them all, to commit to observing one particular signal and not another would invite the agent to correctly falsify only the observed signal. While the type of monitoring may remain unspecified, the principal can still commit to a particular probability of monitoring the agent. An auditor can obscure the parts of a scheme which are exploitable – i.e. the technological aspects of monitoring – while clearly specifying the probability an individual agent is actually monitored. In other words, *how* the principal audits need not be specified, as long as *how likely* the principal audits is specified.

Another reason the principal may not contractually commit to auditing probabilities is if he is incapable of commitment. The principal may prefer not to contract on auditing probabilities if he observes private information about the agent prior to auditing. Consider a tax auditor who must decide whether to investigate a potentially fraudulent return. He may receive an anonymous tip from a whistleblower, collect information which may not be acceptable to present in court, or subjectively feel that a return is suspicious based on experience. If the auditor has contractually committed to particular auditing probabilities, for those probabilities to take advantage of his private information, they must be based on his report and not the information itself. If taxpayers are fully compliant, as implied by the standard revelation principle, the auditor has a credibility problem; he will report any signal that contractually results in a lower auditing probability to save on costs. The auditor's

¹⁴ Another reason to avoid explicit auditing probabilities is transactions costs, as in the incomplete contracts literature (see Tirole (1999) for a review), but the mere existence of transactions costs does not explain why a principal would avoid explicitly committing to auditing probabilities in high stakes environments where the payoffs are too large for transactions costs to matter, e.g. auditing high income brackets.

private information, while beneficial, provides him another way to avoid his commitment to auditing *ex post*, if the private information is used in the contract.

In the more realistic case where taxpayers are not always deterred from fraud, however, the costs borne by auditing may be balanced out by the occasional penalty collected. In particular, forgoing commitment to auditing leads to a mixed strategy equilibrium that makes the principal exactly indifferent between auditing and not auditing. This can solve the principal's issues with credibility. Khalil (1997) offers a strong intuition of why commitment is superior to no commitment generally: The ability to incur auditing costs *ex post* improves incentives *ex ante* and allows auditing to occur along with an enforcement scheme that induces full compliance. If full compliance itself is suboptimal, however, the main issue with forgoing commitment to auditing may be moot.

The modified revelation principle developed by Bester and Strausz (2001) presents a solution to issues in contract theory where the principal cannot credibly commit to future outcomes in the contract, whereby the agent finds it weakly optimal to truthfully reveal his information but may not do so with probability 1. Hence, noncompliance can be an optimal scheme when the principal's private information prevents him from credibly committing to auditing probabilities based on that information.

Khalil, Lawarrée and Scott (2015) also explore issues of private information for the principal, emphasizing the timing of the private information. In their model, the principal obtains a private signal of the agent's type prior to transfers, meaning that transfers can be based on the principal's report. In this case, the principal has the opposite incentive problem: If the agent is penalized due to the private signal, the principal has an incentive to distort his private signal *ex post*. To grant credibility, the principal can "burn money," i.e. grant proceeds of the penalty to a third party. Or, if the private signal occurs prior to production, the principal can more effectively punish the agent by re-scaling the project. In contrast, this paper emphasizes the role of private information when it occurs after production and transfers, but prior to auditing. In effect, the principal cannot use the private signal to alter transfers or the project directly – only to change auditing strategies.

Acemoglu et al (2006) examine a dynamic taxation problem without commitment where the tax authority is not necessarily benevolent, and finds that commitment can lead to distortions. I follow in the tradition of a self-interested principal, in order to focus on the core monitoring problem instead of the traditional efficiency trappings of benevolent planners in the optimal taxation literature. Strausz (2006) examines the role of the timing of partially verifiable information, focusing on the trade-off between accuracy and the ability to manipulate the signal when it arrives later on.

This result runs counter to Finkle and Shin (2007), who show that, if the accuracy of the audit is endogenous and the principal lacks commitment, he will reduce the audit's accuracy to maintain an incentive to audit. In this case, a private signal can improve the principal's incentive to audit by increasing the odds of truly catching an agent red-handed, as opposed to falsely finding an innocent agent guilty.

A key assumption in this paper is the restriction of post-audit transfers to a penalty for catching the agent. If this restriction is lifted, the principal can use a number of methods to offset any differences in auditing costs via post-audit transfers. For instance, if the agent pays all the auditing costs and is fully compliant, the principal is indifferent between auditing and not auditing without the need for noncompliance. Promising a tax refund given a particular signal report can also balance out differences in auditing costs¹⁵. Removing this restriction would allow the principal to effectively violate other limits on the size of the penalty, irrespective of the issues surrounding the private signal. For instance, a small fixed penalty effectively becomes a large transfer-dependent penalty if refunds are allowed.

¹⁵ This works so long as the audited agent would have some income left over to tax were it not for the refund scheme. Refunds would not work in the basic model presented here, but that is an artifact of the simplifying assumptions.

The paper will proceed as follows: Section 2.2 sets up the model and derives the main proposition, that forgoing commitment to auditing is optimal. Section 2.3 offers extensions. Section 2.4 concludes.

2.2 The Model

A principal contracts with an agent to form a productive relationship, where the agent collects θ as productive income. The agent is the residual claimant of θ . The principal then collects a portion of the income t . We interpret t to be a tax, and set limited liability such that $t \leq \theta$ in all cases. Alternatively the agent can reject the relationship and collect an outside option normalized to 0. θ can take two forms, θ_1 or θ_2 , known privately to the agent before signing the contract.

$$P(\theta = \theta_1) = \pi; P(\theta = \theta_2) = 1 - \pi; \theta_2 - \theta_1 = \Delta\theta > 0$$

After the agent's report and payment, the principal can audit the agent's income perfectly at cost c with some probability k . If an agent is caught reporting $\theta = \theta_1$ when $\theta = \theta_2$, the principal may only extract a fixed fine $F, F > \Delta\theta$.

Prior to auditing but after transfers, the principal receives a private signal σ . With probability $q > 0.5$, σ shows the correct type; with probability $1 - q$, it shows the incorrect type. Whether and how the principal can use this signal to inform his auditing strategy depends on the contract chosen.

In the first best, the principal can extract all income such that $t_i = \theta_i, i \in \{1,2\}$. In the second best, the principal has no way to screen the agents, so he offers a pooling contract $t_i = \theta_1$, granting the wealthy agent his full information rent $\Delta\theta$. The timing of the game is as follows:

1. The agent observes θ .

-
2. The principal offers the agent a contract; the agent rejects, ending the game, or accepts the contract and reports his type.
 3. The agent collects θ and provides a transfer t to the principal based off his report.
 4. The principal observes σ and makes a report.
 5. The principal audits the agent with probability k , and collects a penalty upon proof of a false report.

I make the following assumptions to simplify the analysis:

$$(A1): \frac{\pi}{1 - \pi} > \frac{\Delta\theta}{\theta_1}$$

$$(A2): F > \frac{\pi}{\pi - \pi} c$$

$$(A3): q > \frac{\Delta\theta}{F}$$

$$(A4): F - c > c \frac{\pi}{1 - \pi} \frac{1 - q}{q}$$

(A1) ensures the principal never wishes to shut down the low type. (A2) ensures that auditing is profitable. (A3) ensures that auditing is always powerful enough to extract all rent from the high type, as the focus of this model is on the trade-off between auditing costs and shirking. (A4) ensures that allowing noncompliance to alleviate principal incentive issues is feasible, if not necessarily optimal.

2.2.1 Benchmark Contracts

2.2.1.1 Commitment to auditing, no signal

Given the principal can commit to an auditing scheme with no private signal, he can offer a contract according to P_0 :

$$\max_{t_i, k} \pi(t_1 - kc) + (1 - \pi)(t_2)$$

$$IR_1: \theta_1 - t_1 \geq 0$$

$$IR_2: \theta_2 - t_2 \geq 0$$

$$IC_2: \theta_2 - t_2 \geq \theta_2 - t_1 - kF$$

$$NN: k \geq 0$$

The derivation of the optimal contract is in Appendix 2A. Due to (A2), principal prefers to use the audit, and the contract is as follows:

$$t_i = \theta_i$$

$$k = \frac{\Delta\theta}{F}$$

The transfers are first best, but the principal incurs an auditing cost of $\pi \frac{\Delta\theta}{F} c$.

2.2.1.2 Commitment to auditing with a public signal

If the signal observed by the principal is publically verifiable, he can commit to auditing based on the signal and need not rely on his own report. k_i is the contracted probability of auditing given the principal receives a signal of i . (A3) ensures that $k_1 = 0$ in equilibrium.

$$\max_{t, k} \pi(t_1 - qk_1c - (1 - q)k_2c) + (1 - f_1)(t_2) \text{ s. t.}$$

$$IR_1: \theta_1 - t_1 \geq 0$$

$$IR_2: \theta_2 - t_2 \geq 0$$

$$IC_2: \theta_2 - t_2 \geq \theta_2 - t_1 - (qk_2 + (1 - q)k_1)F$$

$$NN_1: k_1 \geq 0$$

IR_1 binds. If it did not, the principal could increase t_1 , slackening IC_2 and increasing the objective function, until it did. As auditing is optimal due to (A2), IR_2 binds. If it did not, the principal could increase k_2 until it did. The restriction on q ensures that IR_2 will bind before $k_2 = 1$.

Given both IR constraints bind, IC_2 becomes:

$$\frac{\Delta\theta}{F} \leq qk_2 + (1 - q)k_1$$

The total probability of a tax-evading agent being audited off the equilibrium path is the same as in P_0 , but the actual probability the principal engages in an audit decreases. This is because the agent complies, and the signal improves the efficiency of the audit.

$$k_1 = 0; k_2 = \frac{\Delta\theta}{Fq}$$

Note that the cost of auditing is $\frac{1-q}{q} \frac{\Delta\theta}{F} c$, which is less than the cost of auditing in the standard commitment contract, because $\frac{1-q}{q} < 1$.

2.2.1.3 Auditing without commitment, with a private signal

Suppose the principal cannot or does not commit to auditing probabilities k_1 and k_2 , and must decide them ex-post. For the principal to audit with some probability between 0 and 1 given a particular signal occurs, he must be indifferent to auditing. This is achieved by allowing the agent to shirk with probability x . Note that this guarantees the principal is *not* indifferent to auditing given the other signal occurs. In particular, if the principal is

indifferent between auditing and not auditing given $\sigma = 2$, then he strictly prefers not to audit when $\sigma = 1$. (A3) ensures that the principal will be indifferent to auditing when $\sigma = 2$ and will strictly prefer not to audit when $\sigma = 1$. Ultimately, the principal will set $k_1 = 0$, and he faces the problem P_1 :

$$\max_{t_i, k_i, x} (\pi + (1 - \pi)x)(t_1 - [0]) + (1 - \pi)x(t_2) \text{ s. t.}$$

$$IR_1: \theta_1 - t_1 \geq 0$$

$$IR_2: \theta_2 - t_2 \geq 0$$

$$IC_2: \theta_2 - t_2 = \theta_2 - t_1 - qk_2P$$

$$PIC: c = F \frac{(1 - \pi)qx}{\pi(1 - q) + (1 - \pi)qx'}$$

And relevant LLCs. The ICs now represent binding indifference conditions for the agent (IC2) and principal (PIC) to optimally randomize shirking and auditing, respectively. The $[0]$ in the objective function represents the 0 profit auditing scheme faced by the principal. Importantly, this means changing k_2 does not affect the objective function.

The IR constraints bind. If IR_1 did not bind, the principal could increase t_1 and decrease k_2 , maintaining IC_2 , which would increase the objective function and tighten IR_1 until it was binding. If IR_2 did not bind, the principal could increase t_2 and k_2 , maintaining IC_2 , which would increase the objective function and tighten IR_2 until it was binding.

Noting that the IR constraints bind, $t_2 = \theta_2$; $t_1 = \theta_1$; $t_2 - t_1 = \Delta\theta$

Re-stating indifference conditions:

$$c((1 - \pi)qx + \pi(1 - q)) = F(1 - \pi)qx$$

$$c\pi(1 - q) = x(1 - \pi)q(F - c)$$

$$x = \frac{\pi}{1 - \pi} \frac{c}{F - c} \frac{1 - q}{q}$$

$$k_2 = \frac{\Delta\theta}{Fq}$$

As $q \rightarrow 1$, $x \rightarrow 0$, indicating that there is some q^* where no commitment with a private signal is superior to commitment without any signal.

Lemma 2.1: Forgoing commitment to auditing is superior to commitment to auditing without a signal for $q \geq q^*$.

The sole cost of the no-commitment contract is noncompliance, as auditing entails exactly zero expected profit for the principal. The sole cost of the standard commitment contract is auditing costs, as there is full compliance with costly auditing. Both contracts entail first best transfers given the agent tells the truth, so we can derive the critical q by comparing noncompliance costs with auditing costs:

$$(1 - \pi)x\Delta\theta \leq \pi \frac{\Delta\theta c}{F}$$

$$\pi \frac{\Delta\theta c}{F - c} \frac{1 - q}{q} \leq \pi \frac{\Delta\theta c}{F}$$

$$(1 - q)F \leq q(F - c)$$

$$F \leq q(2F - c)$$

$$q \geq q^* = \frac{F}{F + (F - c)}$$

This is feasible given (A2).

2.2.2 Auditing with commitment and a private signal

Implicit in the no-commitment contract was the principal's proper reporting of the private signal. The principal had no incentive to lie about the signal, because the principal's report of the signal had no effect on outcomes without commitment. In the case of full compliance and commitment, the principal will prefer to report any signal where he gets to audit less often, because auditing is purely a cost *ex post*. Because of this, the Principal Incentive Compatibility constraints are made explicit:

$$PIC_1: k_1(\alpha_1 F - c) \geq k_2(\alpha_1 F - c)$$

$$PIC_2: k_2(\alpha_2 F - c) \geq k_1(\alpha_2 F - c)$$

Where α_i indicates the probability $\theta = \theta_2$ given the agent reports θ_1 and the signal states θ_i .

With full compliance, $\alpha_i = 0 \forall i$, since the agent never misreports. Therefore, the *PIC* states that $k_1 = k_2$ for any $c > 0$. Note that the private signal occurs after standard tax transfers, so the private signal can only inform auditing. Therefore, if the principal induces full compliance and commits to an audit, he cannot use the private signal. Per Bester and Strausz (2001), the optimal contract may induce a mixed strategy with compliance and noncompliance so long as compliance is a best response. I continue to define x as the probability the high type shirks, and full compliance is equivalent to $x = 0$. In this case, we can define α_1 and α_2 in terms of x :

$$\alpha_2 = \frac{(1 - \pi)qx}{(1 - \pi)qx + \pi(1 - q)}$$

$$\alpha_1 = \frac{(1 - \pi)(1 - q)x}{(1 - \pi)(1 - q)x + \pi q}$$

Lemma 2.2: The principal can satisfy the Principal Incentive Compatibility constraints in one of two ways: Either set $k_1 = k_2$, or set $k_2 > k_1$ and $\alpha_1 \leq \frac{c}{F} \leq \alpha_2$. The minimum shirking required to satisfy these constraints is then $x = 0$ for $k_1 = k_2$ and $\alpha_1 < \frac{c}{F} = \alpha_2$ for $k_2 > k_1$.

First, notice that if $k_1 = k_2$, both PIC_1 and PIC_2 hold exactly.

Next, notice that $\alpha_2 \geq \alpha_1$:

$$\alpha_2 - \alpha_1 = (1 - \pi)x \left[\frac{q}{(1 - \pi)qx + \pi(1 - q)} - \frac{1 - q}{(1 - \pi)(1 - q)x + \pi q} \right]$$

$$\alpha_2 - \alpha_1 = \frac{(1 - \pi)x[q(1 - q)(1 - \pi)x + \pi q^2 - q(1 - q)(1 - \pi)x - \pi(1 - q)^2]}{((1 - \pi)qx + \pi(1 - q))((1 - \pi)(1 - q)x + \pi q)}$$

$$\alpha_2 - \alpha_1 = \frac{(1 - \pi)\pi x[2q - 1]}{((1 - \pi)qx + \pi(1 - q))((1 - \pi)(1 - q)x + \pi q)} \geq 0$$

Suppose $k_2 > k_1$. The only way for PIC_2 to hold is to set x such that $\alpha_2 \geq \frac{c}{F}$. For PIC_1 to hold simultaneously, x must be low enough that $\alpha_1 \leq \frac{c}{F}$.

Suppose $k_1 > k_2$. The only way for PIC_1 to hold is to set $\alpha_1 \geq \frac{c}{F}$, which implies that $x > 0$. This means $\alpha_2 > \alpha_1$, and hence $\alpha_2 > \frac{c}{F}$. This violates PIC_2 .

Lemma 2.3: Commitment to auditing with full compliance and a private signal is identical to commitment to auditing without a signal.

From Lemma 2.2, we know that there are two ways to satisfy the PIC constraints. The only way that sets $x = 0$ is to set $k_1 = k_2$, which means that the principal cannot use the signal when the agent always complies, and the problem reverts to P_0 .

2.2.2.1 Commitment to Auditing with Noncompliance

In the case where the principal commits to auditing probabilities and $x > 0$, the Principal Incentive Compatibility constraints are transformed per Lemma 2.2:

$$\alpha_1 \leq \frac{c}{F} \leq \alpha_2$$

The principal's problem P_2 is as follows:

$$\begin{aligned} \max_{t, k, x} & (\pi + (1 - \pi)x)t_1 + k_1 \left((1 - \pi)(1 - q)xF - (\pi q + (1 - \pi)x(1 - q))c \right) \\ & + k_2 \left((1 - \pi)qx F - ((1 - \pi)qx + \pi(1 - q))c \right) \\ & + (1 - \pi)(1 - x)t_2 \quad s. t. \end{aligned}$$

$$IR_1: \theta_1 - t_1 \geq 0$$

$$IR_2: \theta_2 - t_2 \geq 0$$

$$IC_2: \theta_2 - t_2 = \theta_2 - t_1 - (qk_2 + (1 - q)k_1)F$$

$$NN_1: k_1 \geq 0$$

$$NN_2: k_2 \geq 0$$

$$NN_3: 1 - k_2 \geq 0$$

$$NN_4: 1 - k_1 \geq 0$$

$$PIC_2: \frac{(1 - \pi)qx}{(1 - \pi)qx + \pi(1 - q)} \geq \frac{c}{F}$$

$$PIC_1: \frac{c}{F} \geq \frac{(1 - \pi)(1 - q)x}{(1 - \pi)(1 - q)x + \pi q}$$

I ignore the NN_2 , NN_3 , NN_4 , and PIC_1 constraints and show that they are satisfied in the optimal solution. The full derivation of the optimal contract is in the appendix.

Full solution:

$$k_1 = 0$$

$$k_2 = \frac{\Delta\theta}{Fq}$$

$$t_i = \theta_i$$

$$\frac{\pi}{1-\pi} \frac{1-q}{q} \frac{c}{F-c} = x$$

Notice that the solution to P_2 is the same as the solution to P_1 .

Proposition 2.1: Forgoing commitment to auditing is optimal for $q \geq q^*$

This follows from Lemma 2.1 and the solutions of P_1 and P_2 . First, by Lemma 2.1, note that the optimal full-compliance contract is inferior to auditing without commitment with a private signal when $q \geq q^*$ per Lemma 2.1. Then, note that the optimal contract without full compliance is identical to auditing without commitment.

2.3 Extensions

2.3.1 Productive effort, transfer-dependent penalties

The results of the model are generally invariant to the underlying production model, as long as the private signal is generated after production and transfers and before auditing. In addition, the penalties can be transfer-dependent so long as there are some reasonable

restrictions on how the transfers and penalties can interact with the audit. If the agent can effectively pay for the audit, for instance, then the principal lacks an incentive issue, but this requires complicated rebate schemes in practice.

2.3.2 The model

A principal hires an agent to produce output $x = \theta e$ at cost $\frac{e^2}{2}$, where output (x) is observable but input (e) and productivity (θ) are not. θ takes two forms, θ_1 and θ_2 , $\left(\frac{\theta_1}{\theta_2}\right)^2 = \gamma < 1$, $P(\theta = \theta_1) = \pi$, $P(\theta = \theta_2) = 1 - \pi$.

The principal offers a transfer t dependent upon verifiable reports and information. After production and transfers occur, the principal may audit the agent at cost c to verify the agent's type θ with perfect accuracy. The agent has an outside option of value 0, and limited liability such that $t \geq 0$, and the maximum penalty cannot be greater than t .

All else remains the same.

2.3.3 Benchmark contracts

2.3.3.1 Second Best

If auditing is too expensive, the principal may resort to the second best contract.

$$\max_{t_i, e_i} \pi(\theta_1 e_1 - t_1) + (1 - \pi)(\theta_2 e_2 - t_2) \text{ s. t.}$$

$$IR_1: t_1 - \frac{e_1^2}{2} \geq 0$$

$$IC_2: t_2 - \frac{e_2^2}{2} \geq t_1 - \gamma \frac{e_1^2}{2}$$

The derivation is in Appendix 2C, and the solution is as follows:

$$e_2 = \theta_2$$

$$e_1 = \theta_1 \frac{\pi}{\pi + (1 - \pi)(1 - \gamma)}$$

$$t_1 = \frac{e_1^2}{2}$$

$$t_2 = \frac{\theta_2^2}{2} + \frac{e_1^2}{2}(1 - \gamma)$$

2.3.3.2 Commitment to auditing, no signal

If the principal wishes to audit, but lacks or opts not to use the private signal, he solves P_0 :

$$\max_{t_i, e_i, k} \pi(\theta_1 e_1 - t_1 - kc) + (1 - \pi)(\theta_2 e_2 - t_2) \text{ s. t.}$$

$$IR_1: t_1 - \frac{e_1^2}{2} \geq 0$$

$$IR_2: t_2 - \frac{e_2^2}{2} \geq 0$$

$$IC_2: t_2 - \frac{e_2^2}{2} \geq t_1 - \gamma \frac{e_1^2}{2} - kt_1$$

The derivation is in the appendix. Here is the optimal contract:

$$e_i = \theta_i \text{ [E1]}$$

$$t_i = \frac{\theta_i^2}{2} \text{ [E2]}$$

$$k = 1 - \gamma \text{ [E3]}$$

Because the penalty is based on the size of the transfer to the low type, there is no incentive to distort the low type's effort. With a fixed penalty, I conjecture that the principal would distort the low type's effort.

2.3.4 Auditing without commitment, with a private signal

A private signal changes the indifference conditions faced by a principal who cannot commit. In particular, given there is some shirking $x > 0$, the principal cannot be indifferent both when $\sigma = 2$ and $\sigma = 1$. I will consider the case where $q > 1 - \gamma$, i.e. the principal is indifferent between auditing and not auditing when $\sigma = 2$. Hence, the principal strictly prefers not to audit when $\sigma = 1$.

$$IC_2: t_2 - \frac{e_2^2}{2} = t_1(1 - qk_2) - \gamma \frac{e_1^2}{2} \quad [E4]$$

$$PIC: c = \frac{xq(1 - \pi)}{xq(1 - \pi) + \pi(1 - q)} t_1 \quad [E5]$$

Re-phrasing:

$$c(xq(1 - \pi) + \pi(1 - q)) = xq(1 - \pi)t_1$$

$$x(t_1 - c)q(1 - \pi) = c\pi(1 - q)$$

$$x = \frac{\pi}{1 - \pi} \frac{c}{t_1 - c} \frac{1 - q}{q} \quad [E6]$$

Notice that, since $1 - q < q$, the probability of shirking is smaller than auditing without commitment with no private signal, all else equal. I label the above problem P_1 . The solution can be characterized as follows:

$$e_2 = \theta_2 \quad [E7]$$

$$t_2 = \frac{\theta_2^2}{2} \quad [E8]$$

$$e_1 = \theta_1 \frac{(t_1 - c)(qt_1 - (2q - 1)c)}{(t_1 - c)(qt_1 - (2q - 1)c) - c(1 - q) \left(\frac{\theta_2^2}{2} - (\theta_1 e_1 - t_1) \right)} \quad [E9]$$

$$t_1 = \frac{e_1^2}{2} \quad [E10]$$

Notice the enforced overproduction from the low type, which is moderated by the strength of the private signal. As $q \rightarrow 1$, $x \rightarrow 0$ and $e_1 \rightarrow \theta_1$.

2.3.5 Commitment to auditing with a private signal

As before, for the principal to honestly report his private signal, he must have an incentive to do so. The only way to commit to auditing strategies, enforce full compliance, and report his private signal honestly is not to use the private signal. This is reflected by the Principal Incentive Compatibility constraints:

$$PIC_2: k_2 \left(c - \frac{xq(1 - \pi)}{xq(1 - \pi) + \pi(1 - q)} t_1 \right) \geq k_1 \left(c - \frac{xq(1 - \pi)}{xq(1 - \pi) + \pi(1 - q)} t_1 \right)$$

$$PIC_1: k_1 \left(c - \frac{x(1 - q)(1 - \pi)}{x(1 - q)(1 - \pi) + \pi q} t_1 \right) \geq k_2 \left(c - \frac{x(1 - q)(1 - \pi)}{x(1 - q)(1 - \pi) + \pi q} t_1 \right),$$

Where k_i is the probability of auditing given $\sigma = i, i \in \{1, 2\}$

If $x = 0$, the constraints simplify to:

$$k_2 \geq k_1$$

$$k_1 \geq k_2,$$

Meaning $k_1 = k_2$, and the principal must ignore the private signal.

Increasing k_2 is preferable to increasing k_1 , because k_2 provides a more powerful incentive. If the signal were public, the principal would set $k_1 = 0$ and weight all of the auditing when $\sigma = 2$, and enforce full compliance. However, under full compliance, the principal will have a desire to report $\sigma = 1$ if that results in a lower probability of auditing *ex post*. Increasing x , while imposing significant shirking cost, allows the principal to effectively use the signal. Hence, PIC_2 binds, and

$$k_2 \left(c - \frac{xq(1-\pi)}{xq(1-\pi) + \pi(1-q)} t_1 \right) = k_1 \left(c - \frac{xq(1-\pi)}{xq(1-\pi) + \pi(1-q)} t_1 \right).$$

Given $k_2 > k_1$, simplifies to

$$c = \frac{xq(1-\pi)}{xq(1-\pi) + \pi(1-q)} t_1.$$

The agent must abide by a familiar indifference constraint to randomly deviate:

$$IC_2: t_2 - \frac{e_2^2}{2} = t_1(1 - qk_1) - \gamma \frac{e_1^2}{2}.$$

Hence, the problem is identical to P_1 .

The logic behind **Proposition 2.1** holds in this setting as well, leading to a similar proof that the no-commitment contract is optimal:

Step 1: Commitment to auditing without full compliance is identical to lacking commitment.

Step 2: There is some $q^* < 1$ where noncompliance is superior to forgoing the private signal. To see this, consider the case where $q \rightarrow 1$. Recall from equations [E6] to [E9] that, as $q \rightarrow 1$, $x \rightarrow 0$ and $e_1 \rightarrow \theta_1$, and auditing costs go to zero. Hence, the contract approaches first best quantities, transfers, and profits. Also notice that commitment to auditing with full compliance has first best quantities and transfers, but entails a first-order auditing cost, and is hence clearly inferior when $q \rightarrow 1$.

However, as $q \rightarrow 0.5$, the no-commitment contract approaches the standard no-commitment contract without auditing, which is clearly inferior to a commitment contract with full compliance. I conjecture that profits are monotonically increasing in q , and hence there is an intermediate value q^* that separates the optimality of the no-commitment contract and the full compliance contract.

2.4 Conclusion

This paper has shown that, when the principal has valuable private information affecting a potential audit, he faces an issue of credibility, even when he can commit to particular auditing probabilities. The primary benefit of contractual commitment to a particular auditing scheme is the ability to reliably audit *ex post* at a cost, which improves incentives *ex ante*. If the principal wishes to use the private signal, such commitment alone lacks the power to provide this reliability. Private information acts as a way for the principal to avoid his commitment in favor of cost savings.

This explains, in part, why auditing organizations who appear to have the proper tools for commitment do not formally commit to auditing. Renegotiation, transparency, technology and transactions costs seem to be lesser issues for many auditing cases, but private information prevents such auditors from being fully credible, leaving a complete lack of commitment and transparency as a second-best solution.

This result emphasizes the issues caused by the late timing of a private signal's arrival. If the signal arrived before transfers and the agent enjoyed a positive outside option, the principal could solve his incentive issues trivially. The private-information-based incentive issues faced by the principal in Khalil, Lawarrée and Scott (2013) are countervailing to the ones presented here and could easily balance each other out; the result would resemble their costly certification extension, but with no possibility of collusion.

While the literature on auditing often avoids issues of commitment, Bester and Strausz (2001) provide a convenient method for solving single-agent mechanism design problems where the principal has a wide variety of issues with credibility performing future actions. This paper takes advantage of their method in an applied way to show that one issue with credibility can prevent other commitment devices from providing their value.

Chapter 3

The Impact of Capital Punishment on Child Rape Offenses

3.1 Introduction

The vast literature on the deterrent effect of the death penalty¹⁶ primarily examines the effect of the death penalty on the rate of murder. In contrast, we study the effect of capital punishment for repeated child rape on the rate of child rape. The death penalty for child rape was first implemented in the U.S. in 1995, and continued in various states until it was ruled unconstitutional in 2008. We attempt to estimate the effect of this policy on incidences of child rape using crime panel data, state-level variation in implementation of the death penalty for child rape, information on the relationship between the victims, and the natural experiment of the Supreme Court's 2008 ruling striking down all the statutes.

Economic theories of crime beginning with Becker (1968) suggest that, all else equal, an increase in the level of punishment should decrease criminal activity by making it more costly. Using the victim-offender relationship to identify whether the offender is a stranger, an acquaintance or a family member, we estimate the effect of the death penalty on the rate of child rape at the state level. We find that the death penalty leads to a decrease in child rape reports where the offender is not a family member, which suggests a strong deterrent effect on these groups. We find no evidence of an effect for family members.

¹⁶ Cameron (1994) provides an extensive review of older articles, and Donohue and Wolfers (2006) provide a more recent critique of the literature.

The paper is structured as follows. In Section I, we briefly describe the existing economic literature on empirical studies on capital punishment. Section II provides the theoretical framework and the model predictions. In Section III, we present the identification strategy and our hypothesis. Section IV describes the data. The main estimates are presented in Section V. Section VI concludes.

3.2 Background

3.2.1 Evolution of Capital Punishment for Child Rape

In 1995, Louisiana became the first state in recent history¹⁷ to allow the death penalty for child rape offenses. Montana and Georgia followed Louisiana's child rape statute soon after. Table 1 contains the enacting time for states with this type of capital punishment. Kennedy Patrick, one of two men ever sentenced to death under Louisiana's statute allowing capital punishment for rape offenses of children under 12 years old, sought protection under the Eighth Amendment. In *Kennedy v. Louisiana*, the Supreme Court of the United States held such capital punishment unconstitutional. The sudden change in punishment provides a natural experiment in which we can estimate the effect of the death penalty on child rape reports.

3.2.2 Theories and Empirical Studies of Capital Punishment

In Becker's model of rational criminal activity, a potential criminal commits crimes only when the expected utility of committing the crime is higher than the expected utility of not committing the crime. Therefore, *ceteris paribus*, an increase in the expected sanction should reduce crime. This effect is called deterrence. In the past twenty years, a good deal of economic research has been devoted to estimating deterrence. Nevertheless the deterrent effect of capital punishment has been in hot debate since Sellin's (1959) and

¹⁷ Older articles such as Partington (1965) indicate that some states imposed the death penalty for rape and related offenses in the 1960s and earlier. Our analysis focuses on the more recent statutes.

Ehrlich's (1975) work negating and supporting the effect, respectively. Cameron (1994) provides a broad survey of the literature on the subject of capital punishment deterrence.

Economists face various challenges in the empirical analysis of crime. By its very nature crime is imperfectly observed. Reports of violent crime are estimated to comprise about 50% of total violent crime (Truman 2011). The difference between actual crime incidence and reported crime incidence causes measurement error in the observed crime variable. The actual crime rate will be higher than the reported rate, and the arrest rate (number of arrests per total reported crime) will be greater than arrests per total crime. An increase in crime reporting *ceteris paribus* increases the reported crime rate and decreases the reported arrest rate. This measurement error could have a downward bias on the estimated effect of arrest rate on crime rate. Levitt (1998b) shows the measurement error in reported crime in fact does not have substantial influence on the observed relationship between arrest rate and crime rate.

Allen (2007) studies the incentives for a rape victim to report his or her attacker. Using National Crime Survey from 1973 to 1983, Allen found that the rape victim is more likely to report if the offender is a stranger. The decision to report given the offender is a stranger does not affect the victim's access to social support. Allen's empirical work suggests a family offender reduces the victim's private social support and therefore the likelihood to report is smaller given the offender is a family member. While the reporting effect for family members is ambiguous, we expect no change in reporting behavior regarding strangers.

The general non-reporting rate for all sexual assaults is estimated as 65% for 2006-2010 based on the National Crime Victimization Survey (Langton et al. 2012). Of those non-reports, 28% either feared reprisal from the offender or did not want to get the offender in trouble. Levitt (1998a) demonstrated that an increase in the size of the police force had underestimated effects on property crime due to a change in reporting behavior. As the

police force becomes more effective, victims tend to report crimes more often, as they see an increase in the likelihood that their report will lead to a successful arrest.

Many studies of crime prevention suffer from the potential simultaneity of policy and crime. For instance, an increase in crime may draw attention from the government, leading to a harsher imposed law. A naïve estimate of the effect of such a policy would be biased downwards. However, sex-related laws are usually inspired by one horrible crime (Prescott and Rockoff 2011). That one crime, however, doesn't usually affect aggregate statistics. Because of this we argue that the imposition of capital punishment for child sex offenders can be treated as exogenous.

One issue specific to the death penalty is the lack of significant variation in execution rates. Several studies work around this issue by estimating deterrent effects of related independent variables with more variation. Katz et al. (2003) present evidence that the death rate in prison deters more criminal behavior than the death penalty. Mocan and Gittings (2003) estimate the effect of commuting prisoners from a death sentence. They theorize that the act of commutation decreases the subjective probability of being executed given conviction and therefore reduces the deterrent effect of the death penalty. This theory bears out in their estimation. Dezhbakhsh et al. (2003) use long panel data to overcome the issues of low variation.

This paper focuses on the variation in policy rather than actual death rates, as few repeat child sex offenders subject to the death penalty are actually sentenced to death. It is likely that the death penalty is used by a prosecutor to encourage a harsher plea bargain or sentence than would exist without the death penalty, so focusing on actual execution rates would underestimate the increase in punishment caused by the death penalty. This intuition is also useful when considering the incentives to report an offender; the death penalty implies an increase in the level of punishment, but not likely a state execution.

The common crime data include Uniform Crime Reports (UCR), National Crime Victimization Survey (NCVS), and National Incidence-Based Reporting System (NIBRS).

UCR does not separate offenses by age of victims. We cannot estimate how many child rape offenses were deterred by death penalty of child rapists. NCVS is a nationwide household survey. To keep confidentiality, the data is not available at the state level, so we cannot identify the effect of state-based policy changes with this data. Also, the respondent must be at least 12 years old which only covers the edge of our target sample population. NIBRS is comprehensive and incident-oriented. Law enforcement agencies participating NIBRS maintain a database of details in each reported criminal incident. This data set contains six segment levels: administrative, offense, property, victim, offender, and arrestee. South Carolina is the first state certified to report NIBRS to FBI in 1991. 32 states were certified by 2011. One of the disadvantages of NIBRS is not every law enforcement agency participates in the program; even for certified states, the percentage of population covered by NIBRS varies. In addition, NIBRS, like all other crime datasets maintained by police, provides only the reported crimes. Still, it is the best data set for our purposes. Further discussion of crime datasets can be found in Levitt (2007) and our data section below.

3.3 The Decision to Offend

Our model of criminal behavior is inspired by Becker (1968) and can be considered a special case. We simplify the potential offender's utility function to be constant in the utility of crime and linear (and negative) in punishment level. We simplify the state's punishment choice set to two levels, the "base" punishment level and the marginally changed punishment level. At first, we focus our analytical efforts on analyzing the potential offender's perception of his probability of being caught. We assume an individual considers his utility of committing a particular crime as follows:

$$E\{U_{i,c,y}(\text{Crime})\} = K_{i,c,y} - \alpha R_{c,y} \widehat{P}_{i,c,y}, \quad (1)$$

where i denotes individuals, c denotes the county, y denotes the year, K is individual i 's direct utility of committing the crime in year t , $R_{c,y}$ is the punishment imposed on arrested

criminals in state s in year t , $\widehat{P}_{i,c,y}$ is individual i 's estimation of the probability of arrest given he commits the crime in state s in year t , and $\alpha > 0$. If $K > 0$, then the criminal faces a tradeoff between the direct utility of criminal activity and the potential disutility of punishment if caught. Further, we assume the following:

$$R_{s,t} = Q + I_{c,y}, \quad (2)$$

where Q is the “normal” punishment for the crime and I indicates whether (1) or not (0) county c has imposed additional punishment D (in this case, the death penalty) in year y . The structure of R is imposed for simplicity; for now, all states punish the crime in the same way, and have a choice to impose a fixed amount of further punishment. Finally, we assume the following about P :

$$\widehat{P}_{i,c,y} = \tau_1 A_{c,y-1} + \eta_{i,c,y}, \quad (3)$$

where A is the state's previous year's arrest rate, and η is a random variable with mean μ_η and variance σ_η^2 , truncated such that $0 < \widehat{P}_{i,c,y} < 1$. One can interpret η as a combination of other unobserved factors, including error, in the individual's estimation of his probability of capture. We posit that the true probability of capture is positive even if the previous year's arrest rate was zero, which occurs in some county-years, and that the offender is aware of this. This implies that $E(\eta_{i,c,y}) > 0$. We expect the arrest rate to indicate police effectiveness, interpreted here as a positive τ_1 . If $\tau_1 < 1$, arrest rate indicates something about the police that decreases the offender's subjective probability of capture.

Plug (2) and (3) into (1) to get:

$$E\{U_{i,c,y}(\text{Crime})\} = K_{i,c,y} - \alpha * (Q + I_{c,y} * D) * (\tau_1 * A_{c,y-1} + \eta_{i,c,y}). \quad (4)$$

Individual i will only commit the crime when his expected utility of criminal activity is weakly positive:

$$K_{i,c,y} \geq \alpha(Q + I_{c,y}D)(\tau_1 A_{c,y-1} + \eta_{i,c,y}). \quad (5)$$

Now we assume for simplicity of exposition that, within a particular state, K is distributed uniformly from j to k , and j and k are “broad enough,” that is, some individuals will always refrain from committing the crime, and some individuals will always commit the crime. The probability individual i commits the crime is as follows:

$$P(\text{Crime}|s, t) = \left(k - \alpha(Q + I_{c,y}D)(\tau_1 A_{c,y-1} + \eta_{i,c,y}) \right) (k - j)^{-1}. \quad (6)$$

We can re-write (6) as follows:

$$P(\text{Crime}|s, t) = \left[\frac{k - \alpha Q \eta_{i,c,y}}{k - j} \right] - \left[\frac{\alpha D \eta_{i,c,y} + \alpha D^2}{k - j} \right] I_{c,y} - \left[\frac{\alpha Q \tau_1}{k - j} \right] A_{c,y-1} - \left[\frac{\alpha \tau_1 D}{k - j} \right] [I_{c,y} A_{c,y-1}] \quad (7)$$

The aggregated realizations of criminal activity become the per capita crime rate. We can use Equation (7) to create a simple estimation equation:

$$E(\text{Crime}_{s,t}) = c + \beta_0 I_{c,y} + \beta_1 A_{c,y-1} + \beta_2 I_{c,y} A_{c,y-1} + \varepsilon, \quad (8)$$

where ε is white noise and c is a constant. We can relax a strong assumption and allow other covariates to influence the potential criminal’s decision to commit the crime. As long as these variables enter the utility function (4) *directly* – local environmental effects directly influence the potential offender’s utility of crime irrespective of the punishment scheme – these are essentially a part of $K_{s,t}$ as previously constructed, and we can rewrite Equation (8) as follows:

$$E(\text{Crime}_{c,y}) = c + \delta_c + \gamma_y + \beta_0 I_{c,y} + \beta_1 A_{c,y-1} + \beta_2 I_{c,y} A_{c,y-1} + \Theta * \mathbf{X} + \varepsilon, \quad (9)$$

where \mathbf{X} is a vector of agency-monthly covariates, δ_s represents agency fixed effects, and γ_t represents a time fixed effects. A richer model where one potentially commits a crime

based on the *difference* in utility between “free” life and “caught” life may require additional interaction terms between X , $A_{c,y-1}$, and $I_{c,y}$, but for now we focus on the simpler model.

3.4 Identification Strategy and Hypotheses

Our simplified, linear model predicts that criminals weight the severity of a punishment by the likelihood of getting caught. We expect rational potential criminals to consider the expected severity of a crime as a function of the punishment and the probability of being caught. If the probability of capture is close to zero, then the effect of a change in the punishment scheme is going to be relatively low, no matter how drastic the punishment system is.

We expand our model to better fit the highly nonlinear count data, with a general form that appears as follows:

$$E(\text{crime}_{c,y}|\mathbf{X}) = f(\mathbf{X}; \beta) \quad (10)$$

In this case, the direct coefficient on the death penalty, β_0 , still represents the unobserved components in the offender’s estimated probability of capture, including any change in the probability the offender is caught given the change in policy due to reporting behavior. Additionally, β_2 , the coefficient of the interaction between the law and arrest rate, represents how the magnitude of the deterrent effect of the death penalty depends on the efficacy of the police force. We hypothesize that, under conditions where reporting behavior does not change with the level of punishment, $\beta_0 < 0$. In addition, insofar as arrest rate represents an effective police department, we hypothesize that $\beta_2 < 0$. Formally, our hypotheses are as follows:

$$H_0: \beta_0 = 0; \quad H_A: \beta_0 < 0$$

$$H_0: \beta_2 = 0; \quad H_A: \beta_2 < 0$$

Families can have different incentives to report, so we focus on testing our hypotheses in the case where the offender is a stranger or an acquaintance.

3.4 Data

We use the National Incident-Based Reporting System (NIBRS) data from 2003 to 2012. Every law enforcement agency participating in the NIBRS provides a detailed report of all recorded incidents in the agency's jurisdiction. As of February 2008, NIBRS covered about one quarter of the U.S. population, one quarter of the U.S. reported crime and 37% of law enforcement agencies.

Every incident may involve multiple victims and multiple offenses. If one offense falls into forcible rape or statutory rape, then we classify this incident as one rape offense. In the case of multiple victims, each victim may involve different offenses, e.g. robbery and assault. If more than one victim is involved in a rape offense, then we classify this incident as multiple rape offenses.

We classify three groups of the relationship between victim and offender: family, known, and stranger. This classification follows the definition in NIBRS¹⁸. Table 2 shows the mean of child rape offense rate by offender-victim relationships, average of child rape arrestees per 100,000 residents and exceptional clearance rates.

The average annual income per capita by county is obtained from the Bureau of Economic Analysis. The average annual unemployment rate by county is collected from the Bureau of Labor Statistics. The source of ethnicity and gender-age data by county is from the U.S. Census Bureau.

¹⁸ "Family" relationship includes immediate family, step family, child of boyfriend/girlfriend and extended family. "Known" relationship includes acquaintances, friends, neighbors, babysitters, boyfriend/girlfriend, or employees. "Stranger" relationship specifically indicates strangers. We include child of boyfriend/girlfriend in the family category because the incentive for the child's primary caretaker to report his or her significant other is similar to reporting a family member.

3.5 Empirical Results

3.5.1 Empirical Framework

We apply a negative binomial model with fixed effects to estimate the deterrent effect. Our data is over-dispersed, i.e. its variance is greater than its mean, due to the abundance of county years with zero reports of child rape¹⁹. We estimate the following equations:

$$Crime_{cy} = Crime(\lambda_{cy}, r) \quad (11)$$

$$\begin{aligned} \ln(\lambda_{cy}) = & \beta_0 + \beta_1 Law_{cy} + \beta_2 Arr_{cy-1} + \beta_3 Dec_{cy-1} \\ & + \beta_4 Ref_{cy-1} + \gamma \delta_y + \eta \delta_r + \mathbf{X}'\boldsymbol{\theta} + \varepsilon_{cy} \end{aligned} \quad (12)$$

where c indexes county and y indexes year. r is the over-dispersion parameter, and λ_{cy} is the expected count of child rape. Pop is the population of county c in year y . $Crime$ is the count of the child rape by county c in year y . Law is a dummy variable indicating the death penalty for child rape in a particular county year. Arr is the previous county year's count of arrests for child rape per 100,000 residents²⁰. Dec is the previous county year's count of the instances where the district attorney declined to prosecute a known suspect, per 100,000 residents. Ref is the previous year's count of the instances where an arrest could not be made because the victim refused to cooperate, per 100,000 residents. Arr , Dec , and Ref all predict to the offender's probability of capture. In the case where the count of crimes is zero for a particular county year, we use the previous year's arrest, decline, and refuse rates. In the case where the previous year's rates are not available, we set the rates to zero.

\mathbf{X} is a vector of covariates, including the non-child rape rate and demographic data such as the natural log of income per capita, unemployment rate, share of white people,

¹⁹ Refer to Figure 1 for an analysis of the distribution of the dependent variable.

²⁰ Arrests are invariably linked to crimes, so that the population-based arrest rate may be linked to any persistent criminality in a county not accounted for in the fixed effects. In an alternative specification, we follow Levitt (1998b) by using clearance rates, i.e. arrests per crime, as a measure of police effectiveness.

share of black people, and five age groups by gender. We include time fixed effects (δ_y) to control for the changes in crime over years and region fixed effects (δ_r) to control for heterogeneity within a geographical region. See Table 3 for a description of regions, as defined by the FBI's Uniform Crime Report.

3.5.2 Estimates of the Deterrent Effect of the Capital Punishment on Child Rape

Table 4 shows the main result. Implementing the death penalty reduces the reports of child rape by 37%²¹. Stranger reports are reduced by 67%, while acquaintance reports are reduced by 57%. The effect of the death penalty on family child rape reports is not statistically significant.

The positive effects of *Dec* and *Ref* present further evidence that probability of capture measures heavily into the decision to offend. If the district attorney declines to prosecute, this implies that either the case put together by the police department is too weak to pursue, or the prosecution only picks cases that are certain to win. Either way, the offender in a county with a high decline rate knows that his probability of being convicted is relatively low. In a similar fashion, if a particular county tends to have victims who refuse to cooperate, the probability of the offender ultimately being convicted decreases.

In table 5 we include an interaction effect between *Law* and *Arr*. We find no evidence that the arrest rate augments the deterrent effect of the death penalty in our main specification. The effect of the death penalty remains similar to that in our main specification.

3.5.3 Alternative Specifications and Robustness Checks

²¹ Note that the coefficients in Table 4 predict $\ln(\lambda)$, which is the log of the expected value of the count of crime. Hence we can interpret the coefficients similarly to those in a log-linear regression, which is reflected in more direct calculations of marginal effects. Refer to Table 9 and Table 10 for more detailed marginal effects.

Arrest rates are often set in terms of crimes to more accurately represent a police department's effectiveness in dealing with reported crimes. In addition, setting the arrest rate in terms of population may capture a persistent criminality in a particular county not explained by fixed effects, as arrests are correlated with crimes. We change *Arr*, *Dec*, and *Ref* to be measured per crime report as an alternative specification in Tables 6 and Table 7. Clearance rates may contain bias in the denominator due to measurement error in crime reporting. Levitt (1998b) demonstrates that the deterrent effect is the most likely explanation for the persistent negative relationship between clearance rates and property crimes, as opposed to incapacitation or measurement error in the denominator.

We find qualitatively similar results, except that we find significant negative coefficients in the interaction between *Law* and *Arr*. This suggests that a police department's ability to turn reported crimes into arrests may in fact augment the effect of the death penalty, as we hypothesized in Section III.

To ensure that offenders were not merely hiding their crimes more effectively in response to the death penalty, we tested two effects. First, we examined child rapes where the victim/offender relationship was missing or unknown to see if strangers and acquaintances could somehow commit the crime in a way that would limit the police's knowledge of the details. The results in column 1 of Table 8 show that this section of crime has an even larger deterrent effect. Second, as the punishment for a crime increases, the incentive to murder the victim increases. The victim is a potential witness, and murdering the witness has less of a marginal deterrent effect. We test to see whether the law led to an increase in child murder, and we found no effect, shown in column 2 of Table 8.

3.6 Conclusion

The Supreme Court's 2008 ruling striking down the death penalty for any offense where a life was not taken provided a natural experiment where punishment for repeated child rape was exogenously reduced. We find that capital punishment in the case of repeated child rape decreases the reports of overall child rape. So long as the incentive to

report a stranger or acquaintance does not change with the level of punishment, this suggests a large deterrent effect of the death penalty.

We also find evidence that police effectiveness has a significant impact on child rape, irrespective of the level of punishment. When we separate our analysis by victim-offender relationship, the deterrent effect holds for offenders who are not family members, but the results for family members are ambiguous.

While our results suggest a general effect for countries capable of implementing the death penalty, it is important to consider the varied benefits and costs of such a serious policy. We have covered some here – the death penalty deters strangers and acquaintances from committing child rape – but many other aspects of the death penalty, including lives lost, false positives, and other systemic effects of execution require serious consideration.

Table 3.1: Death Penalty Statutes for the Rape of a Child (before *Kennedy v. Louisiana*)

States	Offenses	Enacted	Notes
Texas	Second conviction for rape of a child under 14; first offense could have occurred prior to law's passage.	2007	Passed by legislature; signed by governor on July 16, 2007
Oklahoma	Rape or forcible sodomy of a victim under 14 where the defendant had a prior conviction of sexual abuse of a person under 14.	2006	
South Carolina	Repeat offenders of criminal sexual conduct with a minor under 11.	2006	
Montana	Second conviction for sexual intercourse without consent accompanied by serious bodily injury.	1997	Statute covers all ages of victims, but has specific provisions if the victim is 12 or under.
Louisiana	Aggravated rape of a child under 13.	1995	Statute upheld in <i>State v. Kennedy</i> on May 22, 2007; similar statute upheld in <i>State v. Wilson</i> , 685 So.2d 1063, 1073 (La. 1996) but defendant did not receive a death sentence. Victim under age 13, rather than 12, was added more recently (Acts 2003, 2006).
Georgia	Carnal knowledge of a female who is less than 10 presumes force.	1999	In 2006, Georgia's legislature revoked its general capital rape statute, but it is unclear whether the rape of a minor could be pursued as a capital crime.

Source: Death Penalty Information Center

Table 3.2: Summary Statistics for Reported Child Rape

	Mean	Std. Dev.
Offenses/ 100,000 people:		
Total	15.66	19.36
Committed by family members	5.15	9.7
Committed by acquaintance	8.08	12.94
Committed by strangers	0.28	1.6
Arrestees/ 100,000 people	1.25	6.14
Exceptional clearance:		
Prosecution declined/ 100,000 people	0.22	2.69
Victim refuse to cooperate/ 100,000 people	0.06	0.73

Note: The data unit is county-year.

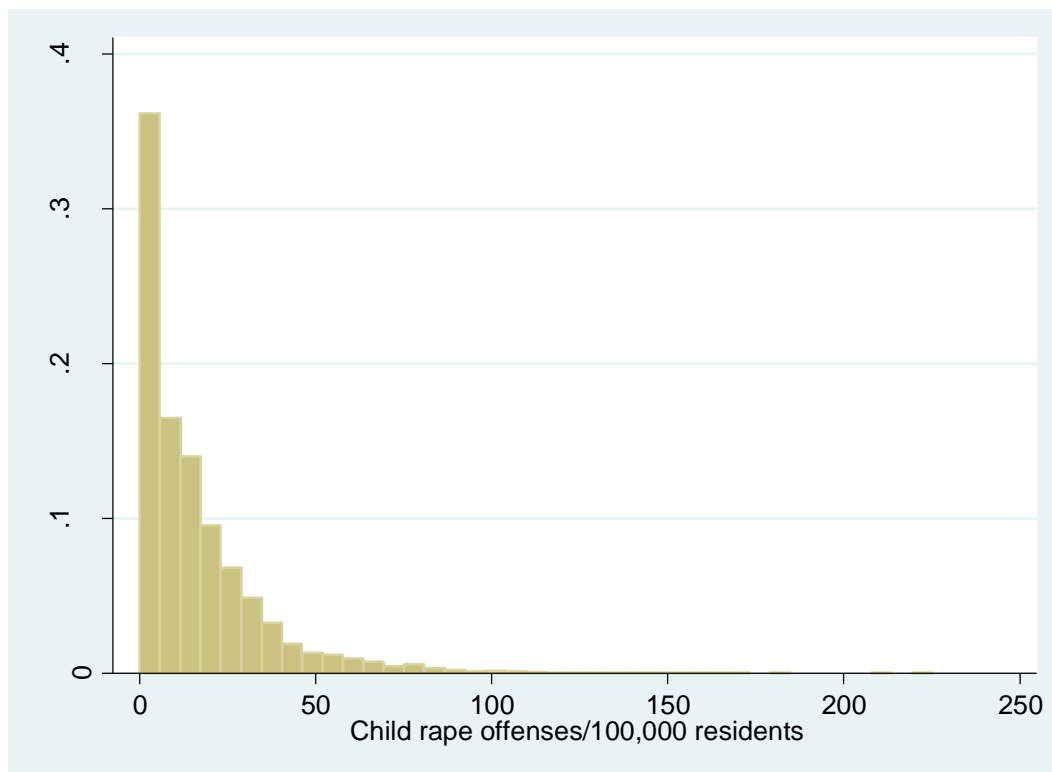


Figure 3.1: Histogram of Average Child Rape Rate by County-Year

Table 3.3: States and Counties in Sample

State	Division	Number of counties in sample	Sample period
Alabama	East South Central	7	2006-2012
Arizona	Mountain	7	2004-2012
Arkansas	West South Central	10	2003-2012
Colorado	Mountain	10	2003-2012
Connecticut	New England	10	2003-2012
Delaware	South Atlantic	10	2003-2012
Georgia	South Atlantic	5	2004-2008
Idaho	Mountain	9	2003-2012
Illinois	East North Central	7	2006-2012
Iowa	West North Central	10	2003-2012
Kansas	West North Central	10	2003-2012
Kentucky	East South Central	10	2003-2012
Louisiana	West South Central	8	2003-2012
Maine	New England	9	2004-2012
Massachusetts	New England	10	2003-2012
Michigan	East North Central	10	2003-2012
Mississippi	East South Central	4	2009-2012
Missouri	West North Central	4	2006-2012
Montana	Mountain	8	2005-2012
Nebraska	West North Central	7	2003-2012
New Hampshire	New England	10	2003-2012
North Dakota	West North Central	10	2003-2012
Ohio	East North Central	7	2003-2012
Oklahoma	West South Central	1	2008-2012
Oregon	Pacific	10	2003-2012
Rhodes Island	New England	9	2004-2012
South Carolina	South Atlantic	10	2003-2012
South Dakota	West North Central	10	2003-2012
Tennessee	East South Central	10	2003-2012
Texas	West South Central	2	2003-2012
Utah	Mountain	10	2003-2012
Vermont	New England	10	2003-2012
Virginia	South Atlantic	10	2003-2012
Washington	Pacific	3	2006-2012
West Virginia	South Atlantic	10	2003-2012
Wisconsin	East North Central	4	2004-2012

Table 3.4: The Deterrent Effect on Child Rape

	Total	Family	Known	Stranger
Law	-0.367*** (0.090)	0.136 (0.111)	-0.571*** (0.097)	-0.670*** (0.219)
Arrest rate (t-1)	-0.009*** (0.002)	-0.010*** (0.003)	-0.006** (0.003)	-0.027*** (0.010)
Prosecution declined rate(t-1)	0.094*** (0.012)	0.102*** (0.013)	0.093*** (0.013)	0.075*** (0.016)
Victim refuse to cooperate rate(t-1)	0.051*** (0.017)	0.048*** (0.017)	0.056*** (0.017)	0.076*** (0.026)

Note: Robust standard errors are in parentheses. The denominator of arrest rate or the exceptional clearance rate is 100,000 people. We control unemployment rate, log of income per capita, share of white people, share of black people, ratio of gender-age groups, non-child rape/100,000 people, division and year fixed effect. The number of observation is 9,517 and observations are at annual county level. *, **, *** denote the significance level at the 10%, 5%, 1%, respectively.

Table 3.5: The Deterrent Effect on Child Rape Including Interaction Term

	Total	Family	Known	Stranger
Law	-0.371*** (0.091)	0.135 (0.112)	-0.594*** (0.098)	-0.692*** (0.219)
Arrest rate (t-1)	-0.009*** (0.002)	-0.010*** (0.003)	-0.007** (0.003)	-0.027*** (0.010)
Law x Arrest rate (t-1)	0.011 (0.049)	0.005 (0.084)	0.051 (0.046)	0.069 (0.057)
Prosecution declined rate(t-1)	0.094*** (0.012)	0.102*** (0.013)	0.093*** (0.013)	0.075*** (0.016)
Victim refuse to cooperate rate(t-1)	0.051*** (0.017)	0.048*** (0.017)	0.057*** (0.017)	0.076*** (0.026)

Note: Robust standard errors are in parentheses. The denominator of arrest rate or the exceptional clearance rate is 100,000 people. We control unemployment rate, log of income per capita, share of white people, share of black people, ratio of gender-age groups, non-child rape/100,000 people, division and year fixed effect. The number of observation is 9,517 and observations are at annual county level. *, **, *** denote the significance level at the 10%, 5%, 1%, respectively.

Table 3.6: The Deterrent Effect on Child Rape with Different Arrest Measure

	Total	Family	Known	Stranger
Law	-0.411*** (0.094)	0.076 (0.116)	-0.621*** (0.101)	-0.726*** (0.221)
Arrest rate (t-1)	-0.241*** (0.086)	-0.447*** (0.073)	-0.156 (0.100)	-0.872*** (0.207)
Prosecution declined rate(t-1)	1.175*** (0.193)	1.282*** (0.205)	1.208*** (0.198)	1.264*** (0.269)
Victim refuse to cooperate rate(t-1)	0.564*** -0.411***	0.444* 0.076	0.742*** -0.621***	1.620*** -0.726***

Note: Robust standard errors are in parentheses. The denominator of arrest rate or the exceptional clearance rate is the number of child rape offenses. We control unemployment rate, log of income per capita, share of white people, share of black people, ratio of gender-age groups, non-child rape/100,000 people, division and year fixed effect. The observations are at annual county level. *, **, *** denote the significance level at the 10%, 5%, 1%, respectively.

Table 3.7: The Deterrent Effect on Child Rape with Different Arrest Measure and Interaction Term

	Total	Family	Known	Stranger
Law	-0.388*** (0.096)	0.088 (0.118)	-0.621*** (0.103)	-0.772*** (0.224)
Arrest rate (t-1)	-0.235*** (0.086)	-0.442*** (0.073)	-0.156 (0.101)	-0.898*** (0.212)
Law x Arrest rate (t-1)	-0.661 (0.573)	-0.339 (0.744)	0.011 (0.571)	1.387* (0.803)
Prosecution declined rate(t-1)	1.182*** (0.193)	1.286*** (0.206)	1.208*** (0.198)	1.250*** (0.267)
Victim refuse to cooperate rate(t-1)	0.562*** (0.201)	0.443* (0.230)	0.742*** (0.213)	1.622*** (0.393)

Note: Robust standard errors are in parentheses. The denominator of arrest rate or the exceptional clearance rate is the number of child rape offenses. We control unemployment rate, log of income per capita, share of white people, share of black people, ratio of gender-age groups, non-child rape/100,000 people, division and year fixed effect. The observations are at annual county level. *, **, *** denote the significance level at the 10%, 5%, 1%, respectively.

Table 3.8: Robustness checks

	(1)	(2)
	Child rape – missing or unknown	Child murder
Law	-0.905*** (0.205)	-0.083 (0.207)
Arrest rate(t-1) ^a	-0.023*** (0.006)	-0.071 (0.050)
Prosecution declined rate(t-1)	0.052*** (0.011)	
Victim refuse to cooperate rate(t-1)	0.038** (0.019)	

Note: ^a The arrest rate is the number of arrest for child murder divided by the number of child murder offenses in the second column. Robust standard errors are in parentheses. *, **, *** denote the significance level at the 10%, 5%, 1%, respectively.

Table 3.9: Marginal Effect of Death Penalty by Arrest Rate in Table 5

Arrest/100,000 people	(1) Total	(2) Family	(3) Known	(4) Stranger
1	-0.360*** (0.095)	0.140 (0.130)	-0.543*** (0.101)	-0.623*** (0.221)
2	-0.348*** (0.121)	0.145 (0.189)	-0.492*** (0.123)	-0.554** (0.236)
3	-0.337** (0.159)	0.150 (0.263)	-0.441*** (0.156)	-0.485* (0.263)
4	-0.325 (0.201)	0.155 (0.341)	-0.391** (0.194)	-0.417 (0.299)
5	-0.314 (0.246)	0.159 (0.422)	-0.340 (0.236)	-0.348 (0.340)
6	-0.303 (0.292)	0.164 (0.504)	-0.289 (0.279)	-0.279 (0.386)
7	-0.291 (0.340)	0.169 (0.587)	-0.239 (0.323)	-0.210 (0.434)
8	-0.280 (0.387)	0.174 (0.670)	-0.188 (0.368)	-0.142 (0.484)
9	-0.268 (0.435)	0.179 (0.753)	-0.137 (0.413)	-0.073 (0.536)
10	-0.257 (0.484)	0.184 (0.837)	-0.086 (0.458)	-0.004 (0.588)
11	-0.246 (0.532)	0.189 (0.921)	-0.036 (0.503)	0.064 (0.641)
12	-0.234 (0.581)	0.194 (1.005)	0.015 (0.549)	0.133 (0.695)
13	-0.223 (0.629)	0.199 (1.089)	0.066 (0.595)	0.202 (0.750)
14	-0.211 (0.678)	0.204 (1.173)	0.117 (0.641)	0.271 (0.804)

Note: Robust standard errors are in parentheses. *, **, *** denote the significance level at the 10%, 5%, 1%, respectively.

Table 3.10: Marginal Effect of Death Penalty by Arrest Rate in Table 7

Arrest/# of crimes	(1) Total	(2) Family	(3) Known	(4) Stranger
0.1	-0.454*** (0.099)	0.054 (0.125)	-0.620*** (0.107)	-0.633*** (0.219)
0.2	-0.520*** (0.130)	0.020 (0.169)	-0.619*** (0.137)	-0.495** (0.242)
0.3	-0.586*** (0.175)	-0.014 (0.229)	-0.618*** (0.181)	-0.356 (0.286)
0.4	-0.653*** (0.225)	-0.048 (0.295)	-0.617*** (0.230)	-0.217 (0.344)
0.5	-0.719** (0.279)	-0.082 (0.365)	-0.616** (0.283)	-0.078 (0.409)
0.6	-0.785** (0.333)	-0.116 (0.436)	-0.615* (0.337)	0.060 (0.479)
0.7	-0.851** (0.388)	-0.150 (0.508)	-0.614 (0.391)	0.199 (0.552)
0.8	-0.917** (0.444)	-0.184 (0.581)	-0.612 (0.447)	0.338 (0.626)
0.9	-0.983** (0.500)	-0.217 (0.654)	-0.611 (0.503)	0.476 (0.702)
1.0	-1.049* (0.557)	-0.251 (0.728)	-0.610 (0.559)	0.615 (0.779)

Note: Robust standard errors are in parentheses. *, **, *** denote the significance level at the 10%, 5%, 1%, respectively.

Bibliography

- Acemoglu, Daron, Michael Golosov, and Aleh Tsyvinski. 2006. "Markets versus Governments: Political Economy of Mechanisms." National Bureau of Economic Research.
- Acemoglu, Daron, and Thierry Verdier. 2000. "The Choice between Market Failures and Corruption." *American Economic Review*, 194–211.
- Ahlin, Christian, and Pinaki Bose. 2007. "Bribery, Inefficiency, and Bureaucratic Delay." *Journal of Development Economics* 84 (1): 465–86.
- Allen, W. David. 2007. "The Reporting and Underreporting of Rape." *Southern Economic Journal*, 623–41.
- Auriol, Emmanuelle. 2006. "Corruption in Procurement and Public Purchase." *International Journal of Industrial Organization* 24 (5): 867–85.
- Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." In *The Economic Dimensions of Crime*, 13–68. Springer.
- Bénabou, Roland, and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *The American Economic Review* 96 (5): 1652–78.
- Bester, Helmut, and Roland Strausz. 2001. "Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case." *Econometrica* 69 (4): 1077–98.
- Cameron, Samuel. 1994. "A Review of the Econometric Evidence on the Effects of Capital Punishment." *The Journal of Socio-Economics* 23 (1–2): 197–214.
- Che, Yeon-Koo. 1995. "Revolving Doors and the Optimal Tolerance for Agency Collusion." *The Rand Journal of Economics*, 378–97.
- Cooray, Arusha, and Friedrich Schneider. 2016. "Does Corruption Promote Emigration? An Empirical Examination." *Journal of Population Economics* 29 (1): 293–310.
- De Has, Hein. 2007. "Remittances, Migration and Social Development. A Conceptual Review of the Literature, Nueva York, United Nations." *Research Institute for Social Development (Social Policy and Development Programme Paper 34)*.

-
- Dezhbakhsh, Hashem, Paul H. Rubin, and Joanna M. Shepherd. 2003. "Does Capital Punishment Have a Deterrent Effect? New Evidence from Postmoratorium Panel Data." *American Law and Economics Review* 5 (2): 344–76.
- Dimant, Eugen, Tim Krieger, and Daniel Meierrieks. 2013. "The Effect of Corruption on Migration, 1985–2000." *Applied Economics Letters* 20 (13): 1270–74.
- Dincer, Oguzhan C., and Burak Gunalp. 2012. "Corruption and Income Inequality in the United States." *Contemporary Economic Policy* 30 (2): 283–92.
- Donohue III, John J., and Justin Wolfers. 2006. "Uses and Abuses of Empirical Evidence in the Death Penalty Debate." National Bureau of Economic Research.
- Ehrlich, Isaac. 1975. "The Deterrent Effect of Capital Punishment: A Question of Life and Death." *The American Economic Review* 65 (3): 397–417.
- Farlee, Mitchell A. 2010. "Disclosure and Secrecy in Employee Monitoring." *Journal of Management Accounting Research* 22 (1): 187–208.
- Figalli, Alessio, Young-Heon Kim, and Robert J. McCann. 2011. "When Is Multidimensional Screening a Convex Program?" *Journal of Economic Theory* 146 (2): 454–78.
- Finkle, Aaron, and Dongsoo Shin. 2007. "Conducting Inaccurate Audits to Commit to the Audit Policy." *International Journal of Industrial Organization* 25 (2): 379–89.
- Gupta, Sanjeev, Hamid Davoodi, and Rosa Alonso-Terme. 2002. "Does Corruption Affect Income Inequality and Poverty?" *Economics of Governance* 3 (1): 23–45.
- Katz, Lawrence, Steven D. Levitt, and Ellen Shustorovich. 2003. "Prison Conditions, Capital Punishment, and Deterrence." *American Law and Economics Review* 5 (2): 318–43.
- Khalil, Fahad. 1997. "Auditing without Commitment." *The RAND Journal of Economics*, 629–40.
- Khalil, Fahad, and Jacques Lawarree. 2001. "Catching the Agent on the Wrong Foot: Ex Post Choice of Monitoring." *Journal of Public Economics* 82 (3): 327–47.
- Khalil, Fahad, and Jacques Lawarree. 2006. "Incentives for Corruptible Auditors in the Absence of Commitment." *The Journal of Industrial Economics* 54 (2): 269–91.
- Khalil, Fahad, Jacques Lawarree, and Troy J. Scott. 2015. "Private Monitoring, Collusion, and the Timing of Information." *The RAND Journal of Economics* 46 (4): 872–90.

-
- Khalil, Fahad, Jacques Lawarrée, and Sungho Yun. 2010. "Bribery versus Extortion: Allowing the Lesser of Two Evils." *The Rand Journal of Economics* 41 (1): 179–98.
- Kim, Mike. 2010. *Escaping North Korea: Defiance and Hope in the World's Most Repressive Country*. Rowman & Littlefield.
- Kofman, Fred, and Jacques Lawarree. 1993. "Collusion in Hierarchical Agency." *Econometrica: Journal of the Econometric Society*, 629–56.
- Kofman, Fred, and Jacques Lawarrée. 1996. "On the Optimality of Allowing Collusion." *Journal of Public Economics* 61 (3): 383–407.
- Lambert-Mogiliansky, Ariane. 1998. "On Optimality of Illegal Collusion in Contracts." *Review of Economic Design* 3 (4): 303–28.
- Langton, Lynn, Marcus Berzofsky, Christopher P. Krebs, and Hope Smiley-McDonald. 2012. *Victimizations Not Reported to the Police, 2006-2010*. US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.
- Levitt, Steven D. 1998a. "The Relationship between Crime Reporting and Police: Implications for the Use of Uniform Crime Reports." *Journal of Quantitative Criminology* 14 (1): 61–81.
- . 1998b. "Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?" *Economic Inquiry* 36 (3): 353–72.
- Manelli, Alejandro M., and Daniel R. Vincent. 2007. "Multidimensional Mechanism Design: Revenue Maximization and the Multiple-Good Monopoly." *Journal of Economic Theory* 137 (1): 153–85.
- Mittendorf, Brian. 2008. "Infectious Ethics: How Upright Employees Can Ease Concerns of Tacit Collusion." *Journal of Law, Economics, and Organization* 24 (2): 356–70.
- Mocan, H. Naci, and R. Kaj Gittings. 2003. "Getting off Death Row: Commuted Sentences and the Deterrent Effect of Capital Punishment." *Journal of Law and Economics* 46 (2): 453–78.
- Mookherjee, Dilip, and Ivan Paak-Liang Png. 1995. "Corruptible Law Enforcers: How Should They Be Compensated?" *The Economic Journal*, 145–59.
- Olsen, Trond E., and Gaute Torsvik. 1998. "Collusion and Renegotiation in Hierarchies: A Case of Beneficial Corruption." *International Economic Review*, 413–38.

-
- Partington, Donald H. 1965. "Incidence of the Death Penalty for Rape in Virginia, The." *Wash. & Lee L. Rev.* 22: 43.
- Prescott, J.J., and Jonah E. Rockoff. 2011. "Do Sex Offender Registration and Notification Laws Affect Criminal Behavior?" *Journal of Law and Economics* 54 (1): 161–206. doi:10.1086/658485.
- Rochet, Jean-Charles, and Philippe Choné. 1998. "Ironing, Sweeping, and Multidimensional Screening." *Econometrica*, 783–826.
- Sellin, Thorsten, and American Law Institute. 1959. *The Death Penalty : A Report for the Model Penal Code Project of the American Law Institute*. Philadelphia: Executive Office, American Law Institute.
- Severinov, Sergei, and Raymond Deneckere. 2006. "Screening When Some Agents Are Nonstrategic: Does a Monopoly Need to Exclude?" *The RAND Journal of Economics* 37 (4): 816–40.
- Strausz, Roland. 1997. "Delegation of Monitoring in a Principal-Agent Relationship." *The Review of Economic Studies* 64 (3): 337–57.
- . 2006. "Interim Information in Long-Term Contracts." *Journal of Economics & Management Strategy* 15 (4): 1041–67.
- Tirole, Jean. 1999. "Incomplete Contracts: Where Do We Stand?" *Econometrica* 67 (4): 741–81.
- Truman, Jennifer L. 2011. *Criminal Victimization (2009)*. DIANE Publishing.
- Yun, Sungho. 2012. "Costs of Engaging in Corruption: Equilibrium with Extortion and Framing." *The BE Journal of Economic Analysis & Policy* 12 (1): 1935–1682.

Appendix 1A: KLY ($q = 0$)

Comparing the benchmark contracts to the results of KLY

The results contained in this section are mostly similar to those developed by Khalil, Lawarrée and Yun (2010), due to these papers sharing assumptions on auditing technology and the structure and timing of corruption. In both this benchmark model and KLY, allowing extortion is never optimal, and allowing bribery can be superior to deterring all corruption or avoiding auditing altogether.

Due to the differences in model construction, however, the results laid out above do bear some differences to KLY, in particular related to the Least-Cost Corruption-Proof (LCCP) contract. In KLY, auditing while deterring all corruption is only useful when the accuracy of the audit is strong enough, and the principal never achieves first best profits so long as the audit is not perfect. In this model, the opposite is true: The LCCP contract is always superior to forgoing auditing, and with a strong enough signal, the principal can achieve the first best.

In KLY's moral hazard model, the principal has a way to punish shirkers without auditing – namely, the principal can withhold payment unless the agent produces high output. This strategy is constrained primarily by the agent's limited liability, such that the principal employs a maximal deterrence scheme without the use of the signal. In other words, the principal sets wages to zero upon observing low output to incentivize high effort. In this setting, rewarding an agent who produced low output depending on the auditor's report can be costly, as expected transfers given the low output occurred must increase. Hence, the signal must be strong enough before the LCCP contract truly punishes shirking more than the second best.

In the case of this adverse selection model, the principal has fewer tools to screen or punish the agent. Without auditing, the principal's only observed outcome is the agent's report; output is not observed, nor will it vary after the contract is signed. Therefore, the

high type receives rent, and limited liability and individual rationality constraints for the high type are nonbinding. When the agent is caught red-handed by an auditor, her transfers to the principal, and hence her punishment, will increase relative to the second best, from $\theta_1 - c$ to θ_2 . Notably, this punitive measure has no effect on low type transfers. Because of this, even an auditor with a weak signal strictly improves incentives. Notice that this change makes the LCCP a more attractive contract in this setting, and yet allowing bribery will still be superior given the signal is accurate enough.

The moral hazard model in KLY requires three assumptions to hold in order for the principal to never achieve the first best: Shirking entails no cost to the agent, the agent is protected by limited liability, and the agent is risk averse. If shirking required some costly effort, even if it was small, the principal could threaten to withhold compensation for that effort in suspicious states of the world without violating limited liability. This would provide strong enough incentives to deter shirking for a strong, imperfect signal. Similarly, if the agent's limited liability were reduced, the principal could offer negative transfers for catching the agent red-handed, which would strongly incentivize the agent not to shirk. Finally, risk neutrality in the agent would allow the principal to use the second best contract to obtain first best profits, as the principal could withhold all transfers until a high output was realized without bearing costs of compensating for risk.

In this model, risk neutrality plays less of a role, as the principal has no tool to vary transfers in the second best, and in the LCCP contract all agents receive a static benefit on the equilibrium path. Additionally, the agent's limited liability is similar to that in KLY, in that the agent cannot lose money. However, in this model the agent faces a fixed cost of entering the relationship, c , which allows high types to face additional punitive measures when shirking. This makes the model more similar to the extension in KLY where the agent has a large outside option which must be compensated.

The differences in these models leads to the LCCP being more attractive and the first best being achievable. Hence, we focus our attention on the parameter space where

this model bears closest resemblance to KLY, as detailed in parameter assumptions (A1) through (A3).

Derivation:

Note that the principal can still set $w_{1,2} = t_{1,2} + w_{1,1} - t_{1,1}$ to make $T_M - t_{1,2} + w_{1,2} = 0$. Increasing $w_{1,2}$ and $t_{1,2}$ has no effect on net transfers except to slacken IC_2 , so the principal still sets $t_{1,2} = \theta_2$ and $w_{1,2} = \theta_2 + w_{1,1} - t_{1,1}$. This logic holds for *all derivations* of the contract, and hence this maximal deterrence strategy for catching the agent red handed, along with offering the auditor a large wage for catching the agent off the equilibrium path, will be assumed for all contracts analyzed in the future. We can then focus on the key choice variables: Namely, the transfers that will occur on the equilibrium path, and whether and how to allow bribery.

Case 2: $T_M = t_{1,1} - w_{1,1}$

Here, the principal solves the following problem:

$$\max_{t_i, w_i} f_1(t_{1,1} - w_{1,1}) + (1 - f_1)(t_2) \text{ s. t.}$$

$$IR_1: \theta_1 - c - pt_{1,1} - (1 - p) \left(t_{1,\emptyset} + \lambda(t_{1,1} - w_{1,1} - t_{1,\emptyset} + w_{1,\emptyset}) \right) \geq 0$$

$$IR_2: \theta_2 - c - t_2 \geq 0$$

$$IC_2: \theta_2 - c - t_2$$

$$\geq \theta_2 - c - p \left(t_{1,2} + \lambda(t_{1,1} - w_{1,1} - t_{1,2} + w_{1,2}) \right)$$

$$- (1 - p) \left(t_{1,\emptyset} + \lambda(t_{1,1} - w_{1,1} - t_{1,\emptyset} + w_{1,\emptyset}) \right)$$

$$NE_1: w_{1,1} \geq w_{1,\emptyset}$$

Re-stating constraints:

$$IR_1: \theta_1 - c - (p + (1 - p)\lambda)t_{1,1} - (1 - p)(1 - \lambda)t_{1,\emptyset} + (1 - p)\lambda(w_{1,1} - w_{1,\emptyset}) \geq 0$$

$$IC_2: u_2 \geq (1 - p)\theta_2 - c - (1 - p)(1 - \lambda)t_{1,\emptyset} - (1 - p)\lambda(t_{1,1} - w_{1,1} + w_{1,\emptyset})$$

Re-stating the maximization problem:

$$\max_{t_i, w_i} L = f_1(t_{1,1} - w_{1,1}) + (1 - f_1)(t_2)$$

$$+ \phi_1 \left(\theta_1 - c - (p + (1 - p)\lambda)t_{1,1} - (1 - p)(1 - \lambda)t_{1,\emptyset} + (1 - p)\lambda(w_{1,1} - w_{1,\emptyset}) \right)$$

$$+ \phi_2 \left(p\theta_2 - t_2 + (1 - p)(1 - \lambda)t_{1,\emptyset} + (1 - p)\lambda \left(t_{1,1} - (w_{1,1} - w_{1,\emptyset}) \right) \right)$$

$$+ \phi_3 (\theta_2 - c - t_2)$$

$$+ \phi_4 (w_{1,1} - w_{1,\emptyset})$$

$$+ \phi_5 (w_{1,\emptyset})$$

$$+ \phi_6 (\theta_1 - t_{1,\emptyset})$$

First Order Conditions:

$$[t_2]: (1 - f_1) - \phi_2 - \phi_3 = 0$$

$$[t_{11}]: f_1 - \phi_1(p + (1 - p)\lambda) + \phi_2(1 - p)\lambda = 0$$

$$[t_{1\emptyset}]: -\phi_1(1 - p)(1 - \lambda) + \phi_2(1 - p)(1 - \lambda) - \phi_6 = 0$$

$$[w_{11}]: -f_1 + \phi_1(1 - p)\lambda - \phi_2(1 - p)\lambda + \phi_4 = 0$$

$$[w_{1\emptyset}]: -\phi_1(1 - p)\lambda + \phi_2(1 - p)\lambda - \phi_4 + \phi_5 = 0$$

Solving:

$$\phi_5 = f_1; w_{1,\emptyset} = 0$$

$$\phi_1(1-p)\lambda - \phi_2(1-p)\lambda + \phi_4 = f_1$$

$$\phi_1(p + (1-p)\lambda) - \phi_2(1-p)\lambda = f_1$$

$$\phi_4 = \phi_1 p; w_{1,1} = 0$$

$$\phi_1 = \frac{f_1 + \phi_2(1-p)\lambda}{p + (1-p)\lambda}$$

$$\phi_6 = \phi_2(1-p)(1-\lambda) - \frac{f_1 + \phi_2(1-p)\lambda}{p + (1-p)\lambda} (1-p)(1-\lambda)$$

$$\phi_6 \frac{p + (1-p)\lambda}{(1-p)(1-\lambda)} = \phi_2(p + (1-p)\lambda) - f_1 - \phi_2(1-p)\lambda$$

$$\phi_6 \frac{p + (1-p)\lambda}{(1-p)(1-\lambda)} = \phi_2 p - f_1$$

$$\phi_2 = 1 - f_1 - \phi_3$$

$$(1 - f_1 - \phi_3)p - f_1 = \phi_6 \frac{p + (1-p)\lambda}{(1-p)(1-\lambda)}$$

$$p = \frac{f_1}{1 - f_1 - \phi_3} + \phi_6 \frac{p + (1-p)\lambda}{(1-p)(1-\lambda)}$$

If $\phi_6 > 0$, $t_{1,\emptyset} = \theta_1$. For this to be true,

$$p > \frac{f_1}{1 - f_1}$$

Which is assumed in (A3). If $p < \frac{f_1}{1-f_1}$, the principal prefers to implement the LCCP contract.

Re-stating IR_1 :

$$\theta_1 - c - (p + (1-p)\lambda)t_{1,1} - (1-p)(1-\lambda)t_{1,\emptyset} = 0$$

$$\theta_1(p + (1-p)\lambda) - t_{1,1}(p + (1-p)\lambda) - c = 0$$

$$t_{1,1} = \theta_1 - \frac{c}{p + (1-p)\lambda}$$

$$p\theta_2 - t_2 + (1-p)(1-\lambda)t_{1,\emptyset} + (1-p)\lambda(t_{1,1}) = 0$$

$$t_2 = p\theta_2 + (1-p)\theta_1 - \frac{(1-p)\lambda}{p + (1-p)\lambda}c$$

Full solution for $\phi_6 > 0$:

$$w_{1,1} = w_{1,\emptyset} = 0$$

$$t_{1,2} = \theta_2$$

$$w_{1,2} = \theta_2 - \theta_1 + \frac{c}{p + (1-p)\lambda}$$

$$t_{1,1} = \theta_1 - \frac{c}{p + (1-p)\lambda}$$

$$t_{1,\emptyset} = \theta_1$$

$$t_2 = p\theta_2 + (1-p)\theta_1 - \frac{(1-p)\lambda}{p + (1-p)\lambda}c$$

$$u_2 = (1-p)\Delta\theta - \frac{p}{p+(1-p)\lambda}c$$

This is the solution so long as

$$(1-p)\Delta\theta \geq \frac{p}{p+(1-p)\lambda}c$$

Which is true as long as (A2) holds.

If $\phi_3 > 0$, $\phi_6 = 0$, $t_{1,\emptyset} < \theta_1$. In this case $t_{1,\emptyset}$ is determined by the minimum amount it can be while still having a binding IR_2 . Solving:

$$\mathbf{u}_2 = \mathbf{0}; \mathbf{t}_2 = \boldsymbol{\theta}_2 - \mathbf{c}$$

$$(1-p)\theta_2 - c - (1-p)(1-\lambda)t_{1,\emptyset} - (1-p)\lambda t_{1,1} = 0$$

$$\theta_1 - c - (p+(1-p)\lambda)t_{1,1} - (1-p)(1-\lambda)t_{1,\emptyset} = 0$$

$$t_{1,1} = \frac{\theta_1 - c - (1-p)(1-\lambda)t_{1,\emptyset}}{p+(1-p)\lambda}$$

$$(1-p)\theta_2 - c - (1-p)(1-\lambda)t_{1,\emptyset} - (1-p)\lambda \frac{\theta_1 - c - (1-p)(1-\lambda)t_{1,\emptyset}}{p+(1-p)\lambda} = 0$$

$$(1-p)\theta_2 - c \left(1 - \frac{(1-p)\lambda}{p+(1-p)\lambda}\right) - t_{1,\emptyset}(1-p)(1-\lambda) \left(1 - \frac{(1-p)\lambda}{p+(1-p)\lambda}\right)$$

$$(1-p)\theta_2 - \frac{(1-p)\lambda}{p+(1-p)\lambda}\theta_1 - \frac{p}{p+(1-p)\lambda}c - (1-p)(1-\lambda)\frac{p}{p+(1-p)\lambda}t_{1,\emptyset} = 0$$

$$\frac{(1-p)(1-\lambda)p}{p+(1-p)\lambda}t_{1,\emptyset} = (1-p)\theta_2 - \frac{(1-p)\lambda}{p+(1-p)\lambda}\theta_1 - \frac{p}{p+(1-p)\lambda}c$$

$$t_{1,\emptyset} = \frac{(1-p)(p+(1-p)\lambda)}{(1-p)(1-\lambda)p} \theta_2 - \frac{(1-p)\lambda}{(1-p)(1-\lambda)p} \theta_1 - \frac{p}{(1-p)(1-\lambda)p} c$$

$$t_{1,\emptyset} = \frac{p+(1-p)\lambda}{p(1-\lambda)} \theta_2 - \frac{\lambda}{p(1-\lambda)} \theta_1 - \frac{c}{(1-p)(1-\lambda)}$$

$$t_{1,\emptyset} = \frac{p+(1-p)\lambda}{p(1-\lambda)} \Delta\theta + \theta_1 - \frac{c}{(1-p)(1-\lambda)}$$

$$t_{1,1} = \frac{\theta_1 - c - (1-p)(1-\lambda) \left(\frac{p+(1-p)\lambda}{p(1-\lambda)} \Delta\theta + \theta_1 - \frac{c}{(1-p)(1-\lambda)} \right)}{p+(1-p)\lambda}$$

$$t_{1,1} = \theta_1 - \frac{(1-p)(1-\lambda)}{p+(1-p)\lambda} \Delta\theta$$

$$w_{1,1} = w_{1,\emptyset} = 0$$

$$t_{1,2} = \theta_2$$

$$w_{1,2} = \frac{\Delta\theta}{p+(1-p)\lambda}$$

Case 2: $T_M = t_{1,\emptyset} - w_{1,\emptyset}$

$$\max_{t_i, w_i} f_1(t_{1,\emptyset} - w_{1,\emptyset}) + (1-f_1)(t_2) \text{ s.t.}$$

$$IR_1: \theta_1 - c - p \left(t_{1,1} + \lambda(t_{1,\emptyset} - w_{1,\emptyset} - t_{1,1} + w_{1,1}) \right) - (1-p)t_{1,\emptyset} \geq 0$$

$$IR_2: \theta_2 - c - t_2 \geq 0$$

$$IC_2: \theta_2 - c - t_2 \geq \theta_2(1-p) - c - (1-p)t_{1,\emptyset}$$

$$NE_1: w_{1,1} \geq w_{1,\emptyset}$$

Suppose NE_1 does not bind. Then the principal can freely decrease $w_{1,1}$, slackening IR_1 , until NE_1 does bind. Hence $w_{1,1} = w_{1,\emptyset}$. With that in mind, increasing $w_{1,\emptyset}$ only decreases the principal's profits, so the principal sets $w_{1,1} = w_{1,\emptyset} = 0$. $t_{1,1}$ only tightens IR_1 , so the principal will set it as low as possible while still maintaining $T_M = t_{1,\emptyset} - w_{1,\emptyset}$. This means $t_{1,1} = t_{1,\emptyset}$. This is the LCCP contract.

Case 3: $T_M = t_{1,2} - w_{1,2}$

$$\max_{t_i, w_i} f_1(t_{1,2} - w_{1,2}) + (1 - f_1)(t_2) \text{ s. t.}$$

$$\begin{aligned} IR_1: \theta_1 - c - p \left(t_{1,1} + \lambda(t_{1,2} - w_{1,2} - t_{1,1} + w_{1,1}) \right) \\ - (1 - p) \left(t_{1,\emptyset} + \lambda(t_{1,2} - w_{1,2} - t_{1,\emptyset} + w_{1,\emptyset}) \right) \geq 0 \end{aligned}$$

$$IR_2: \theta_2 - c - t_2 \geq 0$$

$$IC_2: \theta_2 - c - t_2 \geq \theta_2 - c - pt_{1,2} - (1 - p) \left(t_{1,\emptyset} + \lambda(t_{1,2} - w_{1,2} - t_{1,\emptyset} + w_{1,\emptyset}) \right)$$

$$NE_1: w_{1,1} \geq w_{1,\emptyset}$$

Holding $t_{1,1} - w_{1,1}$ constant, IR_1 slackens when the principal decreases $t_{1,1}$ and $w_{1,1}$. So the principal will decrease both until NE_1 binds and $w_{1,1} = w_{1,\emptyset}$. With that in mind, decreasing $t_{1,1}$ only slackens IR_1 , so the principal will decrease $t_{1,1}$ up until the point where $t_{1,2} - w_{1,2} = t_{1,1} - w_{1,1}$. With that in mind, the principal can set $t_{1,2} = \theta_2$ and $w_{1,1} = \theta_2 - t_{1,1} + w_{1,1}$ like in the previous two cases, and the principal can re-write the problem as follows:

$$\max_{t_i, w_i} f_1(t_{1,1} - w_{1,1}) + (1 - f_1)(t_2)$$

$$IR_1: \theta_1 - c - pt_{1,1} - (1 - p) \left(t_{1,\emptyset} + \lambda(t_{1,1} - w_{1,1} - t_{1,\emptyset} + w_{1,\emptyset}) \right) \geq 0$$

$$IR_2: \theta_2 - c - t_2 \geq 0$$

$$IC_2: \theta_2 - c - t_2 \geq \theta_2(1-p) - c - (1-p) \left(t_{1,\emptyset} + \lambda(t_{1,1} - w_{1,1} - t_{1,\emptyset} + w_{1,\emptyset}) \right)$$

$$NE_1: w_{1,1} \geq w_{1,\emptyset}$$

This is the same as Case 1.

Hence, we conclude that allowing bribery is superior to deterring all corruption, given the parameter constraints.

Appendix 1B: LCCP ($q > 0$)

Proof (LCCP) contract: In this case, all the CIC and NE constraints are included.

Re-stating the problem with relevant constraints:

$$\max_{t,w} f_1 \left(q(pT_{H1,1} + (1-p)T_{H1,\emptyset}) + (1-q)(pT_{S1,1} + (1-p)T_{S1,\emptyset}) \right) \\ + (1-f_1)(qt_{H2} + (1-q)t_{S2}) \text{ s. t.}$$

$$IR_{H1}: \theta_1 - c - pt_{H1,1} - (1-p)t_{H1,\emptyset} \geq 0$$

$$IR_{H2}: \theta_2 - c - t_{H2} \geq 0$$

$$IR_{S1}: \theta_1 - c - pt_{S1,1} - (1-p)t_{S1,\emptyset} \geq 0$$

$$IR_{S2}: \theta_2 - c - t_{S2} \geq 0$$

$$IC_{S1 \rightarrow H1}: \theta_1 - c - pt_{S1,1} - (1-p)t_{S1,\emptyset} \geq \theta_1 - c - pt_{H1,1} - (1-p)t_{H1,\emptyset}$$

$$IC_{S2 \rightarrow H1}: \theta_2 - c - t_{S2} \geq \theta_2 - c - pt_{H1,2} - (1-p)t_{H1,\emptyset}$$

$$IC_{S2 \rightarrow S1}: \theta_2 - c - t_{S2} \geq \theta_2 - c - pt_{S1,2} - (1-p)t_{S1,\emptyset}$$

$$IC_{S2 \rightarrow H2}: \theta_2 - c - t_{S2} \geq \theta_2 - c - t_{H2}$$

$$CIC_{H1,j,j'}: T_{H1,j} \geq T_{H1,j'}, j, j' \in \{1, 2, \emptyset\}$$

$$CIC_{S1,j,j'}: T_{S1,j} \geq T_{S1,j'}, j, j' \in \{1, 2, \emptyset\} NE_{H1,j}: w_{H1,j} \geq w_{H1,\emptyset}$$

$$NE_{S1,j}: w_{S1,j} \geq w_{S1,\emptyset}$$

$$NonNeg_{H1,j}: w_{H1,j} \geq 0$$

$$NonNeg_{S1,j}: w_{S1,j} \geq 0$$

$$LLC_{H1,2}: t_{H1,2} \leq \theta_2$$

$$LLC_{S1,2}: t_{S1,2} \leq \theta_2$$

$$LLC_{H1,\emptyset}: t_{H1,\emptyset} \leq \theta_1$$

$$LLC_{S1,\emptyset}: t_{S1,\emptyset} \leq \theta_1$$

$$LLC_{H1,1}: t_{H1,1} \leq \theta_1$$

$$LLC_{S1,1}: t_{S1,1} \leq \theta_1$$

$$LLC_{H2}: t_{H2} \leq \theta_2$$

$$LLC_{S2}: t_{S2} \leq \theta_2$$

We will ignore the following constraints as non-binding:

$$IC_{S1 \rightarrow H1}, IR_{S2}, IC_{S2 \rightarrow H2}, NE_{H1,2} \text{ and } NE_{S1,2}, NonNeg_{H1,2} \text{ and } NonNeg_{S1,2}$$

Also all the LLCs *except* $LLC_{H1,2}$, and $LLC_{S1,2}$

The CIC constraints can be re-written as follows:

$$t_{H1,\emptyset} - w_{H1,\emptyset} = t_{H1,1} - w_{H1,1} = t_{H1,2} - w_{H1,2}$$

$$t_{S1,\emptyset} - w_{S1,\emptyset} = t_{S1,1} - w_{S1,1} = t_{S1,2} - w_{S1,2}$$

Plugging in the CICs, the Langrangian is:

$$\begin{aligned}
L = & f_1 \left(q(t_{H1,1} - w_{H1,1}) + (1 - q)(t_{S1,1} - w_{S1,1}) \right) + (1 - f_1)(qt_{H2} + (1 - q)t_{S2}) \\
& + \phi_1 \left(\theta_1 - c - t_{H1,1} - (1 - p)(w_{H1,\emptyset} - w_{H1,1}) \right) \\
& + \phi_2(\theta_2 - c - t_{H2}) \\
& + \phi_3 \left(\theta_1 - c - t_{S1,1} - (1 - p)(w_{S1,\emptyset} - w_{S1,1}) \right) \\
& + \phi_4 \left(-t_{S2} + pt_{H1,2} + (1 - p) \left(t_{H1,1} + (w_{H1,\emptyset} - w_{H1,1}) \right) \right) \\
& + \phi_5 \left(-t_{S2} + pt_{S1,2} + (1 - p) \left(t_{S1,1} + (w_{S1,\emptyset} - w_{S1,1}) \right) \right) \\
& + \phi_6(w_{H1,1} - w_{H1,\emptyset}) \\
& + \phi_7(w_{S1,1} - w_{S1,\emptyset}) \\
& + \phi_8(w_{H1,1}) \\
& + \phi_9(w_{S1,1}) \\
& + \phi_{10}(\theta_2 - t_{H1,2}) \\
& + \phi_{11}(\theta_2 - t_{S1,2})
\end{aligned}$$

First order conditions:

$$[t_{H1,1}]: f_1 q - \phi_1 + \phi_4(1 - p) = 0$$

$$[t_{S1,1}]: f_1(1 - q) - \phi_3 + \phi_5(1 - p) = 0$$

$$[t_{H2}]: (1 - f_1)q - \phi_2 = 0$$

$$[t_{S2}]: (1 - f_1)(1 - q) - (\phi_4 + \phi_5) = 0$$

$$[t_{H1,2}]: \phi_4 p - \phi_{10} = 0$$

$$[t_{S1,2}]: \phi_5 p - \phi_{11} = 0$$

$$[w_{H1,1}]: -f_1 q + \phi_1(1 - p) - \phi_4(1 - p) + \phi_6 + \phi_8 = 0$$

$$[w_{H1,\emptyset}]: -\phi_1(1 - p) + \phi_4(1 - p) - \phi_6 = 0$$

$$[w_{S1,1}]: -f_1(1 - q) + \phi_3(1 - p) - \phi_5(1 - p) + \phi_7 + \phi_9 = 0$$

$$[w_{S1,\emptyset}]: -\phi_3(1 - p) + \phi_5(1 - p) - \phi_7 = 0$$

Along with the conditions on the constraints.

Solving:

$$\phi_1 = f_1 q + \phi_4(1 - p)$$

$$\phi_2 = (1 - f_1)q$$

$$\phi_3 = f_1(1 - q) + \phi_5(1 - p)$$

$$\phi_4 + \phi_5 = (1 - f_1)(1 - q)$$

$$\phi_{10} = \phi_4 p$$

$$\phi_{11} = \phi_5 p$$

$$\phi_6 = f_1 q + (\phi_4 - \phi_1)(1 - p) - \phi_8$$

$$\phi_6 = (\phi_4 - \phi_1)(1 - p)$$

$$\phi_7 = f_1(1 - q) + (\phi_5 - \phi_3)(1 - p) - \phi_9$$

$$\phi_7 = (\phi_5 - \phi_3)(1 - p)$$

$$\phi_8 = f_1 q$$

$$\phi_9 = f_1(1 - q)$$

$$w_{H1,1} = w_{S1,1} = w_{H1,\emptyset} = w_{S1,\emptyset} = 0$$

Given $\phi_4 > 0$ and $\phi_5 > 0$,

$$t_{H1,1} = t_{H1,\emptyset} = \theta_1 - c$$

$$t_{S1,1} = t_{S1,\emptyset} = \theta_1 - c$$

$$t_{H1,2} = t_{S1,2} = \theta_2$$

$$w_{H1,2} = w_{S1,2} = \Delta\theta + c$$

$$t_{H2} = \theta_2 - c$$

$$t_{S2} = p\theta_2 + (1 - p)(\theta_1 - c) = \theta_2 - c - (\Delta\theta(1 - p) - pc)$$

Note that the optimal solution satisfies all ignored constraints.

Consider the case where ϕ_4 or ϕ_5 were not greater than 0, i.e. $IC_{S2 \rightarrow S1}$ or $IC_{S2 \rightarrow H1}$ did not bind. First, if neither constraint was binding, the principal could increase t_{S2} without issue so long as the constraints we ignored do not bind instead. Double checking with simplified versions of the constraints:

$$IC_{S2 \rightarrow H2}: t_{H2} \geq t_{S2}$$

Because IR_{H2} binds, this is true so long as $t_{S2} \leq \theta_2 - c$, i.e. if S2 has any rent whatsoever.

The same holds for $IR_{S2}: t_{S2} \geq 0$

Now suppose $IC_{S2 \rightarrow S1}$ was binding, while $IC_{S2 \rightarrow H1}$ was not. That is to say, $\phi_4 = 0, \phi_5 > 0$. That would mean potentially two things: $t_{H1,2} > t_{S1,2}$, and/or $t_{H1,\emptyset} > t_{S1,\emptyset}$. This follows from the IC constraints, re-stated here:

$$IC_{S2 \rightarrow H1}: \theta_2 - c - t_{S2} \geq \theta_2 - c - pt_{H1,2} - (1-p)t_{H1,\emptyset}$$

$$IC_{S2 \rightarrow S1}: \theta_2 - c - t_{S2} \geq \theta_2 - c - pt_{S1,2} - (1-p)t_{S1,\emptyset}$$

First order conditions show that $LLC_{S1,2}$ is determined by $\phi_{11} = \phi_5 p$, so if $\phi_5 > 0, \phi_{11} > 0$, and that $LLC_{S1,2}$ binds. By $LLC_{H1,2}, t_{H1,2} \leq \theta_2$, so $t_{H1,2} \leq \theta_2 = t_{S1,2}$. Therefore, it must be that $t_{H1,\emptyset} > t_{S1,\emptyset}$. We know from first order conditions without any assumptions concerning ϕ_4 and ϕ_5 that $w_{H1,1} = w_{H1,\emptyset} = 0$ and $w_{S1,1} = w_{S1,\emptyset} = 0$. To maintain CICs, that means $t_{H1,\emptyset} = t_{H1,1}$ and $t_{S1,\emptyset} = t_{S1,1}$. Therefore, the principal could increase $t_{S1,1}$ and $t_{S1,\emptyset}$ equally, increase the objective function, and slacken $IC_{S2 \rightarrow S1}$ without harming any constraints. In particular, $IC_{S1 \rightarrow H1}$ requires that $t_{H1,\emptyset} \geq t_{S1,\emptyset}$, which is satisfied.

Now suppose $IC_{S2 \rightarrow H1}$ was binding, while $IC_{S2 \rightarrow S1}$ was not. That is, $\phi_4 > 0$ and $\phi_5 = 0$.

That would mean potentially two things: $t_{S1,2} > t_{H1,2}$, and/or $t_{S1,\emptyset} > t_{H1,\emptyset}$. This follows from the IC constraints, as before.

First order conditions show that $LLC_{H1,2}$ is determined by $\phi_{10} = \phi_4 p$, so if $\phi_4 > 0, \phi_{10} > 0$, and that $LLC_{H1,2}$ binds. By $LLC_{S1,2}, t_{S1,2} \leq \theta_2$, so $t_{S1,2} \leq \theta_2 = t_{H1,2}$. Therefore, it must be that $t_{S1,\emptyset} > t_{H1,\emptyset}$. We know from first order conditions without any assumptions concerning ϕ_4 and ϕ_5 that $w_{H1,1} = w_{H1,\emptyset} = 0$ and $w_{S1,1} = w_{S1,\emptyset} = 0$. To maintain CICs, that means $t_{H1,\emptyset} = t_{H1,1}$ and $t_{S1,\emptyset} = t_{S1,1}$. Therefore, the principal in this case could

increase $t_{H1,1}$ and $t_{H1,\emptyset}$ equally, increase the objective function, and slacken $IC_{S2 \rightarrow H1}$ without harming any constraints.

Appendix 1C: Allowing Bribery

Preliminary steps: possible corruption cases

(C1) *The coalition cannot gain by reporting or moving away from a signal of 2 through bribery. This implies that without loss of generality, we can assume there is no bribery after a signal of 2.*

Consider the maximand and constraints potentially related to a signal of 2:

$$f_1 \left(q(pT_{H1,1,F} + (1-p)T_{H1,\emptyset}) + (1-q)(T_{S1,M}) \right) + (1-f_1)(qt_{H2} + (1-q)t_{S2})$$

$$IC_{S2 \rightarrow H1}: \theta_2 - c - t_{S2} \geq \theta_2 - c - pt'_{H1,2} - (1-p)t'_{H1,\emptyset}$$

$$IC_{S2 \rightarrow S1}: \theta_2 - c - t_{S2} \geq \theta_2 - c - pt'_{S1,2} - (1-p)t'_{S1,\emptyset}$$

$$IR_{S1}: \theta_1 - c - pt'_{S1,1} - (1-p)t'_{S1,\emptyset} \geq 0$$

$$IC_{S1 \rightarrow H1}: \theta_1 - c - pt'_{S1,1} - (1-p)t'_{S1,\emptyset} \geq \theta_1 - c - pt'_{H1,1} - (1-p)t'_{H1,\emptyset}$$

Note that $t'_{i,2} = t_{i,2} + \lambda(T_{i,M} - T_{i,2}) = t_{i,2} + \lambda(t_{i,M} - w_{i,M} - t_{i,2} + w_{i,2})$.

$T_{i,2} = \min\{T_{i,1}, T_{i,\emptyset}\}$, where $i \in \{H1, S1\}$, which means that *the coalition cannot gain by reporting or moving away from a signal of 2 through bribery.*

If $T_{S1,2} < \min\{T_{S1,1}, T_{S1,\emptyset}\}$, then $t_{S1,1}$ is not in the objective function because bribery leads to a report of 2. Decreasing $t_{S1,1}$ only slackens IR_{S1} and $IC_{S1 \rightarrow H1}$, so the principal will do so until $T_{S1,2} = T_{S1,1}$, which means $T_{S1,2} = \min\{T_{S1,1}, T_{S1,\emptyset}\}$.

If $T_{S1,2} > \min\{T_{S1,1}, T_{S1,\emptyset}\}$, then $w_{S1,2}$ is not in the objective function. Increasing $w_{S1,2}$ only slackens $IC_{S2 \rightarrow S1}$, and so the principal will increase $w_{S1,2}$ until $T_{S1,2} = \min\{T_{S1,1}, T_{S1,\emptyset}\}$. In either case, the principal sets $T_{S1,2} = \min\{T_{S1,1}, T_{S1,\emptyset}\}$.

Next we will show that $T_{H1,2} = \min\{T_{H1,1}, T_{H1,\emptyset}\}$. Since honest types cannot bribe and truth-telling is induced, $T_{H1,2}$ never occurs in equilibrium.

If $T_{H1,2} < \min\{T_{H1,1}, T_{H1,\emptyset}\}$, then decreasing $w_{H1,2}$ only slackens $IC_{S2 \rightarrow H1}$ and $IC_{S1 \rightarrow H1}$. So the principal will decrease $w_{H1,2}$ until $T_{H1,2} = \min\{T_{H1,1}, T_{H1,\emptyset}\}$.

If $T_{H1,2} > \min\{T_{H1,1}, T_{H1,\emptyset}\}$, then increasing $w_{H1,2}$ only slackens $IC_{S2 \rightarrow H1}$. So the principal will increase $w_{H1,2}$ until $T_{H1,2} = \min\{T_{H1,1}, T_{H1,\emptyset}\}$.

Therefore, we have shown that $w_{i,2} = t_{i,2} - \min\{T_{i,1}, T_{i,\emptyset}\}$ for $i \in \{H1, S1\}$.

(C2) Maximal deterrence after a signal of 2: the principal will always set $t_{i,2} = \theta_2$.

Given that $T_{i,2} = \min\{T_{i,1}, T_{i,\emptyset}\}$, then increasing $t_{i,2}$ and $w_{i,2}$ such that $T_{i,2}$ is constant only slackens $IC_{S2 \rightarrow i}$ without directly affecting the objective function. Therefore, the principal will set $t_{i,2} = \theta_2$.

Note this also means that $t'_{i,2} = t_{i,2} = \theta_2$.

(C3) The IR constraints of honest types are binding.

Notice that honest types lack IC constraints by definition. This means that a transfer from a particular honesty type only ever increases the objective function, tightens its own IR constraint, and slackens other IC constraints. Therefore, if an IR of an honest type is slack, the principal can increase the honest type's transfers until it binds. Note that this means $t_{H2} = \theta_2 - c$, and $pt_{H1,1,F} + (1-p)t_{H1,\emptyset} = \theta_1 - c$.

(C4) Allowing extortion and framing is suboptimal

We prove this in steps.

Step 1 shows that allowing extortion/framing is suboptimal for H1, taking the following sub-steps:

Step 1a describes a general transfer scheme given framing occurs for H1.

Step 1b shows that, if framing occurs for H1, the principal is weakly better off setting $t_{H1,1} = t_{H1,\emptyset}$.

Step 1c describes a scheme that disallows framing which is strictly superior to the scheme in step 1b. Hence, allowing framing in H1 is suboptimal.

Step 2 shows that allowing extortion/framing is suboptimal for S1 using similar steps.

Step 1a

If $w_{i,\emptyset} > w_{i,1}$, then extortion/framing occurs, $t_{i,j,F} = t_{i,\emptyset}$ and $T_{i,j,F} = T_{i,\emptyset}$.

Additionally, when $w_{i,\emptyset} > w_{i,1}$, $t'_{i,j} = t_{i,\emptyset} + \lambda(T_{i,M} - T_{i,\emptyset})$.

Re-writing IR_{H1} given framing occurs in H1:

$$IR_{H1}: \theta_1 - c - t_{H1,\emptyset} \geq 0$$

H1's portion of the objective function:

$$T_{H1,\emptyset} = t_{H1,\emptyset} - w_{H1,\emptyset}$$

IC constraints which involve H1:

$$IC_{S1 \rightarrow H1}: \theta_1 - c - pt'_{S1,1} - (1-p)t'_{S1,\emptyset} \geq \theta_1 - c - pt'_{H1,1} - (1-p)t'_{H1,\emptyset}$$

$$IC_{S2 \rightarrow H1}: \theta_2 - c - t_{S2} \geq \theta_2 - c - p\theta_2 - (1-p)t'_{H1,\emptyset}$$

Note that IR_{H1} and the objective function are unaffected by $t_{H1,1}$.

Step 1b

First, suppose $T_{H1,M} = T_{H1,1}$. Then $t_{H1,1} < t_{H1,\emptyset}$, because $w_{H1,\emptyset} > w_{H1,1}$.

In this case, $t'_{H1,\emptyset} = t_{H1,\emptyset}(1-\lambda) + \lambda t_{H1,1} + \lambda(w_{H1,\emptyset} - w_{H1,1})$. Re-writing the IC constraints:

$$\begin{aligned} IC_{S1 \rightarrow H1}: & (p + (1-p)\lambda)t_{H1,1} + (1-p)\left((1-\lambda)t_{H1,\emptyset} + \lambda(w_{H1,\emptyset} - w_{H1,1})\right) \\ & \geq pt'_{S1,1} + (1-p)t'_{S1,\emptyset} \end{aligned}$$

$$IC_{S2 \rightarrow H1}: p\theta_2 + (1-p)\left((1-\lambda)t_{H1,\emptyset} + \lambda t_{H1,1} + \lambda(w_{H1,\emptyset} - w_{H1,1})\right) \geq t_{S2}$$

Notice that increasing $t_{H1,1}$ slackens both IC constraints so long as $T_{H1,M} = T_{H1,1}$. The principal can increase $t_{H1,1}$ and keep $T_{H1,M} = T_{H1,1}$ so long as $t_{H1,1} \leq t_{H1,\emptyset} - (w_{H1,\emptyset} - w_{H1,1})$.

Now, suppose $T_{H1,M} = T_{H1,\emptyset}$. Re-writing the IC constraints:

$$IC_{S1 \rightarrow H1}: t_{H1,\emptyset} \geq pt'_{S1,1} + (1-p)t'_{S1,\emptyset}$$

$$IC_{S2 \rightarrow H1}: p\theta_2 + (1-p)t_{H1,\emptyset} \geq t_{S2}$$

In this case, $t_{H1,1}$ has no effect whatsoever, and so setting it to $t_{H1,1} = t_{H1,\emptyset}$ does not harm the principal. Therefore, the principal weakly prefers to set $t_{H1,1} = t_{H1,\emptyset}$ when allowing framing to occur in H1.

The case of $T_{H1,M} = T_{H1,2}$ is equivalent to either the case where $T_{H1,M} = T_{H1,1}$ or $T_{H1,M} = T_{H1,\emptyset}$, because (C1) ensures that $T_{H1,2} = \min(T_{H1,1}, T_{H1,\emptyset})$.

Step 1c

Now consider a contract with equivalent transfers to those in Step 1b, but decrease $w_{H1,\emptyset}$ such that $w_{H1,\emptyset} = w_{H1,1}$, meaning framing is deterred. Because $t_{H1,1} = t_{H1,\emptyset}$, IR_{H1} remains the same:

$$pt_{H1,1} + (1-p)t_{H1,\emptyset} = t_{H1,\emptyset}$$

Re-writing IC constraints:

$$IC_{S1 \rightarrow H1}: \theta_1 - c - pt'_{S1,1} - (1-p)t'_{S1,\emptyset} \geq \theta_1 - c - pt'_{H1,1} - (1-p)t'_{H1,\emptyset}$$

Because $t_{H1,1} = t_{H1,\emptyset}$ and $w_{H1,1} = w_{H1,\emptyset}$, $t'_{H1,1} = t'_{H1,\emptyset} = t_{H1,\emptyset}$. Re-stating the IC:

$$\theta_1 - c - p't_{S1,1} - (1-p)t'_{S1,\emptyset} \geq \theta_1 - c - t_{H1,\emptyset}$$

Which is the same constraint as the one in Step 1b.

$$IC_{S2 \rightarrow H1}: \theta_2 - c - t_{S2} \geq \theta_2 - c - p\theta_2 - (1-p)t'_{H1,\emptyset} = \theta_2(1-p) - c - (1-p)t_{H1,\emptyset}$$

Which is the same constraint as the one in Step 1b.

Now consider H1's portion of the objective function:

$$pT_{H1,1} + (1-p)T_{H1,\emptyset} = t_{H1,\emptyset} - w_{H1,\emptyset}$$

Because $w_{H1,\emptyset}$ decreased, the objective function increased without affecting other constraints, relative to the weakly optimal framing scheme. Therefore, allowing framing in H1 is never optimal.

Step 2a

Note that, if extortion occurs, $t'_{S1,1} = t'_{S1,\emptyset} = t_{S1,\emptyset} + \lambda(T_{S1,M} - T_{S1,j}), j \in \{1, \emptyset\}$, and $w_{S1,\emptyset} > w_{S1,1}$.

Re-writing IR_{S1} and $IC_{S1 \rightarrow H1}$ given extortion/framing is allowed in S1:

$$IR_{S1}: \theta_1 - c - t'_{S1,\emptyset} \geq 0$$

$$IC_{S1 \rightarrow H1}: \theta_1 - c - t'_{S1,\emptyset} \geq \theta_1 - c - pt'_{H1,1} - (1-p)t'_{H1,\emptyset}$$

S1's portion of the objective function is $T_{S1,M}$.

IC constraints which involve S1:

$$IC_{S2 \rightarrow S1}: \theta_2 - c - t_{S2} \geq \theta_2 - c - p\theta_2 - (1-p)t'_{S1,\emptyset}$$

Step 2b

First, suppose $T_{S1,M} = T_{S1,1}$. In this case, $t'_{S1,\emptyset} = t_{S1,\emptyset}(1-\lambda) + \lambda t_{S1,1} + \lambda(w_{S1,\emptyset} - w_{S1,1})$.

Re-writing the IR and IC constraints related to S1:

$$IR_{S1}: \theta_1 - c - t_{S1,\emptyset}(1-\lambda) - \lambda t_{S1,1} - \lambda(w_{S1,\emptyset} - w_{S1,1}) \geq 0$$

$$IC_{S1 \rightarrow H1}: pt'_{H1,1} + (1-p)t'_{H1,\emptyset} \geq t_{S1,\emptyset}(1-\lambda) + \lambda t_{S1,1} + \lambda(w_{S1,\emptyset} - w_{S1,1})$$

$$IC_{S2 \rightarrow S1}: p\theta_2 + (1-p)\left((1-\lambda)t_{S1,\emptyset} + \lambda t_{S1,1} + \lambda(w_{S1,\emptyset} - w_{S1,1})\right) \geq t_{S2}$$

Increasing $t_{S1,1}$ tightens IR_{S1} and $IC_{S1 \rightarrow H1}$ while slackening $IC_{S2 \rightarrow S1}$. Increasing $t_{S1,1}$ and decreasing $t_{S1,\emptyset}$ so that $t'_{S1,\emptyset}$ remains the same will not affect these constraints. It will, however, increase $T_{S1,M} = T_{S1,1}$ by increasing $t_{S1,1}$. So the principal will at least increase $t_{S1,1}$ and decrease $t_{S1,\emptyset}$ until $T_{S1,1} = T_{S1,\emptyset}$.

Now suppose $T_{S1,M} = T_{S1,\emptyset}$. In this case, $t'_{S1,\emptyset} = t_{S1,\emptyset}$. Increasing $t_{S1,1}$ in this case has no effect on any constraints or the objective function, so it is weakly optimal to set $t_{S1,\emptyset} = t_{S1,1}$ in the case where extortion is allowed. Also notice that, if $t_{S1,\emptyset} = t_{S1,1}$ and $w_{S1,\emptyset} > w_{S1,1}$, $T_{S1,\emptyset} < T_{S1,1}$.

The case of $T_{S1,M} = T_{S1,2}$ is equivalent to one of the prior cases due to (C1) setting $T_{S1,2} = T_{S1,j} = \min(T_{S1,1}, T_{S1,\emptyset})$.

Re-writing the constraints where $t_{S1,1} = t_{S1,\emptyset}$:

$$IR_{S1}: \theta_1 - c - t_{S1,\emptyset} \geq 0$$

$$IC_{S1 \rightarrow H1}: pt'_{H1,1} + (1-p)t'_{H1,\emptyset} \geq t_{S1,\emptyset}$$

$$IC_{S2 \rightarrow S1}: p\theta_2 + (1-p)t_{S1,\emptyset} \geq t_{S2}$$

Step 2c

Now consider a contract with equivalent transfers to those in Step 2b, but decrease $w_{S1,\emptyset}$ such that $w_{S1,\emptyset} = w_{S1,1}$, meaning extortion/framing is deterred. Because $t_{S1,1} = t_{S1,\emptyset}$, IR_{S1} and $IC_{S1 \rightarrow H1}$ remain the same:

$$t'_{S1,1} = t'_{S1,\emptyset} = t_{S1,\emptyset}$$

Re-writing IR and IC constraints:

$$IC_{S1 \rightarrow H1}: \theta_1 - c - t_{S1,\emptyset} \geq \theta_1 - c - pt'_{H1,1} - (1-p)t'_{H1,\emptyset}$$

$$IR_{S1}: \theta_1 - c - t_{S1,\emptyset} \geq 0$$

$$IC_{S2-S1}: p\theta_2 + (1-p)t_{S1,\emptyset} \geq t_{S2}$$

These are the same constraints as those in Step 2b.

Now consider S1's portion of the objective function:

$$T_{S1,M} = T_{S1,1} = T_{S1,\emptyset} = t_{S1,\emptyset} - w_{S1,\emptyset}$$

Because $w_{S1,\emptyset}$ decreased relative to the contract in Step 2b, the objective function increased without affecting other constraints. Therefore, allowing framing in H1 is never optimal.

(C8) $T_{i,1} \leq T_{i,\emptyset}$ and $w_{i,1} = w_{i,\emptyset} = 0$, and therefore $t_{i,1} \leq t_{i,\emptyset}$.

If bribery is allowed for a particular type, the minimum net transfer for a particular low type will always occur on a signal of 1, and auditor wages on the equilibrium path will be set to 0.

If $T_{S1,M} = T_{S1,\emptyset}$, then decreasing $t_{S1,1}$ has no effect except to slacken IR_{S1} and $IC_{S1 \rightarrow H1}$. Therefore, the principal will decrease $t_{S1,1}$ until $T_{S1,1} = T_{S1,\emptyset}$. This, along with (C1), effectively satisfies all CIC constraints for S1, so corruption is deterred. The principal can then increase both $T_{S1,1}$ and $T_{S1,\emptyset}$ by decreasing $w_{S1,1}$ and $w_{S1,\emptyset}$ equally until $w_{S1,\emptyset} = 0$. At that point, if $w_{S1,1} > w_{S1,\emptyset}$, the principal can decrease $t_{S1,1}$ and $w_{S1,1}$ equally without affecting constraints. Therefore, the principal will set $w_{S1,1} = w_{S1,\emptyset} = 0$, and $t_{S1,1} = t_{S1,\emptyset}$, deterring all bribery by S1.

If $T_{S1,M} = T_{S1,1}$, increasing $w_{S1,\emptyset}$ and $t_{S1,\emptyset}$ equally such that $T_{S1,\emptyset}$ remains constant only serves to slacken $IC_{S2 \rightarrow S1}$. Increasing $w_{S1,\emptyset}$ causes $NE_{S1,1}$ to bind, such that $w_{S1,\emptyset} = w_{S1,1}$. In this case, decreasing both $w_{S1,\emptyset}$ and $w_{S1,1}$ proportionally has no effect on incentives and increases the objective function, so the principal will set $w_{S1,1} = w_{S1,\emptyset} = 0$.

If $T_{H1,M} = T_{H1,\emptyset}$, if $w_{H1,\emptyset} > 0$, decreasing $w_{H1,\emptyset}$ increases the objective function and slackens $IC_{S1 \rightarrow H1}$, so the principal will set it to 0. Decreasing $w_{H1,1}$ and $t_{H1,1}$ equally, holding $T_{H1,1}$ constant, slackens IR_{H1} , while doing nothing else, so the principal will set $w_{H1,1}$ to 0.

Re-writing $IC_{S1 \rightarrow H1}$:

$$\theta_1 - c - pt'_{S1,1} - (1-p)t'_{S1,\emptyset} \geq \theta_1 - c - p(t_{H1,1} - \lambda(t_{H1,\emptyset} - t_{H1,1})) - (1-p)t_{H1,\emptyset}$$

$$\theta_1 - c - pt'_{S1,1} - (1-p)t'_{S1,\emptyset} \geq \theta_1 - c - (p - p\lambda)t_{H1,1} - (1 - p + p\lambda)t_{H1,\emptyset}$$

Re-writing $IC_{S2 \rightarrow H1}$:

$$\theta_2 - c - t_{S2} \geq \theta_2(1-p) - c - (1-p)t_{H1,\emptyset}$$

Noting IR_{H1} and H1's portion of the objective function:

$$IR_{H1}: pt_{H1,1} + (1-p)t_{H1,\emptyset} = \theta_1 - c$$

$$\Pi|H1 = pt_{H1,1} + (1-p)t_{H1,\emptyset}$$

Re-writing the right hand side of $IC_{S1 \rightarrow H1}$:

$$p\lambda(t_{H1,1} - t_{H1,\emptyset})$$

Decreasing $t_{H1,1}$ and increasing $t_{H1,\emptyset}$ to maintain IR_{H1} slackens $IC_{S1 \rightarrow H1}$ and $IC_{S2 \rightarrow H1}$ without affecting the objective function. Therefore, the principal will set $t_{H1,\emptyset} = t_{H1,1}$ and $w_{H1,\emptyset} = w_{H1,1} = 0$, deterring all bribery by anyone mimicking H1.

If $T_{H1,M} = T_{H1,1}$, decreasing $w_{H1,1}$ increases the objective function and slackens $IC_{S1 \rightarrow H1}$ and $IC_{S2 \rightarrow H1}$, so the principal will decrease $w_{H1,1}$ until $NE_{H1,1}$ binds. Since $w_{H1,1} = w_{H1,\emptyset}$,

these wages serve no incentive effect and decreasing them to 0 will only increase the objective function.

In sum, when $T_{i,M} = T_{i,\emptyset}$, the principal will set $T_{i,1} = T_{i,2} = T_{i,\emptyset}$, which means bribery will only possibly occur when $T_{i,M} = T_{i,1}$, or equivalently $T_{i,M} = T_{i,2}$.

Optimal contract if all types are included

Problem and solution (Cases 1-9)

Then the maximization problem becomes:

$$\begin{aligned} \max_{t,w} f_1 & (q(pt_{H1,1} + (1-p)t_{H1,\emptyset}) + (1-q)t_{S1,1}) \\ & + (1-f_1)(q(\theta_2 - c) + (1-q)t_{S2}) \text{ s. t.} \end{aligned}$$

$$IR_{H1}: \theta_1 - c - pt_{H1,1} - (1-p)t_{H1,\emptyset} = 0$$

$$\begin{aligned} IC_{S1 \rightarrow H1}: \theta_1 - c - (p + (1-p)\lambda)t_{S1,1} - (1-p)(1-\lambda)t_{S1,\emptyset} \\ \geq \theta_1 - c - (p + (1-p)\lambda)t_{H1,1} - (1-p)(1-\lambda)t_{H1,\emptyset} \end{aligned}$$

$$IC_{S2 \rightarrow S1}: \theta_2 - c - t_{S2} \geq \theta_2 - c - p\theta_2 - (1-p)(\lambda t_{S1,1} + (1-\lambda)t_{S1,\emptyset})$$

$$IC_{S2 \rightarrow H1}: \theta_2 - c - t_{S2} \geq \theta_2 - c - p\theta_2 - (1-p)(\lambda t_{H1,1} + (1-\lambda)t_{H1,\emptyset})$$

$$LLC_{S1,\emptyset}: \theta_1 \geq t_{S1,\emptyset}$$

$$LLC_{H1,\emptyset}: \theta_1 \geq t_{H1,\emptyset}$$

$$t_{H1,1} \leq t_{H1,\emptyset}$$

$$t_{S1,1} \leq t_{S1,\emptyset}$$

We ignore the other constraints. It can be verified later that they are satisfied by the optimal contract.

The Lagrangian for the maximization problem:

$$\begin{aligned}
L = & f_1(q(pt_{H1,1} + (1-p)t_{H1,\emptyset}) + (1-q)t_{S1,1}) + (1-f_1)(q(\theta_2 - c) + (1-q)t_{S2}) \\
& + \phi_1(\theta_1 - c - pt_{H1,1} - (1-p)t_{H1,\emptyset}) \\
& + \phi_2(-(p + (1-p)\lambda)t_{S1,1} - (1-p)(1-\lambda)t_{S1,\emptyset} + (p + (1-p)\lambda)t_{H1,1} \\
& + (1-p)(1-\lambda)t_{H1,\emptyset}) \\
& + \phi_3(\theta_2 p - t_{S2} + (1-p)(\lambda t_{S1,1} + (1-\lambda)t_{S1,\emptyset})) \\
& + \phi_4(\theta_2 p - t_{S2} + (1-p)(\lambda t_{H1,1} + (1-\lambda)t_{H1,\emptyset})) \\
& + \phi_5(\theta_1 - t_{S1,\emptyset}) \\
& + \phi_6(\theta_1 - t_{H1,\emptyset}) \\
& + \phi_7(t_{H1,\emptyset} - t_{H1,1}) \\
& + \phi_8(t_{S1,\emptyset} - t_{S1,1})
\end{aligned}$$

where, $\phi_i \geq 0$ are the non-negative multipliers associated with $R_{H1}, IC_{S1 \rightarrow H1}, IC_{S2 \rightarrow S1}, IC_{S2 \rightarrow H1}, LLC_{S1,\emptyset}, LLC_{H1,\emptyset}$, and the $t_{i,1} \leq t_{i,\emptyset}$ constraints respectively.

First order conditions:

$$(1) \quad \frac{\partial L}{\partial t_{H1,1}} = f_1 qp - \phi_1 p + \phi_2(p + (1-p)\lambda) + \phi_4(1-p)\lambda - \phi_7 = 0$$

$$(2) \quad \frac{\partial L}{\partial t_{H1,\emptyset}} = f_1 q(1-p) - \phi_1(1-p) + \phi_2(1-p)(1-\lambda) + \phi_4(1-p)(1-\lambda) - \phi_6 + \phi_7 = 0$$

$$(3) \quad \frac{\partial L}{\partial t_{S1,1}} = f_1(1-q) - \phi_2(p + (1-p)\lambda) + \phi_3(1-p)\lambda - \phi_8 = 0$$

$$(4) \quad \frac{\partial L}{\partial t_{S1,\emptyset}} = -\phi_2(1-p)(1-\lambda) + \phi_3(1-p)(1-\lambda) - \phi_5 + \phi_8 = 0$$

$$(5) \quad \frac{\partial L}{\partial t_{S2}} = (1-f_1)(1-q) - \phi_3 - \phi_4 = 0$$

Notes:

$\phi_6 > 0$ and $\phi_7 > 0$ violates IR_{H1}

$$t_{H1,1} = t_{H1,\emptyset} = \theta_1$$

$$\theta_1 - c - \theta_1 < 0$$

$\phi_5 > 0$ and $\phi_8 > 0$ violates $IC_{S1 \rightarrow H1}$

$$\theta_1 - c - \theta_1 \geq \theta_1 - c - p t_{H1,1} - (1-p)(\lambda t_{H1,1} + (1-\lambda)t_{H1,\emptyset})$$

$$\theta_1 \leq p t_{H1,1} + (1-p)(\lambda t_{H1,1} + (1-\lambda)t_{H1,\emptyset})$$

Using $t_{H1,\emptyset} \geq t_{H1,1}$, IR_{H1} , $t_{H1,1} \leq \theta_1 - c$ and also, $t_{H1,\emptyset} \leq \theta_1$,

$$\theta_1 \leq p t_{H1,1} + (1-p)(\lambda t_{H1,1} + (1-\lambda)t_{H1,\emptyset}) < \theta_1$$

Solving:

$$\phi_3 = (1-f_1)(1-q) - \phi_4$$

$$f_1 q - \phi_1 + \phi_2 + \phi_4(1-p) = \phi_6$$

$$f_1(1-q) - \phi_2 + \phi_3(1-p) = \phi_5$$

$$f_1(1-q) - \phi_2 + \phi_4(1-f_1)(1-q)(1-p) = \phi_5$$

$$f_1 - \phi_1 + (\phi_4 + \phi_3)(1-p) = (\phi_5 + \phi_6)$$

$$\phi_1 = f_1 + (1-f_1)(1-q)(1-p) - (\phi_5 + \phi_6)$$

Case 1: $\phi_5 > 0, \phi_6 > 0, \phi_7 = \phi_8 = 0$ (main case)

By (1), $\phi_1 > 0$

By (3), $\phi_2 > 0$

By (4), $\phi_3 > 0$

$$f_1 q + \phi_2 \frac{p+(1-p)\lambda}{p} + \phi_4 \frac{(1-p)\lambda}{p} = \phi_1, \text{ from (1)}$$

$$f_1 q + \phi_2(1-\lambda) + \phi_4(1-\lambda) - \frac{\phi_6}{1-p} = \phi_1, \text{ from (2).}$$

$\phi_2 \frac{p+(1-p)\lambda}{p} + \phi_4 \frac{(1-p)\lambda}{p} = \phi_2 \frac{p(1-\lambda)}{p} + \phi_4 \frac{p(1-\lambda)}{p} - \frac{\phi_6}{1-p}$, combining the above two conditions. Rewriting:

$$\phi_2 \frac{\lambda}{p} + \frac{\phi_6}{1-p} = \phi_4 \frac{p-\lambda}{p}$$

For this case to be valid, $p > \lambda$ and $\phi_4 > 0$. If $p \leq \lambda$, we are not in case 1.

Solution for case 1 (assuming $p > \lambda$):

$$t_{S1,\emptyset} = t_{H1,\emptyset} = \theta_1$$

$$t_{S1,1} = t_{H1,1} = \theta_1 - \frac{c}{p}$$

$$\begin{aligned} t_{S2} &= \theta_2 - c - \left(\Delta\theta(1-p) - \frac{p - (1-p)\lambda}{p} c \right) \\ &= \theta_2 - c - \left(\Delta\theta(1-p) - c + \frac{(1-p)}{p} \lambda c \right) \end{aligned}$$

Case 2: $\phi_7 > 0, \phi_8 > 0, \phi_6 = \phi_5 = 0$ (LCCP)

$$\phi_7 > 0 \Rightarrow t_{S1,1} = t_{S1,\emptyset}$$

$$\phi_8 > 0 \Rightarrow t_{H1,1} = t_{H1,\emptyset}$$

Since $\phi_1 = f_1 + (1 - f_1)(1 - q)(1 - p) > 0$, IR_{H1} binds.

Since $\phi_2 > 0$ by (4), we can rewrite

$$IC_{S1 \rightarrow H1}: \theta_1 - c - t_{S1,1} = \theta_1 - c - t_{H1,1}$$

$$\text{So } t_{S1,1} = t_{S1,\emptyset} = t_{H1,1} = t_{H1,\emptyset} = \theta_1 - c$$

Because the transfers are identical, the RHS of both IC_{S2} s are identical. Hence, both constraints bind, and $\phi_3 > 0$ and $\phi_4 > 0$. Otherwise, the principal can increase t_{S2} and gain.

This is the LCCP contract.

Cases 3-8 will lead to contradictions, in large part because the low types should be pooled. First, leaving either $IC_{S2 \rightarrow S1}$ or $IC_{H2 \rightarrow H1}$ slack leaves room for the principal to reduce either S1 rent or overall bribery without affecting other incentives. Second, if both are binding

and the low types are not pooled, the principal will have an opportunity to increase expected transfers in one or another type without harming incentives by pooling them. Case 3: $\phi_8 > 0, \phi_6 > 0, \phi_7 = \phi_5 = 0$ (contradiction)

$$t_{H1,\emptyset} = \theta_1$$

$$t_{S1,\emptyset} = t_{S1,1}$$

Due to binding IR_{H1} , $t_{H1,1} = \theta_1 - \frac{c}{p}$

Re-writing $IC_{S1 \rightarrow H1}$:

$$\theta_1 - c - t_{S1,1} \geq \theta_1 - c - (p + (1-p)\lambda) \left(\theta_1 - \frac{c}{p} \right) - (1-p)(1-\lambda)\theta_1$$

$$t_{S1,1} \leq \theta_1 - \frac{p + (1-p)\lambda}{p} c = \theta_1 - c - \frac{(1-p)\lambda}{p} c < \theta_1 - c.$$

Then, we can compare the *RHS* of $IC_{S2 \rightarrow H1}$ and $IC_{S2 \rightarrow S1}$ to show that $IC_{S2 \rightarrow H1}$ is slack since its *RHS* is strictly larger:

$$\begin{aligned} \theta_2 - c - p\theta_2 - (1-p)(\lambda t_{S1,1} + (1-\lambda)t_{S1,\emptyset}) \\ > \theta_2 - c - p\theta_2 - (1-p)(\lambda t_{H1,1} + (1-\lambda)t_{H1,\emptyset}) \end{aligned}$$

$$(\lambda t_{S1,1} + (1-\lambda)t_{S1,\emptyset}) < (\lambda t_{H1,1} + (1-\lambda)t_{H1,\emptyset})$$

$$t_{S1,1} < \theta_1 - c.$$

Hence $\phi_4 = 0$. We have $\phi_3 = (1 - f_1)(1 - q)$, by (5), and from (1), we have

$\phi_1 = f_1 q + \phi_2 \frac{p + (1-p)\lambda}{p}$. And from (2), we have:

$$\begin{aligned}\phi_6 &= f_1 q(1-p) - f_1 q(1-p) - \phi_2 \frac{1-p}{p}(p + (1-p)\lambda) + \phi_2(1-p)(1-\lambda) \\ &= \phi_2 \frac{1-p}{p}(-\lambda) < 0\end{aligned}$$

The Case 3 assumptions lead to a contradiction.

Case 4: $\phi_7 > 0, \phi_5 > 0, \phi_8 = \phi_6 = 0$ (contradiction)

By (3), $\phi_2 > 0$

By (4), $\phi_3 > 0$

$$t_{S1,\emptyset} = \theta_1$$

$t_{H1,1} = t_{H1,\emptyset} = \theta_1 - c$, which also means the *RHS* of $IC_{S1 \rightarrow H1}$ is zero.

Re-writing $IC_{S1 \rightarrow H1}$:

$$\theta_1 - c - (p + (1-p)\lambda)t_{S1,1} - (1-p)(1-\lambda)\theta_1 = 0$$

$$t_{S1,1} = \theta_1 - \frac{c}{p + (1-p)\lambda}$$

This implies that $IC_{S2 \rightarrow S1}$ is slack since the *RHS* of $IC_{S2 \rightarrow S1}$ is strictly larger than the *RHS* of $IC_{S1 \rightarrow H1}$:

$$\Delta\theta(1-p) - \frac{p}{p + (1-p)\lambda}c > \Delta\theta(1-p) - pc$$

Therefore, $\phi_3 = 0$, which contradicts (4).

Case 5: $\phi_5 > 0, \phi_6 = \phi_7 = \phi_8 = 0$ (contradiction)

$$t_{S1,1} < t_{S1,\emptyset} = \theta_1; t_{H1,1} < t_{H1,\emptyset} < \theta_1$$

By (C3) and (1), $\phi_1 > 0$

By (3), $\phi_2 > 0$

By (4), $\phi_3 > 0$

Re-writing constraints:

$$\begin{aligned} IC_{S1 \rightarrow H1}: & (p + (1-p)\lambda)t_{S1,1} + (1-p)(1-\lambda)\theta_1 \\ & = (p + (1-p)\lambda)t_{H1,1} + (1-p)(1-\lambda)t_{H1,\emptyset} \end{aligned}$$

$$IC_{S2 \rightarrow S1}: \theta_2 - c - t_{S2} = \theta_2(1-p) - c - (1-p)(\lambda t_{S1,1} + (1-\lambda)\theta_1)$$

$$IC_{S2 \rightarrow H1}: \theta_2 - c - t_{S2} \geq \theta_2(1-p) - c - (1-p)(\lambda t_{H1,1} + (1-\lambda)t_{H1,\emptyset})$$

Solving from these constraints:

$$t_{S1,1} = t_{H1,1} + \frac{(1-p)(1-\lambda)}{p + (1-p)\lambda} (t_{H1,\emptyset} - \theta_1)$$

$$\lambda t_{H1,1} + (1-\lambda)t_{H1,\emptyset} \geq \lambda \left(t_{H1,1} + \frac{(1-p)(1-\lambda)}{p + (1-p)\lambda} (t_{H1,\emptyset} - \theta_1) \right) + (1-\lambda)\theta_1$$

$$t_{H1,\emptyset} \left(1 - \lambda - \frac{(1-p)\lambda(1-\lambda)}{p + (1-p)\lambda} \right) \geq \theta_1 \left(1 - \lambda - \frac{(1-p)\lambda(1-\lambda)}{p + (1-p)\lambda} \right)$$

$$t_{H1,\emptyset}(1-\lambda) \left(\frac{p}{p + (1-p)\lambda} \right) \geq \theta_1(1-\lambda) \left(\frac{p}{p + (1-p)\lambda} \right)$$

$$t_{H1,\emptyset} \geq \theta_1$$

As established, however, $t_{H1,\emptyset} < \theta_1$, so we arrive at a contradiction.

Case 6: $\phi_6 > 0, \phi_5 = \phi_7 = \phi_8 = 0$ (contradiction)

$$t_{H1,1} < t_{H1,\emptyset} = \theta_1; t_{S1,1} < t_{S1,\emptyset} < \theta_1$$

By (C3) and (1), $\phi_1 > 0$

By (3), $\phi_2 > 0$

By (4), $\phi_3 > 0$

Re-writing constraints:

$$\begin{aligned} IC_{S1 \rightarrow H1}: & (p + (1-p)\lambda)t_{S1,1} + (1-p)(1-\lambda)t_{S1,\emptyset} \\ & = (p + (1-p)\lambda)t_{H1,1} + (1-p)(1-\lambda)\theta_1 \end{aligned}$$

$$IC_{S2 \rightarrow S1}: \theta_2 - c - t_{S2} = \theta_2(1-p) - c - (1-p)(\lambda t_{S1,1} + (1-\lambda)t_{S1,\emptyset})$$

$$IC_{S2 \rightarrow H1}: \theta_2 - c - t_{S2} \geq \theta_2(1-p) - c - (1-p)(\lambda t_{H1,1} + (1-\lambda)\theta_1)$$

Solving from these constraints:

$$t_{H1,1} = t_{S1,1} + \frac{(1-p)(1-\lambda)}{p + (1-p)\lambda} (t_{S1,\emptyset} - \theta_1)$$

$$\lambda t_{H1,1} + (1-\lambda)\theta_1 \geq \lambda t_{S1,1} + (1-\lambda)t_{S1,\emptyset}$$

$$\lambda \left(t_{S1,1} + \frac{(1-p)(1-\lambda)}{p + (1-p)\lambda} (t_{S1,\emptyset} - \theta_1) \right) - (1-\lambda)(t_{S1,\emptyset} - \theta_1) \geq \lambda t_{S1,1}$$

$$\frac{p(1-\lambda)}{p + (1-p)\lambda} (t_{S1,\emptyset} - \theta_1) \geq 0$$

$$t_{S1,\emptyset} \geq \theta_1$$

As established, however, $t_{S1,\emptyset} < \theta_1$, so we arrive at a contradiction.

Case 7: $\phi_7 > 0, \phi_5 = \phi_6 = \phi_8 = 0$ (contradiction)

$$t_{H1,1} = t_{H1,\emptyset}; t_{S1,1} < t_{S1,\emptyset} < \theta_1$$

By (C3), $\phi_1 > 0$

By (3), $\phi_2 > 0$

By (4), $\phi_3 > 0$

Combining IR_{H1} with $IC_{S1 \rightarrow H1}$:

$$(p + (1 - p)\lambda)t_{S1,1} + (1 - p)(1 - \lambda)t_{S1,\emptyset} = t_{H1,1} = t_{H1,\emptyset} = \theta_1 - c$$

Re-stating $IC_{S2 \rightarrow S1}$ and $IC_{S2 \rightarrow H1}$:

$$\theta_2 - c - t_{S2} = \theta_2(1 - p) - c - (1 - p)(\lambda t_{S1,1} + (1 - \lambda)t_{S1,\emptyset})$$

$$\theta_2 - c - t_{S2} \geq \theta_2(1 - p) - c - (1 - p)t_{H1,1}$$

Combining the above:

$$t_{H1,1} \geq \lambda t_{S1,1} + (1 - \lambda)t_{S1,\emptyset}$$

$$(p + (1 - p)\lambda)t_{S1,1} + (1 - p)(1 - \lambda)t_{S1,\emptyset} \geq \lambda t_{S1,1} + (1 - \lambda)t_{S1,\emptyset}$$

$$p(1 - \lambda)(t_{S1,1} - t_{S1,\emptyset}) \geq 0$$

As established, however, $t_{S1,1} < t_{S1,\emptyset}$, so we arrive at a contradiction.

Case 8: $\phi_8 > 0, \phi_5 = \phi_6 = \phi_7 = 0$ (contradiction)

$$t_{S1,1} = t_{S1,\emptyset}; t_{H1,1} < t_{H1,\emptyset} < \theta_1$$

From (C3), $\phi_1 > 0$

From (4), $\phi_2 > 0$

From (5), at least one of ϕ_3 and ϕ_4 is positive.

Re-writing $IC_{S1 \rightarrow H1}$:

$$\begin{aligned} t_{S1,1} &= (p + (1-p)\lambda)t_{H1,1} + (1-p)(1-\lambda)t_{H1,\emptyset} \\ &= \lambda t_{H1,1} + (1-\lambda)t_{H1,\emptyset} + p(1-\lambda)(t_{H1,1} - t_{H1,\emptyset}) \end{aligned}$$

Re-stating $IC_{S2 \rightarrow H1}$ and $IC_{S2 \rightarrow S1}$:

$$\theta_2 - c - t_{S2} \geq \theta_2(1-p) - c - (1-p)(\lambda t_{H1,1} + (1-\lambda)t_{H1,\emptyset})$$

$$\theta_2 - c - t_{S2} \geq \theta_2(1-p) - c - (1-p)t_{S1,1}$$

Re-writing $IC_{S2 \rightarrow S1}$:

$$\begin{aligned} \theta_2 - c - t_{S2} &\geq \theta_2(1-p) - c \\ &\quad - (1-p) \left(\lambda t_{H1,1} + (1-\lambda)t_{H1,\emptyset} + p(1-\lambda)(t_{H1,1} - t_{H1,\emptyset}) \right) \end{aligned}$$

It is then clear that $IC_{S2 \rightarrow S1}$ is more slack than $IC_{S2 \rightarrow H1}$, which must mean that $\phi_3 = 0$ and $\phi_4 = (1-f_1)(1-q)$ from (5).

Combining the above result with (3) and (4):

$$\phi_2 = f_1(1 - q) + (1 - f_1)(1 - q)(1 - p) = (1 - q)(f_1 + (1 - f_1)(1 - p))$$

Re-writing (4):

$$\begin{aligned}\phi_8 &= (1 - p)(1 - \lambda)(\phi_2 - \phi_3) \\ &= (1 - p)(1 - \lambda)(1 - q)(f_1 + (1 - f_1)(1 - p) - (1 - f_1)) \\ &= (1 - p)(1 - \lambda)(1 - q)(f_1 - (1 - f_1)p)\end{aligned}$$

Which is negative by (A3), and so we arrive at a contradiction.

Case 9: $\phi_5 = \phi_6 = \phi_7 = \phi_8 = 0$ (knife edge)

$$t_{i,1} < t_{i,\emptyset} < \theta_1$$

From (4), $\phi_2 = \phi_3$. Re-writing (1),

$$\phi_1 = f_1q + \phi_3 \frac{p + (1 - p)\lambda}{p} + ((1 - f_1)(1 - q) - \phi_3) \frac{(1 - p)\lambda}{p}$$

Re-writing (2),

$$\begin{aligned}\phi_1 &= f_1q + \phi_3(1 - \lambda) + ((1 - f_1)(1 - q) - \phi_3)(1 - \lambda) \\ &= f_1q + (1 - f_1)(1 - q)(1 - \lambda)\end{aligned}$$

Combining the above,

$$\phi_3 + \frac{(1 - f_1)(1 - q)(1 - p)\lambda}{p} = (1 - f_1)(1 - q)(1 - \lambda)$$

$$\phi_3 = \phi_2 = (1 - f_1)(1 - q) \frac{p - \lambda}{p}$$

Given (5),

$$\phi_4 = (1 - f_1)(1 - q) \frac{\lambda}{p}$$

Re-writing (3):

$$f_1(1 - q) - \phi_3(p + (1 - p)\lambda) + \phi_3(1 - p)\lambda = 0$$

$$f_1(1 - q) = \phi_3 p$$

$$(1 - f_1)(1 - q)(p - \lambda) = f_1(1 - q)$$

$$p - \lambda = \frac{f_1}{1 - f_1}$$

In other words, the principal must be exactly indifferent between the LCCP and BA contracts for there to be an interior solution.

Note that the optimal solution satisfies all ignored constraints.

Other notes:

If all types are included, the principal will pool S1 and H1. Hence, if CIC constraints are violated for one low type, they will be violated for both.

Shut down is a Choice: If shut down, only H1

(C5) Complete low type shutdown is inferior to the LCCP contract

If all low types were shut down, high types would gain no rent, resulting in a profit of $(1 - f_1)(\theta_2 - c)$. The second best contract for $q = 0$ pools all agents together, with a resulting profit of $\theta_1 - c$. By (A1), we know that $\theta_1 - c > (1 - f_1)(\theta_2 - c)$, or equivalently

$$\frac{f_1}{1 - f_1} > \frac{\Delta\theta}{\theta_1 - c}$$

Therefore, the second best contract is preferable to shutting down all agents.

Comparing the LCCP profit to the second best profit for $q = 0$,

$$\begin{aligned}\Pi^{LCCP} &= f_1(\theta_1 - c) + (1 - f_1)(\theta_2 - c - q(\Delta\theta(1 - p) - pc)) \\ &= \Pi^{SB} + (1 - f_1)(q\Delta\theta + (1 - q)(p(\Delta\theta + c)))\end{aligned}$$

Therefore, the LCCP contract is always superior to shutting down both low types.

(C6) When H1 is not shut down, shutting down S1 is impossible

Consider the following constraints:

$$IR_{H1}: \theta_1 - c - pt_{H1,1} - (1 - p)t_{H1,\emptyset} \geq 0$$

$$IR_{S1}: \theta_1 - c - pt'_{S1,1} - (1 - p)t'_{S1,\emptyset} \geq 0$$

$$IC_{S1 \rightarrow H1}: \theta_1 - c - pt'_{S1,1} - (1 - p)t'_{S1,\emptyset} \geq \theta_1 - c - pt_{H1,1} - (1 - p)t_{H1,\emptyset}$$

Note that $t'_{i,j} \leq t_{i,j} \forall i, j$; then we can re-write the inequalities:

$$\begin{aligned} \theta_1 - c - pt'_{S1,1} - (1-p)t'_{S1,\emptyset} &\geq \theta_1 - c - pt'_{H1,1} - (1-p)t'_{H1,\emptyset} \\ &\geq \theta_1 - c - pt_{H1,1} - (1-p)t_{H1,\emptyset} \geq 0 \end{aligned}$$

Therefore, if IR_{H1} is upheld, violating IR_{S1} violates $IC_{S1 \rightarrow H1}$ necessarily, and S1 will simply mimic H1.

(C7) Shutting down the high types is suboptimal

H2's transfer is $\theta_2 - c$, i.e. first best transfers. We take $IC_{S2 \rightarrow H2}$ to be non-binding and see later that it is not violated; this means no one profitably attempts to emulate H2, and so there is no reason for shut down.

S2's transfer is $\theta_2 - c - u_{S2}$. In the second best case, u_{S2} is at its highest at $\Delta\theta$, and it is still worthwhile to include S2. Auditing only ever decreases rent, and makes it better to include the type receiving the rent. Therefore, it is better for the principal to include S2.

Hence, the principal has two main options: Allow corruption or not, and shut down H1 or not.

Appendix 1D: Increasing q hurts the principal for small q

Shut Down contract

In this case, bribery is allowed, and H1 is shut down.

$$\max_{t,w} f_1(1-q)(t_{S1,1}) + (1-f_1)(q(\theta_2 - c) + (1-q)t_{S2}) \text{ s. t.}$$

$$IR_{S1}: \theta_1 - c - (p + (1-p)\lambda)t_{S1,1} - (1-p)(1-\lambda)t_{S1,\emptyset} \geq 0$$

$$IC_{S2}: \theta_2 - c - t_{S2} \geq \theta_2(1-p) - c - (1-p)(\lambda t_{S1,1} + (1-\lambda)t_{S1,\emptyset})$$

$$LLC_{S1,\emptyset}: \theta_1 \geq t_{S1,\emptyset}$$

$$NonNeg_{S1}: t_{S1,\emptyset} - t_{S1,1} \geq 0$$

The Lagrangian is:

$$L = f_1(1-q)(t_{S1,1}) + (1-f_1)(q(\theta_2 - c) + (1-q)t_{S2})$$

$$+ \phi_1(\theta_1 - c - (p + (1-p)\lambda)t_{S1,1} - (1-p)(1-\lambda)t_{S1,\emptyset})$$

$$+ \phi_2(\theta_2 p - t_{S2} + (1-p)(\lambda t_{S1,1} + (1-\lambda)t_{S1,\emptyset}))$$

$$+ \phi_3(\theta_1 - t_{S1,\emptyset})$$

$$+ \phi_4(t_{S1,\emptyset} - t_{S1,1})$$

First order conditions:

$$(1) \quad [t_{S1,1}]: f_1(1-q) - \phi_1(p + (1-p)\lambda) + \phi_2(1-p)\lambda - \phi_4 = 0$$

$$(2) \quad [t_{S1,\emptyset}]: -\phi_1(1-p)(1-\lambda) + \phi_2(1-p)(1-\lambda) - \phi_3 + \phi_4 = 0$$

$$(3) \quad [t_{S2}]: (1-f_1)(1-q) - \phi_2 = 0$$

Solving:

Re-stating (3):

$$(4) \quad \phi_2 = (1-f_1)(1-q)$$

Combining (1) and (4):

$$(5) \quad \phi_1 = \frac{(1-q)(f_1+(1-f_1)(1-p)\lambda) - \phi_4}{p+(1-p)\lambda}$$

Combining (2), (4), and (5):

$$\begin{aligned} \phi_3 &= (\phi_2 - \phi_1)(1-p)(1-\lambda) + \phi_4 \\ &= ((1-f_1)(1-q)(p+(1-p)\lambda) - f_1(1-q) \\ &\quad - (1-f_1)(1-p)\lambda(1-q)) \frac{(1-p)(1-\lambda)}{p+(1-p)\lambda} + \phi_4 \left(1 + \frac{(1-p)(1-\lambda)}{p+(1-p)\lambda} \right) \end{aligned}$$

$$(6) \quad \phi_3 = ((1-f_1)p - f_1) \frac{(1-p)(1-\lambda)(1-q)}{p+(1-p)\lambda} + \frac{\phi_4}{p+(1-p)\lambda}$$

Due to binding IR_{S1} , $\phi_3 > 0$ necessitates $\phi_4 = 0$, and $\phi_4 > 0$ necessitates $\phi_3 = 0$. Therefore, by (6), for ϕ_3 to be positive, it must be that $(1-f_1)p - f_1 > 0$, which is implied by (A3).²² (Minor note: This does not imply a border condition for SD and LCCP, because we are assuming H1 is shut down.)

²² If (A3) is instead violated, it must be that $\phi_3 = 0$ and $\phi_4 > 0$, deterring all bribery.

The solution for with $\phi_3 > 0$ is:

$$t_{S1,\emptyset} = \theta_1$$

$$t_{S1,1} = \theta_1 - \frac{c}{p + (1-p)\lambda}$$

$$\theta_2 p + (1-p)(\lambda t_{S1,1} + (1-\lambda)t_{S1,\emptyset}) = t_{S2}$$

$$t_{S2} = \theta_2 p + (1-p) \left(\theta_1 - c \left(\frac{\lambda}{p + (1-p)\lambda} \right) \right) = \theta_2 p + \theta_1 (1-p) - \frac{(1-p)\lambda}{p + (1-p)\lambda} c$$

$$t_{S2} = \theta_2 - c - \left(\Delta\theta(1-p) - \frac{p}{p + (1-p)\lambda} c \right)$$

$$\begin{aligned} \Pi^{SD} &= f_1(1-q) \left(\theta_1 - c - \frac{(1-p)(1-\lambda)}{p + (1-p)\lambda} c \right) \\ &\quad + (1-f_1) \left(\theta_2 - c - (1-q) \left(\Delta\theta(1-p) - \frac{p}{p + (1-p)\lambda} c \right) \right) \end{aligned}$$

Note that the optimal solution satisfies all ignored constraints.

SD of HI optimal for small q

We show this by comparing profits. The profits of the three potentially optimal contracts are as follows:

Least Cost Corruption Proof (LCCP):

$$\Pi^{LCCP} = f_1(\theta_1 - c) + (1-f_1)(\theta_2 - c - (1-q)(\Delta\theta(1-p) - pc))$$

Bribery Allowed (BA):

$$\begin{aligned}\Pi^{BA} = & f_1 \left(\theta_1 - c - (1 - q) \left(\frac{1 - p}{p} c \right) \right) \\ & + (1 - f_1) \left(\theta_2 - c - (1 - q) \left(\Delta\theta(1 - p) - \frac{p - (1 - p)\lambda}{p} c \right) \right)\end{aligned}$$

Bribery Allowed, Shut Down H1 (SD):

$$\begin{aligned}\Pi^{SD} = & f_1(1 - q) \left(\theta_1 - c - \frac{(1 - p)(1 - \lambda)}{p + (1 - p)\lambda} c \right) \\ & + (1 - f_1) \left(\theta_2 - c - (1 - q) \left(\Delta\theta(1 - p) - \frac{p}{p + (1 - p)\lambda} c \right) \right)\end{aligned}$$

We compare the profits of the BA and LCCP contracts and confirm that the critical point where the profits are the same corresponds with our previous analysis.

$$\Pi^{BA} - \Pi^{LCCP} = (1 - f_1)(1 - q) \left(\frac{p - (1 - p)\lambda}{p} - p \right) c - f_1(1 - q) \left(\frac{1 - p}{p} c \right)$$

$$\Pi^{BA} - \Pi^{LCCP} = \frac{1 - q}{p} c [(1 - f_1)(p(1 - p) - \lambda(1 - p)) - f_1(1 - p)]$$

$$\Pi^{BA} - \Pi^{LCCP} = \frac{1 - p}{p} (1 - q) c [(1 - f_1)(p - \lambda) - f_1] = 0$$

$$p = \lambda + \frac{f_1}{1 - f_1} = p^*$$

When $p > p^*$, $\Pi^{BA} > \Pi^{LCCP}$. When $p < p^*$, $\Pi^{LCCP} > \Pi^{BA}$.

This inequality can also determine which profit condition is tighter when comparing profits to SD.

For the purposes of proving that honesty hurts the principal, we will first show that the SD contract is optimal for a range of q , given (A3) holds.

(i) Comparing SD and LCCP contracts:

$$\begin{aligned}\Pi^{SD} - \Pi^{LCCP} &= (1 - f_1)(1 - q) \left(\frac{p}{p + (1 - p)\lambda} - p \right) c \\ &\quad - f_1 \left(q(\theta_1 - c) + (1 - q) \frac{(1 - p)(1 - \lambda)}{p + (1 - p)\lambda} c \right)\end{aligned}$$

$$\begin{aligned}\Pi^{SD} - \Pi^{LCCP} &= (1 - f_1)(1 - q)p(1 - p)(1 - \lambda)c \\ &\quad - f_1(q(p + (1 - p)\lambda)(\theta_1 - c) + (1 - q)(1 - p)(1 - \lambda)c)\end{aligned}$$

$$\Pi^{SD} - \Pi^{LCCP} = ((1 - f_1)p - f_1)(1 - q)(1 - p)(1 - \lambda)c - f_1q(p + (1 - p)\lambda)(\theta_1 - c)$$

For the SD contract to be superior to the LCCP contract, $\Pi^{SD} - \Pi^{LCCP}$ must be positive:

$$\begin{aligned}q &\left(((1 - f_1)p - f_1)(1 - p)(1 - \lambda)c + f_1(p + (1 - p)\lambda)(\theta_1 - c) \right) \\ &< ((1 - f_1)p - f_1)(1 - p)(1 - \lambda)c \\ q &< \frac{((1 - f_1)p - f_1)(1 - p)(1 - \lambda)c}{((1 - f_1)p - f_1)(1 - p)(1 - \lambda)c + f_1(p + (1 - p)\lambda)(\theta_1 - c)} = q_1\end{aligned}$$

By (A3), we know that $(1 - f_1)p > f_1$, which guarantees both the numerator and the denominator are positive, and so $0 < q_1 < 1$.

Note that q_1 is increasing in p . That is to say, while the LCCP and SD contracts both benefit from an increase in p , the SD contract benefits more. The cost of shutting down H1 depends only on $f_1q(\theta_1 - c)$, but the general cost of allowing bribery is highly sensitive to p . The lower p gets, the less often bribery occurs, and the more efficient the SD contract becomes in that respect.

$$\frac{((1-f_1)p-f_1)(1-p)(1-\lambda)c}{((1-f_1)p-f_1)(1-p)(1-\lambda)c+f_1(p+(1-p)\lambda)(\theta_1-c)} = q_1$$

$$q_1 = \frac{f(p)}{f(p)+g(p)}$$

$$f(p) = ((1-f_1)p-f_1)(1-p)(1-\lambda)c$$

$$g(p) = f_1(p+(1-p)\lambda)(\theta_1-c)$$

$$\begin{aligned} f'(p) &= (1-f_1)(1-p)(1-\lambda)c - ((1-f_1)p-f_1)(1-\lambda)c \\ &= (f_1+(1-f_1)(1-2p))(1-\lambda)c \end{aligned}$$

$$g'(p) = f_1(1-\lambda)(\theta_1-c)$$

$$\frac{dq_1}{dp} = \frac{f'(p)g(p) - g'(p)f(p)}{(f(p)+g(p))^2}$$

$$\frac{dq_1}{dp}$$

$$= \frac{(f_1+(1-f_1)(1-2p))(1-\lambda)c f_1(p+(1-p)\lambda)(\theta_1-c) - f_1(1-\lambda)(\theta_1-c)((1-f_1)p-f_1)(1-p)(1-\lambda)c}{\left(\left((1-f_1)p-f_1\right)(1-p)(1-\lambda)c + f_1(p+(1-p)\lambda)(\theta_1-c)\right)^2}$$

$$\frac{dq_1}{dp} = \frac{f_1c(1-\lambda)(\theta_1-c)[(p+(1-p)\lambda) - ((1-f_1)p-f_1)]}{\left(\left((1-f_1)p-f_1\right)(1-p)(1-\lambda)c + f_1(p+(1-p)\lambda)(\theta_1-c)\right)^2}$$

$$\frac{dq_1}{dp} = \frac{f_1c(1-\lambda)(\theta_1-c)[pf_1+f_1+(1-p)\lambda]}{\left(\left((1-f_1)p-f_1\right)(1-p)(1-\lambda)c + f_1(p+(1-p)\lambda)(\theta_1-c)\right)^2} > 0$$

(ii) Comparing SD and BA contracts:

$$\begin{aligned}
\Pi^{SD} - \Pi^{BA} &= f_1 \left((1-q) \frac{(1-p)\lambda}{p(p+(1-p)\lambda)} c - q(\theta_1 - c) \right) \\
&\quad + (1-f_1)(1-q) \left(\frac{p}{p+(1-p)\lambda} - \frac{p-(1-p)\lambda}{p} \right) c \\
&= f_1 \left((1-q) \frac{(1-p)\lambda}{p(p+(1-p)\lambda)} c - q(\theta_1 - c) \right) \\
&\quad + (1-f_1)(1-q) \left(\frac{p^2 - p(p+(1-p)\lambda) + (1-p)\lambda(p+(1-p)\lambda)}{p(p+(1-p)\lambda)} \right) c \\
&= f_1 \left((1-q) \frac{(1-p)\lambda}{p(p+(1-p)\lambda)} c - q(\theta_1 - c) \right) + (1-f_1)(1-q) \left(\frac{(1-p)^2 \lambda^2}{p(p+(1-p)\lambda)} \right) c \\
&= (1-q) \frac{(1-p)\lambda}{p(p+(1-p)\lambda)} (f_1 + (1-f_1)(1-p)\lambda) c - q f_1 (\theta_1 - c)
\end{aligned}$$

For the SD contract to be superior to the BA contract, $\Pi^{SD} - \Pi^{BA}$ must be positive:

$$\begin{aligned}
& q \left(\frac{(1-p)\lambda}{p(p+(1-p)\lambda)} (f_1 + (1-f_1)(1-p)\lambda) c + f_1 (\theta_1 - c) \right) \\
& < \frac{(1-p)\lambda}{p(p+(1-p)\lambda)} (f_1 + (1-f_1)(1-p)\lambda) c \\
q & < \frac{(1-p)\lambda (f_1 + (1-f_1)(1-p)\lambda) c}{(1-p)\lambda (f_1 + (1-f_1)(1-p)\lambda) c + p(p+(1-p)\lambda) f_1 (\theta_1 - c)} = q_2
\end{aligned}$$

Clearly, $0 < q_2 < 1$.

Note that q_2 is decreasing in p . This is because the BA contract is even more sensitive to changes in p than the SD contract, due to the inclusion of H1. Rent to both strategic types are based on the fact that, $(1-p)$ of the time, H1 gets punished and requires compensation for that punishment.

$$\frac{(1-p)\lambda(f_1 + (1-f_1)(1-p)\lambda)c}{(1-p)\lambda(f_1 + (1-f_1)(1-p)\lambda)c + p(p + (1-p)\lambda)f_1(\theta_1 - c)} = q_2$$

$$q_2 = \frac{f(p)}{f(p) + g(p)}$$

$$f(p) = (1-p)\lambda(f_1 + (1-f_1)(1-p)\lambda)c$$

$$g(p) = p(p + (1-p)\lambda)f_1(\theta_1 - c)$$

$$f'(p) = -\lambda(f_1 + (1-f_1)(1-p)\lambda)c - (1-p)\lambda^2(1-f_1)c$$

$$g'(p) = (p + (1-p)\lambda)f_1(\theta_1 - c) + p(1-\lambda)f_1(\theta_1 - c)$$

$$\frac{dq_2}{dp}$$

$$= \frac{(-\lambda(f_1 + (1-f_1)(1-p)\lambda)c - (1-p)\lambda^2(1-f_1)c)p(p + (1-p)\lambda)f_1(\theta_1 - c)}{(f(p) + g(p))^2}$$

$$- \frac{((p + (1-p)\lambda)f_1(\theta_1 - c) + p(1-\lambda)f_1(\theta_1 - c))((1-p)\lambda(f_1 + (1-f_1)(1-p)\lambda)c)}{(f(p) + g(p))^2}$$

Both terms are negative, and so the derivative is negative.

From these two contract comparisons, we now have the following lemma:

Lemma 1

Define $\underline{q} = \min\{q_1, q_2\}$. We know that $0 < \underline{q} < 1$, because both q_1 and q_2 are between 0 and 1. When $0 \leq q < \underline{q}$, the SD contract is the optimal contract.

Proof of Proposition: Increasing q hurts the principal under SD

Proposition 1: The principal's profit decreases in q when $0 < q < \underline{q}$.

The proof takes two steps:

1. Per Lemma 1, the optimal contract is the SD contract when $0 < q < \underline{q}$, where \underline{q} is defined in the previous section and exists between 0 and 1.
2. Given the principal offers the SD contract, assumptions (A1) and (A3) are sufficient to show that the derivative of the principal's profit with respect to q is negative.

Step 1 is evident.

Step 2:

First, we re-write (A1) and (A3):

$$(A1): f_1(\theta_1 - c) > (1 - f_1)\Delta\theta$$

$$(A3): p(1 - f_1) > f_1$$

Now, we take a derivative of the principal's SD contract profit with respect to q :

$$\begin{aligned} \Pi^{SD} = & f_1(1 - q) \left(\theta_1 - c - \frac{(1 - p)(1 - \lambda)}{p + (1 - p)\lambda} c \right) \\ & + (1 - f_1) \left(\theta_2 - c - (1 - q) \left(\Delta\theta(1 - p) - \frac{p}{p + (1 - p)\lambda} c \right) \right) \end{aligned}$$

$$\frac{d\Pi^{SD}}{dq} = -f_1 \left(\theta_1 - c - \frac{(1 - p)(1 - \lambda)}{p + (1 - p)\lambda} c \right) + (1 - f_1) \left(\Delta\theta(1 - p) - \frac{p}{p + (1 - p)\lambda} c \right)$$

$$\frac{d\Pi^{SD}}{dq} = (1 - f_1)\Delta\theta(1 - p) - f_1(\theta_1 - c) + \frac{f_1(1 - p)(1 - \lambda) - (1 - f_1)p}{p + (1 - p)\lambda} c$$

Using (A1) and (A3), we substitute $-(1 - f_1)\Delta\theta$ for $-f_1(\theta_1 - c)$ and $-f_1$ for $-(1 - f_1)p$ to *increase* the right hand side:

$$\begin{aligned} \frac{d\Pi^{SD}}{dq} &< (1 - f_1)\Delta\theta(1 - p) - (1 - f_1)\Delta\theta + \frac{f_1(1 - p)(1 - \lambda) - f_1}{p + (1 - p)\lambda}c \\ &= -(1 - f_1)\Delta\theta - f_1c < 0 \end{aligned}$$

Therefore,

$$\frac{d\Pi^{SD}}{dq} < 0$$

Appendix 2A: Benchmark contracts

Auditing with commitment, without a private signal (P_0)

Converting to a Lagrangian:

$$\begin{aligned}
 L = & f_1(\pi - kc) + (1 - f_1)t_2 \\
 & + \phi_1(\theta_1 - t_1) \\
 & + \phi_2(\theta_2 - t_2) \\
 & + \phi_3(t_1 - t_2 + kF) \\
 & + \phi_4(k)
 \end{aligned}$$

First order conditions:

$$\begin{aligned}
 (1) \quad & \frac{dL}{dt_1} = \pi - \phi_1 + \phi_3(1 - k) = 0 \\
 (2) \quad & \frac{dL}{dt_2} = (1 - \pi) - \phi_2 - \phi_3 = 0 \\
 (3) \quad & \frac{dL}{dk} = -\pi c + \phi_3(\theta_2 - t_1 + p) + \phi_4 = 0
 \end{aligned}$$

Along with the conditions for the constraints.

From (1), IR_1 must bind, which means $t_1 = \theta_1$.

From (2),

$$(4) \quad \phi_2 + \phi_3 = 1 - \pi$$

From (3) and binding IR_1 ,

$$(5) \quad \phi_4 = \pi c - \phi_3 F, \text{ which means that if } \phi_3 = 0, \phi_4 > 0, \text{ and vice versa.}$$

From (4) and (5),

$$(6) \quad \phi_2 = 1 - \pi - \frac{\pi c - \phi_4}{F} = \frac{(1 - \pi)F - \pi c + \phi_4}{F}$$

Meaning $\phi_4 > 0$ if $\pi c > (1 - \pi)(\Delta\theta + p)$, or $\frac{\pi}{1-\pi} > \frac{F}{c}$.

Case 1: $\phi_3 > 0, \phi_4 = 0, \phi_2 > 0$ (solution)

$$\phi_3 = \frac{\pi c}{F}$$

$$\phi_2 = 1 - \pi - \frac{\pi c}{F}$$

$$\theta_2 = t_2$$

$$0 = \Delta\theta - kF$$

$$k = \frac{\Delta\theta}{F}$$

Case 2: $\phi_3 > 0, \phi_4 = 0, \phi_2 = 0$ (knife-edge)

$$\phi_3 = 1 - \pi = \frac{\pi c}{F}$$

This is the knife-edge case on the border of auditing optimality.

Case 3: $\phi_3 = 0, \phi_4 > 0$

$$k = 0$$

$$\phi_1 = \pi$$

$$t_1 = \theta_1$$

$$\phi_4 = \pi c$$

$$\phi_2 = 1 - \pi$$

$$t_2 = \theta_2$$

This violates IC_2 .

Case 3: $\phi_3 > 0, \phi_4 > 0, \phi_2 = 0$

$$k = 0$$

$$t_1 = t_2 = \theta_1$$

$$\phi_3 = 1 - \pi$$

$$\phi_4 = \pi c - (1 - \pi)F$$

As $\frac{c}{F} < \frac{1-\pi}{\pi}$, there is a contradiction. If that assumption were violated, the principal would revert to the second best contract.

Appendix 2B: Auditing with commitment, noncompliance

$$\begin{aligned}
L = & (\pi + (1 - \pi)x)t_1 + k_1 \left((1 - \pi)(1 - q)xF - (\pi q + (1 - \pi)x(1 - q))c \right) \\
& + k_2 \left((1 - \pi)qx F - ((1 - \pi)qx + \pi(1 - q))c \right) \\
& + (1 - \pi)(1 - x)(t_1 + (qk_2 + (1 - q)k_1)F) \\
& \quad + \phi_1(\theta_1 - t_1) \\
& \quad + \phi_2(\theta_2 - t_1 - (qk_2 + (1 - q)k_1)F) \\
& \quad + \phi_3(k_1) \\
& + \phi_4 \left((1 - \pi)qx F - (1 - \pi)qxc - \pi(1 - q)c \right)
\end{aligned}$$

First order conditions:

$$(1) \quad \frac{dL}{dx} = (1 - \pi)t_1 + k_1 \left((1 - \pi)(1 - q)F - (1 - \pi)(1 - q)c \right) + k_2 \left((1 - \pi)qF - (1 - \pi)qc \right) - (1 - \pi)(t_1 + (qk_2 + (1 - q)k_1)F) + \phi_4 \left((1 - \pi)qF - (1 - \pi)qc \right) = 0$$

$$(2) \quad \frac{dL}{dt_1} = \pi + (1 - \pi)x + (1 - \pi)(1 - x) - \phi_1 - \phi_2 = 0$$

$$(3) \quad \frac{dL}{dk_1} = (1 - \pi)(1 - q)x F - (\pi q + (1 - \pi)x(1 - q))c + (1 - \pi)(1 - x)(1 - q)F - \phi_2(1 - q)F + \phi_3 = 0$$

$$(4) \quad \frac{dL}{dk_2} = (1 - \pi)qx F - ((1 - \pi)qx + \pi(1 - q))c + (1 - \pi)(1 - x)qF - \phi_2qF = 0$$

Solving:

From (1),

$$\begin{aligned}
& (1 - \pi)t_1 + k_1((1 - \pi)(1 - q)F - (1 - \pi)(1 - q)c) + k_2((1 - \pi)qF - (1 - \pi)qc) \\
& \quad - (1 - \pi)(t_1 + (qk_2 + (1 - q)k_1)F) + \phi_4((1 - \pi)qF - (1 - \pi)qc) \\
& \quad = 0
\end{aligned}$$

$$\phi_4((1 - \pi)qF - (1 - \pi)qc) = (1 - \pi)c(k_1(1 - q) + k_2q)$$

$$(5) \quad \phi_4 = \frac{(1 - \pi)c(k_1(1 - q) + k_2q)}{(1 - \pi)qF - (1 - \pi)qc} > 0$$

Therefore,

$$(1 - \pi)qx F - (1 - \pi)qxc - \pi(1 - q)c = 0$$

$$x(1 - \pi)q(F - c) = \pi(1 - q)c$$

$$(6) \quad x = \frac{\pi}{1 - \pi} \frac{1 - q}{q} \frac{c}{F - c}$$

From (4),

$$(1 - \pi)qx F - ((1 - \pi)qx + \pi(1 - q))c + (1 - \pi)(1 - x)qF - \phi_2 qF = 0$$

$$\phi_2 qF = (1 - \pi)qF - ((1 - \pi)qx + \pi(1 - q))c$$

Plugging in (6),

$$(7) \quad \phi_2 = (1 - \pi)(1 - x) > 0$$

Therefore,

$$(8) \quad t_1 = \theta_2 - (qk_2 + (1 - q)k_1)F$$

Plugging (7) into (2),

$$\pi + (1 - \pi)x + (1 - \pi)(1 - x) - (1 - \pi)(1 - x) = \phi_1$$

Plugging in (6),

$$\pi \left(1 + \frac{1 - q}{q} \frac{c}{F - c} \right) = \phi_1$$

$$(9) \quad \phi_1 = \frac{\pi}{q(F-c)} (q(F-c) + (1-q)c) > 0$$

Therefore,

$$(10) \quad t_1 = \theta_1$$

Plugging (7) into (3),

$$\begin{aligned} (1-\pi)(1-q)xF - (\pi q + (1-\pi)x(1-q))c + (1-\pi)(1-x)(1-q)F \\ - (1-\pi)(1-x)(1-q)F + \phi_3 &= 0 \\ \phi_3 &= (\pi q + (1-\pi)x(1-q))c - (1-\pi)(1-q)xF \\ \phi_3 &= \pi qc - (1-\pi)(1-q)x(F-c) \end{aligned}$$

Plugging in (6),

$$\begin{aligned} \frac{\pi}{1-\pi} \frac{1-q}{q} \frac{c}{F-c} \\ \phi_3 = \pi qc - \pi \frac{(1-q)^2}{q} c \end{aligned}$$

$$(11) \quad \phi_3 = \frac{\pi c}{q} (q^2 - (1-q)^2) > 0$$

Therefore,

$$k_1 = 0$$

Full solution:

$$t_i = \theta_i$$

$$k_1 = 0$$

$$k_2 = \frac{\Delta\theta}{Fq}$$

$$x = \frac{\pi}{1-\pi} \frac{1-q}{q} \frac{c}{F-c}$$

Appendix 2C: Extensions

Second Best

Lagrangian:

$$\begin{aligned}
 L = & \pi(\theta_1 e_1 - t_1) + (1 - \pi)(\theta_2 e_2 - t_2) \\
 & + \phi_1 \left(t_1 - \frac{e_1^2}{2} \right) \\
 & + \phi_2 \left(t_2 - \frac{e_2^2}{2} - t_1 + \gamma \frac{e_1^2}{2} \right)
 \end{aligned}$$

First order conditions:

$$[t_1]: -\pi + \phi_1 - \phi_2 = 0$$

$$[t_2]: -(1 - \pi) + \phi_2 = 0$$

$$[e_1]: \pi\theta_1 - \phi_1 e_1 + \phi_2 \gamma e_1 = 0$$

$$[e_2]: (1 - \pi)\theta_2 - \phi_2 e_2 = 0$$

Solving:

$$\phi_2 = 1 - \pi$$

$$\phi_1 = 1$$

$$e_2 = \theta_2 \frac{1 - \pi}{\phi_2} = \theta_2$$

$$e_1 = \theta_1 \frac{\pi}{\phi_1 - \phi_2 \gamma} = \theta_1 \frac{\pi}{\pi + (1 - \pi)(1 - \gamma)}$$

Commitment to auditing, no private signal

Lagrangian:

$$L = \pi(\theta_1 e_1 - t_1 - kc) + (1 - \pi)(\theta_2 e_2 - t_2)$$

$$\begin{aligned}
& +\phi_1 \left(t_1 - \frac{e_1^2}{2} \right) \\
& +\phi_2 \left(t_2 - \frac{e_2^2}{2} \right) \\
& +\phi_3 \left(t_2 - \frac{e_2^2}{2} - t_1 + \gamma \frac{e_1^2}{2} + kt_1 \right)
\end{aligned}$$

First order conditions:

$$\begin{aligned}
[k]: & -\pi c + \phi_3 t_1 = 0 \\
[t_1]: & -\pi + \phi_1 - \phi_3(1 - k) = 0 \\
[t_2]: & -(1 - \pi) + \phi_2 + \phi_3 = 0 \\
[e_1]: & \pi \theta_1 - \phi_1 e_1 + \phi_3 \gamma e_1 = 0 \\
[e_2]: & (1 - \pi) \theta_2 - (\phi_2 + \phi_3) e_2 = 0
\end{aligned}$$

Solving:

$$\begin{aligned}
\phi_2 + \phi_3 &= 1 - \pi \\
\phi_3 t_1 &= \pi c \\
\phi_3(1 - k) &= \phi_1 - \pi \\
e_1 &= \theta_1 \frac{\pi}{\phi_1 - \phi_3 \gamma} \\
e_2 &= \theta_2 \\
t_1 &= \frac{e_1^2}{2} = \frac{\theta_1^2 \pi^2}{2(\phi_1 - \phi_3 \gamma)^2} \\
t_2 &= \frac{\theta_2^2}{2}
\end{aligned}$$

$$k = \frac{\left(t_1 - \gamma \frac{e_1^2}{2}\right) - \left(t_2 - \frac{e_2^2}{2}\right)}{t_1} = 1 - \gamma$$

$$\frac{\theta_1^2 \pi^2}{2(\phi_1 - \phi_3 \gamma)^2} = \frac{\pi c}{\phi_3}$$

$$\phi_1 = \phi_3 \gamma + \pi$$

$$e_1 = \theta_1$$

Auditing with a private signal without commitment

$$IC_2: t_2 - \frac{e_2^2}{2} = t_1(1 - qk_2) - \gamma \frac{e_1^2}{2} \quad [E4]$$

$$PIC: c = \frac{xq(1 - \pi)}{xq(1 - \pi) + \pi(1 - q)} t_1 \quad [E5]$$

Re-phrasing:

$$c(xq(1 - \pi) + \pi(1 - q)) = xq(1 - \pi)t_1$$

$$x(t_1 - c)q(1 - \pi) = c\pi(1 - q)$$

$$x = \frac{\pi}{1 - \pi} \frac{c}{t_1 - c} \frac{1 - q}{q} \quad [E6]$$

Notice that, since $1 - q < q$, the probability of shirking is smaller than auditing without commitment with no private signal, all else equal.

I label the following problem P_1 :

$$L = \pi \left(\frac{qt_1 - (2q - 1)c}{q(t_1 - c)} \right) (\theta_1 e_1 - t_1 - [0]) + \left(1 - \pi \left(\frac{qt_1 - (2q - 1)c}{q(t_1 - c)} \right) \right) (\theta_2 e_2 - t_2) \\ + \phi_1 \left(t_1 - \frac{e_1^2}{2} \right)$$

$$+\phi_2 \left(t_2 - \frac{e_2^2}{2} \right)$$

First order conditions:

$$[t_1]: -\pi \frac{c}{(t_1 - c)^2} \frac{1 - q}{q} (\theta_1 e_1 - t_1) - \pi \frac{qt_1 - (2q - 1)c}{q(t_1 - c)} \\ + \pi \frac{c}{(t_1 - c)^2} \frac{1 - q}{q} (\theta_2 e_2 - t_2) + \phi_1 = 0$$

$$[t_2]: - \left(1 - \pi \left(\frac{qt_1 - (2q - 1)c}{q(t_1 - c)} \right) \right) + \phi_2 = 0$$

$$[e_1]: \pi \left(\frac{qt_1 - (2q - 1)c}{q(t_1 - c)} \right) \theta_1 - \phi_1 e_1 = 0$$

$$[e_2]: \left(1 - \pi \left(\frac{qt_1 - (2q - 1)c}{q(t_1 - c)} \right) \right) \theta_2 - \phi_2 e_2 = 0$$

Solving:

$$e_2 = \theta_2 [E7]$$

$$t_2 = \frac{\theta_2^2}{2} [E8]$$

$$e_1 = \theta_1 \frac{\pi}{\phi_1} \left(\frac{qt_1 - (2q - 1)c}{q(t_1 - c)} \right)$$

$$\phi_1 = \frac{\pi}{q(t_1 - c)^2} \left(qt_1(t_1 - c) - (2q - 1)c(t_1 - c) - c(1 - q) \left(\frac{\theta_2^2}{2} - (\theta_1 e_1 - t_1) \right) \right)$$

$$\phi_1 = \frac{\pi}{e_1 q(t_1 - c)} (qt_1 - (2q - 1)c) \theta_1$$

$$\frac{\pi}{q(t_1 - c)^2} \left(qt_1(t_1 - c) - (2q - 1)c(t_1 - c) - c(1 - q) \left(\frac{\theta_2^2}{2} - (\theta_1 e_1 - t_1) \right) \right)$$

$$= \frac{\pi}{e_1 q (t_1 - c)} (qt_1 - (2q - 1)c) \theta_1$$

$$qt_1(t_1 - c) - (2q - 1)c(t_1 - c) - c(1 - q) \left(\frac{\theta_2^2}{2} - (\theta_1 e_1 - t_1) \right)$$

$$= \frac{t_1 - c}{e_1} (qt_1 - (2q - 1)c) \theta_1$$

$$e_1 = \theta_1 \frac{(t_1 - c)(qt_1 - (2q - 1)c)}{(t_1 - c)(qt_1 - (2q - 1)c) - c(1 - q) \left(\frac{\theta_2^2}{2} - (\theta_1 e_1 - t_1) \right)} \quad [E9]$$