View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by ScholarlyCommons@Penn

University of Pennsylvania
ScholarlyCommons

Publicly Accessible Penn Dissertations

2017

Essays On Dynamic Incentive Design

Mustafa Dogan
University of Pennsylvania, mudogan@sas.upenn.edu

Follow this and additional works at: https://repository.upenn.edu/edissertations

Part of the Economics Commons

Recommended Citation
Dogan, Mustafa, "Essays On Dynamic Incentive Design" (2017). *Publicly Accessible Penn Dissertations*. 2259.
https://repository.upenn.edu/edissertations/2259

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/edissertations/2259
For more information, please contact repository@pobox.upenn.edu.

Essays On Dynamic Incentive Design

Abstract
This dissertation consists of three essays that examine incentive problems within various dynamic environments. In Chapter 1, I study the optimal design of a dynamic regulatory system that encourages regulated agents to monitor their activities and voluntarily report their violations. Self-monitoring is a private and costly process, and comprises the core of the incentive problem. There are no monetary transfers. Instead, the regulator (she) uses future regulatory behavior for incentive provision. When the regulator has full commitment power, she can induce costly self-monitoring and revelation of ``bad news'' in the initial phase of the optimal policy. During this phase, the agent is promised a higher continuation utility (in the form of future regulatory approval) each time he discloses ``bad news''. If the regulator internalizes self-monitoring costs, the agent is either blacklisted or whitelisted in the long run. When she does not internalize these costs, blacklisting is replaced by a temporary probation state, and whitelisting becomes the unique long run outcome. This result suggests that whitelisting, which may appear to be a form of regulatory capture, may instead be a consequence of optimal policy.

In Chapter 2, I study the dynamic pricing problem of a durable good monopolist with commitment power, when a new version of the good is expected at some point in the future. The new version of the good is superior to the existing one, bringing a higher flow utility. The buyers are heterogeneous in terms of their valuations and strategically time their purchases. When the arrival is a stationary stochastic process, the corresponding optimal price path is shown to be constant for both versions of the good, hence there is no delay on purchases and time is not used to discriminate over buyers, which is in line with the literature. However, if the arrival of the new version occurs at a commonly known deterministic date, then the price path may decrease over time, resulting in delayed purchases. For both arrival processes, posted prices is a sub-optimal selling mechanism. The optimal one involves bundling of both versions of the good and selling them only together, which can easily be implemented by selling the initial version of the good with a replacement guarantee.

Finally, Chapter 3 examines the question under what conditions can automation be less desirable compared to human labor. We study a firm that has to decide between a human-human team and a human-machine team for production. The effort choice of a human employee is not observed by the manager, therefore the incentives need to be properly aligned. We argue that, despite the desirable benefits resulting from the partial substitution of labor with automated machines such as less costly machine input and reduced scope of moral hazard, the teams with only human employees can, under some conditions, be more preferred over the human-machine teams. This stems from the fact that, in all-human teams, the principal, through the selection of incentive scheme, can control the interaction among the agents and get benefit from the mutual monitoring capacity between them. The automation, however, eliminates this interaction and shuts down a channel that can potentially help to mitigate the overall agency problem.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/2259

Graduate Group
Economics

First Advisor
George J. Mailath

Subject Categories
Economics

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/2259

ESSAYS ON DYNAMIC INCENTIVE DESIGN

Mustafa Dogan

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2017

Supervisor of Dissertation

George J. Mailath, Professor of Economics

Graduate Group Chairperson

Jesus Fernandez-Villaverde, Professor of Economics

Dissertation Committee

George J. Mailath, Professor of Economics

Steven Matthews, Professor of Economics

Mallesh M. Pai, Assistant Professor of Economics

ESSAYS ON DYNAMIC INCENTIVE DESIGN

© COPYRIGHT

2017

Mustafa Dogan

To my parents

ACKNOWLEDGEMENTS

I am extremely thankful to my advisor, Professor George Mailath, for his guidance and encouragement throughout my doctoral studies. I have benefited deeply from his intellect along the years that I spent at Penn. He is, and will always be a reference of academic perfection and diligence to me.

I also thank the other members of my dissertation committee, Steven Mathews, and Mallesh Pai, for their support. I am thankful to my co-author Pinar Yildirim. My research has also benefited from discussions with Aislinn Bohren, David Dillenberger, Ju Hu, Nicholas Janetos, Joonbae Lee, SangMok Lee, Andrew Postlewaite, Francisco Silva, Rakesh Vohra, Yuichi Yamamoto.

Kelly Quinn deserves a special thank for her constant willingness to help.

I am indebted to Ismail Benli, Ihsan Dogan, Selman Erol, Salih Ozaslan, Metin Uyanik, and especially Ramazan Bora for sincere friendship.

Finally, I am grateful to my mother, father and sisters for their true love and limitless faith on me. I wouldn't be me without my family. Their existence has always been an essential constant of my life and motivated me towards the pursuit of wisdom.

ABSTRACT

ESSAYS ON DYNAMIC INCENTIVE DESIGN

Mustafa Dogan

George J. Mailath

This dissertation consists of three essays that examine incentive problems within various dynamic environments. In Chapter 1, I study the optimal design of a dynamic regulatory system that encourages regulated agents to monitor their activities and voluntarily report their violations. Self-monitoring is a private and costly process, and comprises the core of the incentive problem. There are no monetary transfers. Instead, the regulator (she) uses future regulatory behavior for incentive provision. When the regulator has full commitment power, she can induce costly self-monitoring and revelation of "bad news" in the initial phase of the optimal policy. During this phase, the agent is promised a higher continuation utility (in the form of future regulatory approval) each time he discloses "bad news". If the regulator internalizes self-monitoring costs, the agent is either blacklisted or whitelisted in the long run. When she does not internalize these costs, blacklisting is replaced by a temporary probation state, and whitelisting becomes the unique long run outcome. This result suggests that whitelisting, which may appear to be a form of regulatory capture, may instead be a consequence of optimal policy.

In Chapter 2, I study the dynamic pricing problem of a durable good monopolist with commitment power, when a new version of the good is expected at some point in the future. The new version of the good is superior to the existing one, bringing a higher flow utility. The buyers are heterogeneous in terms of their valuations and strategically time their purchases. When the arrival is a stationary stochastic process, the corresponding optimal price path is shown to be constant for both versions of the good, hence there is no delay on purchases and time is not used to discriminate over buyers, which is in line with the literature. However, if the arrival of the new version occurs at a commonly

known deterministic date, then the price path may decrease over time, resulting in delayed purchases. For both arrival processes, posted prices is a sub-optimal selling mechanism. The optimal one involves bundling of both versions of the good and selling them only together, which can easily be implemented by selling the initial version of the good with a replacement guarantee.

Finally, Chapter 3, which is a joint work with Pinar Yildirim, examines the question under what conditions can automation be less desirable compared to human labor. We study a firm that has to decide between a human-human team and a human-machine team for production. The effort choice of a human employee is not observed by the manager, therefore the incentives need to be properly aligned. We argue that, despite the desirable benefits resulting from the partial substitution of labor with automated machines such as less costly machine input and reduced scope of moral hazard, the teams with only human employees can, under some conditions, be more preferred over the human-machine teams. This stems from the fact that, in all-human teams, the principal, through the selection of incentive scheme, can control the interaction among the agents and get benefit from the mutual monitoring capacity between them. The automation, however, eliminates this interaction and shuts down a channel that can potentially help to mitigate the overall agency problem.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv

ABSTRACT v

LIST OF TABLES xi

LIST OF ILLUSTRATIONS xii

CHAPTER 1 : Dynamic Incentives for Self-Monitoring 1

1.1 Introduction 1

1.1.1 Literature Review 6

1.2 Model 9

1.2.1 Actions and Preferences 13

1.2.2 An Alternative Interpretation 14

1.2.3 Policy 15

1.2.4 Stationary Representation 17

1.2.5 Principal's Problem $\mathcal{P}$ 18

1.3 Case 1: Principal Internalizes Cost 19

1.3.1 Observable Information Acquisition 19

1.3.2 Moral Hazard and The Optimal Policy 23

1.3.3 Conditional Representation 24

1.3.4 The problem $\mathcal{P}_D$ 27

1.3.5 The problem $\mathcal{P}_N$ 30

1.3.6 The problem $\mathcal{P}'$ 32

1.4 Case 2: Principal Does Not Internalize Cost 40

1.4.1 Observable Information Acquisition 41

1.4.2 Moral Hazard and the Optimal Policy 44

1.5 Limited Commitment 50
1.6 Concluding Remarks 53

CHAPTER 2 : Product Upgrades and Posted Prices 54
2.1 Introduction 54
2.1.1 Literature Review 57
2.2 Model 58
2.2.1 Stochastic Arrival 60
2.2.2 Deterministic Arrival 61
2.3 Optimal Posted Prices: Stochastic Arrival 62
2.3.1 Benchmark I: Canonical Durable Good Monopoly 62
2.3.2 Benchmark II: Product Upgrade without Anonymous Prices 65
2.3.3 Sales With Posted Prices: Anonymity 67
2.4 Deterministic Arrival 74
2.5 Conclusion 79

CHAPTER 3 : Man vs. Machine: When is automation inferior to human labor? . 81
3.1 Introduction 81
3.2 Literature Review 84
3.3 Model 86
3.3.1 No-Automation Regime 86
3.3.2 Automation Regime: 88
3.4 Optimal Incentive Contract 88
3.4.1 Optimal Incentive Scheme Under Automation 89
3.4.2 Optimal Incentive Scheme Under No-Automation 90
3.4.3 Static Environment 92
3.4.4 Dynamic Environment 93
3.4.5 Collusion and Renegotiation-proofness 102
3.4.6 The Optimal Regime 104

3.4.7 Human teams survive under costless automation 106
3.5 Extensions 108
3.5.1 Heterogeneous Effort Decision 109
3.5.2 First Best Benchmark 110
3.5.3 Human-Machine Team 111
3.5.4 Human-Human Team 112
3.6 Conclusions 116

APPENDIX 120

CHAPTER A : Appendix to Chapter 1 120
A.1 Proof of Lemma 1.1 120
A.2 Proof of Lemma 1.2 121
A.3 Proof of Lemma 1.3 122
A.4 Proof of Lemma 1.4 124
A.5 Proof of Proposition 1.1 125
A.6 Proof of Lemma 1.5 127
A.7 Proof of Proposition 1.2 128
A.8 Proof of Lemma 1.6 130

CHAPTER B : Appendix to Chapter 2 131
B.1 Proof of Lemma 2.1 131
B.2 Proof of Lemma 2.4 132
B.3 Proof of Proposition 2.1 136
B.4 Proof of Lemma 2.5 138
B.5 Proof of Lemma 2.6 139
B.6 Proof of Proposition 2.2 140

CHAPTER C : Appendix to Chapter 3 147
C.1 Proof of Proposition 3.2 147

C.2 Proof of Lemma 3.2 148
C.3 Proof of Lemma 3.3 149
C.4 Proof of Proposition 3.4 156
C.5 Proof of Lemma 3.1 156
C.6 Proof of Proposition 3.5 161
C.7 Proof of Lemma 3.4 162
C.8 Proof of Proposition 3.6 164

BIBLIOGRAPHY 164

LIST OF TABLES

TABLE 1.1 : Alternative Interpretation 15

TABLE 3.1 : Joint Distribution of Signals 87

TABLE B.1 : Sale Policies and Corresponding Profits. 146

LIST OF ILLUSTRATIONS

FIGURE 1.1 : The Stage Game 14
FIGURE 1.2 : Value Function V^* 22
FIGURE 1.3 : Representing Conditional Problem 25
FIGURE 1.4 : Optimal Delegation Decisions in Case 1 35
FIGURE 1.5 : Value Functions V_N, and V_D 38
FIGURE 1.6 : Value Function $\tilde{V}^*$ 44
FIGURE 1.7 : Optimal Delegation Decisions in Case 2 46

FIGURE 3.1 : Optimal Regime and Incentive Scheme 106
FIGURE 3.2 : Optimal Regime, Costless Automation 107

FIGURE C.1 : Restriction on The Incentive Scheme 150

Chapter 1

Dynamic Incentives for Self-Monitoring

1.1 Introduction

The U.S. Environmental Protection Agency has an auditing policy that encourages companies to monitor their ongoing and planned activities that may fall within its authority, and to voluntarily report their violations.[1] This is an example of the framework that I explore in this paper. I am interested in understanding the behavior of regulators in environments where the regulated activity may result in bad outcomes and where there is significant uncertainty. Agents have an advantage in acquiring information about these activities because they have lower costs of monitoring. The regulator, in an efficient regulatory regime, would like to use agents' self-monitoring. I study how regulators can induce economic agents to acquire and disclose costly information about the negative consequences of their activities through the use of future regulatory behavior without resorting to monetary transfers.

I show that, when the regulator has full commitment power, the optimal policy can induce self-monitoring only in an initial phase, which endures over a stochastic number of periods. When it ends, a terminal phase of the policy is initiated and self-monitoring stops.

[1] The auditing policy is defined in Environmental Protection Agency (11 April 2000), titled "Incentives for Self-Policing: Discovery, Disclosure, Correction and Prevention of Violations".

The outcome in this terminal phase is history-dependent and involves either blacklisting the agent or whitelisting him. When the regulator does not internalize self-monitoring costs, blacklisting is replaced by a temporary probation state. The unique long-run outcome is whitelisting in this case. This result suggests that whitelisting, which may appear to be a form of regulatory capture, may instead be a consequence of optimal policy. I also analyze the case in which the regulator's commitment power is limited so that she cannot commit to policies with negative continuation values. In this case, if the expected cost of the social harm is larger than the economic benefits of the projects, then whitelisting never occurs in an optimal policy. Moreover, self-monitoring is sustained over the long term when the regulator does not internalize its costs.

In general, many enforcement authorities adopt self-monitoring practices for various regulatory purposes.[2] A specific practice is the process of issuing licenses for activities with possible environmental consequences. A mining company, for example, in applying for a mining license, may be asked to submit an Environmental Impact Statement and sometimes other supplementary information that requires substantial and costly self-monitoring. The grant of the license empowers the company to operate and contributes to the aggregate economy. Yet, it may also cause some undesirable social consequences that the regulator needs to take into account.[3] To make matters worse, these undesired outcomes oftentimes take a considerable amount of time to become apparent so that it is no more feasible to take ameliorating action.[4] Therefore, investigating these potential harms prior to making a licensing decision is the only convenient policy for the regulator. And these investigations are delegated to the applicant company through the request of Environment Impact Statement.

Incentive divergence is the most prominent feature of the aforementioned settings. The agents prefer to avoid suspension of their activities and also generally prefer to avoid mon-

[2]Securities and Exchange Commission, U.S. Department of Agriculture, U.S. Department of Defense, and Food and Drug Administration are examples.

[3]For example, the mining area may have invisible connections to groundwater resources, in which case mining activities might lead to the production and the spread of hazardous material.

[4]There are many cases, in the mining industry, for example, where the actual damage become apparent only decades after the operations took place.

itoring due to its costs and the possibility of unfavorable signals that it might reveal. On the other hand, the regulators care about efficiency for which monitoring and suspending harm-producing activities are essential. Therefore, the regulator has a complicated policy problem that involves supervising the agents and at the same time incentivizing them to self-monitor.

To study the regulator's problem, I construct a dynamic principal-agent model in which a stream of projects arrives over time, one for each period. The agent (he) wants to undertake the projects, but needs the approval of the regulator (she), who is the principal. The projects are ex-ante identical, yielding the same revenue for the agent. On the other hand, a project might either be harmless or it might result in social costs which would outweigh the value that the project generates. The agent can acquire information about whether the project has social costs by costly self-monitoring, but the efforts spent on self-monitoring are not directly observed by the principal. The principal has the objective of inducing socially preferable outcomes. Her preferences involve the economic benefits as well as the social costs resulting from these projects. Each period, the principal first decides whether or not to ask for self-monitoring, and then chooses whether to approve the project. There is no ex-post monitoring, and the realized harms are never observed. In many settings, the harms occur, or become evident, with a significant lag compared to the economic yields of the projects. To abstract from this reality, I analyze the case in which this lag goes to infinity.

There are no monetary transfers. In many situations, regulatory agencies are limited in their ability to use monetary transfers for various reasons. For example, in some industries there is a legal limitation on the size of the monetary fines that regulators can levy.[5] As a result, monetary transfers are too weak to induce proper behavior, and the regulator needs to use other tools for incentive provision. This paper focuses on the extreme case

[5]In the oil and gas industry of the U.S., there is a daily limit on the maximum amount of fines, and this limit varies across states in the U.S. The total amount of fine the regulating authority collects is is negligible compared to the economic benefit that the companies receive, see for instance E&E Publishing, LLC (November 14, 2011).

of the regulators' limited use of monetary transfers by ruling them out. In the absence of monetary transfers, the regulator provides incentives by linking her decisions over time.

The information structure governing the self-monitoring process takes the form of verifiable "bad news" which are publicly observed. More precisely, there is a unique verifiable signal perfectly revealing bad news and informing about the harm that will occur if the project is undertaken. In case the agent performs self-monitoring in a particular period, conditional on the project being harmful, the signal will be realized with some probability and will be publicly observed. If there is no news, then there are two possibilities from the principal's perspective. First, the agent shirked and did not monitor. Second, the agent acquired information; however, no signal was realized since the project is more likely not to cause harm. There is no direct signal indicating good news. In most of the settings that fit into this paper, the only good news is the absence of bad news. In other words, certifiably disclosing good news is not possible. On the contrary, bad news, in general, provides concrete evidence and detailed description of the harm that will occur if the project is undertaken. Conditional on this information structure, assuming that the signal is publicly observed is without loss of generality. As long as the agent prefers to monitor himself, he also prefers to disclose the signal in case it is realized. Otherwise, he could simply shirk in the first place and eliminate the cost of monitoring. The incentives that induce information acquisition automatically induces the disclosure contingent upon acquisition. Therefore, the signal remains public throughout the discussion in the paper, and hence self-disclosure exogenously occurs conditional on self-monitoring.

Initially, I study the case in which the principal has full commitment power. At the beginning, she commits to history-contingent policy that specifies her decisions regarding self-monitoring requests, and the approval of the projects in each period. I show that self-monitoring is only induced in an initial phase of the optimal policy. During this initial phase, the agent is promised a higher continuation utility (in the form of future regulatory approval) each time he discloses bad news. His current project is less likely to get approved,

but the regulator promises more frequent approval in the future. If he does not disclose any signal, he is downgraded to a lower continuation utility. His current project has higher chances of approval, yet he will be given less frequent approvals in the future. The duration of this phase is stochastic; when it ends, the policy reaches a second phase in which there is no more self-monitoring. The transitional dynamics between the phases and the long-run outcome of the optimal policy depends on whether the principal internalizes the cost of self-monitoring.

If the principal internalizes the costs of self-monitoring, the acquired information is always used in the approval decision. The agent's continuation utility eventually reaches either its minimum or maximum and remains constant. In this stage, the principal permanently rejects or permanently approves projects, that is, the agent is either blacklisted or whitelisted in the long run.[6]

When the principal does not internalize the self-monitoring costs, the content of the information is not always used in the current approval decision, in contrast to the previous case. There is a probation state, which replaces blacklisting, wherein the agent acquires information, but the project is rejected regardless of the outcome. The probation occurs when the agent's continuation utility reaches its minimum in consequence of the agent not disclosing bad news frequent enough. After being initiated, this probationary state repeats until the agent discloses some bad news. Leaving the probation state today does not rule out the possibility of facing it again in the future. The agent's continuation utility eventually reaches its maximum which still puts permanent approval into action, in the long run, the agent is always whitelisted.

The above-mentioned difference in the principal's preferences alters the set of effective incentive devices she is willing to use. When she does not internalize its costs, self-monitoring can purely be used to punish the agent. It is possible for the principal to use

[6]EPA blacklists companies at times, by labeling them as ineligible for federal contracts, subcontracts, grants or loans. See Washington Post (February 5, 1977). The outcome permanent rejection can be considered as the counterpart of blacklisting. .

self-monitoring as punishment, because verifiability ensures that monitoring effort is taken. While the same channel was also feasible when the principal internalized the self-monitoring costs, she preferred not to punish this way because she cared about the cost.

In this model, whitelisting is an outcome of the optimal regulatory policy. Hence, I do not interpret it as a form of regulatory capture even though it shares some of its features. My paper, therefore, suggests that what has been described as regulatory capture in some cases may instead be an outcome of optimal regulatory policy.

Two situations give rise to inefficiencies in this framework. The first one occurs when the agent forgoes information acquisition, and the second one arises when the content of the information is not used efficiently. The first type only appears during the terminal phase of the optimal policy where there is no more self-monitoring. The second type, appears in the initial phase. Its occurrence triggers the terminal phase of the contract, when principal internalizes self-monitoring costs. Therefore, the inefficiencies are back-loaded in this case. However, when the principal does not internalize the self-monitoring costs, the second type of inefficiency occurs in a non-consecutive stochastic order, and its occurrence does not necessarily initiate the terminal phase. In this respect, efficiency will be lost and restored stochastically throughout the optimal policy.

I also study the situation in which the principal has limited commitment power, in that she cannot commit to a policy with a negative continuation value. The results change remarkably. If the expected cost of a project is higher than its economic benefit, the policy does not feature whitelisting. In this case, if the principal does not internalize self-monitoring costs, the policy never reaches a stable outcome and fluctuates over time.

1.1.1 Literature Review

Kaplow and Shavell (1994) is the first study analyzing self-monitoring and self-reporting. They introduce a self-reporting stage into the classical probabilistic law enforcement model

of Becker (1968). By self-reporting a harmful act, the agent is granted a reduction in the sanctions he faces. In contrast to my model, the agent in their paper is initially endowed with the relevant information. Pfaff and Sanchirico (2000) introduce a more general framework in which the agency problem has two tiers: testing for noncompliance and fixing it.[7] There is no information asymmetry to begin with, and the agent needs to exert effort to acquire relevant information. Most of the papers in this literature focus on characterizing the optimal incentive scheme in a static framework. My paper, however, studies the dynamics of a regulatory regime incorporating self-monitoring. In a contemporaneous work, Wang et al. (2016) also study a similar dynamic environment. The main distinction is that the harms are already known to agent and monetary transfers are allowed in their framework. They show that the optimal regime, in order to induce the agent to disclose harms, incorporates a cyclical structure alternating between rewarding self-disclosure and initiating inspections. Departing from theirs, my paper provides some explanation for practices such as blacklisting which arise from dynamic consequences.

In its use of non-monetary intertemporal incentives as a disciplining device, this paper relates to several different branches of literature. In mechanism design, Horner and Guo (2015) (HG) analyze a dynamic allocation problem in the absence of monetary transfers. The principal is interested in efficiency, which requires that the principal allocate the good only if the agent has a high valuation. The optimal mechanism follows a history-dependent rule which eventually converges to permanent allocation or permanent rejection of allocation. In the literature on relational contracts, Li et al. (2015) analyze the evolution of power inside organizations within the context of what they call a repeated trust game. The efficient equilibrium has a structure similar to that of HG, incorporating a bipolar long run outcome with permanent punishment and permanent rewards for the agent. Both of these papers assumes that the agent is initially informed about the state variable. In my paper, however, the state variable is initially unknown to both (the principal and the agent); but, the agent can acquire information about it at some cost. The effort spent on information

[7] Also see Short and Toffel (2008), Innes (1999a), and Innes (1999b).

acquisition is not observed by the principal, and the agency problem is moral hazard instead of adverse selection unlike HG. The fact that the relevant information comes with a cost changes the dynamic structure of the optimal contract/policy. More precisely, when the principal does not internalize the cost of information acquisition, the long-run outcome is unique, and permanent punishment is never a part of the optimal policy in contrast to HG and (Li et al., 2015). Moreover, if the principal has a limited commitment power, the optimal policy does not reach a stable outcome, instead it fluctuates over time.

Lipnowski and Ramos (2015) consider the repeated game version of HG. The efficient equilibria in their framework have a unique long-run outcome featuring a permanent punishment for the agent, for much the same reason permanent does not occur in the limited commitment section of my paper. In contrast to Lipnowski and Ramos (2015), I show that the optimal policy does not necessarily reach a stable outcome in this case, when the principal does not internalize the self-monitoring costs. Battaglini (2005) focuses on the same allocation problem as in HG without ruling out monetary transfers. The principal is a profit-maximizing monopolist. In this paper, the inefficiencies are entirely front-loaded. The use of monetary transfers played an important role on this significant difference, as it alters the natural way of providing incentives.

This paper also relates to the literature on relationship formation and trading favors. Möbius (2001), Hauser and Hopenhayn (2008), and Abdulkadiroğlu and Bagwell (2013) are examples. Within a repeated game setting, players facilitate cooperation by providing favors to each other. The favor provision ability changes over time and is privately observed. Inter-temporal incentives are utilized to elicit proper behavior and sustain cooperative gains. Möbius (2001) suggests a simple chip mechanism which keeps track of the difference in the number of favors provided. There is a maximum number of chips that can be maintained in an equilibrium. Hauser and Hopenhayn (2008) improve on this by considering a more general set of mechanisms. The optimal policy in my paper can also be interpreted as a chips mechanism, in which several different factors affecting the amount of chips that the

agent has for the next period. First, the quantity of the agent's chips expands over time at a constant rate that is equal to the inverse of the discount factor. Second, the amount of chips that the agent has for the next period diminishes at an amount that is proportional to the approval rate in the current period. Finally, the agent receives a fixed amount of additional chips for each piece of bad news he discloses.

My paper also relates to the literature on linked decisions. Jackson and Sonnenschein (2007), within a static environment, showed that linking multiple independent decisions can help overcome incentive constraints. See also Cohn (2010), Hortala-Vallve (2010), and Fang and Norman (2006).

1.2 Model

There is a principal (she) and an agent (he) interacting within a discrete time infinite horizon setting, and δ is the common discount factor. A stream of projects arrives over time, one for each period $t = 1, 2, ..., \infty$. The agent would like to undertake each of these projects for which he needs the approval of the principal.

Approving a project yields a positive value $v \in (0, 1)$ to the agent. In addition to this value, each project may cause a social harm depending on its type θ, which takes its values from the binary space $\Theta = \{\theta_g, \theta_b\}$. If $\theta = \theta_b$, then the project is "bad", producing harm with a magnitude normalized to 1. Otherwise, if $\theta = \theta_g$, then the project is "good", and does not produce any harm. The type of the project is initially unknown to the principal and the agent, and $\mu = P(\theta = \theta_b)$ is the common prior about it. The project types are independently and identically distributed over time; hence, the project arriving at the beginning of each period is believed to be a bad one with probability $\mu \in (0, 1)$.

The principal is interested in efficiency, and wants to maximize the social surplus in her decision. The surplus resulting from the approval of the project is equal to v or $v - 1$

depending on whether the project is good or bad, respectively. The agent, on the other hand, only cares about the value that the project generates for him, hence his utility increases by v each time a project is approved irrespective of its type. Rejecting the project causes a loss due to the forgone value v, yet, at the same time, prevents the production of probable harms. Therefore, from an ex-post point of view, she wants to grant an approval for a project only if it is a good one.

At each period, the agent can acquire information about the type of the project at cost c prior to the principal's approval decision. The information acquisition process, which is also referred as self-monitoring, is governed by the following information structure. There is a unique verifiable signal "s" which perfectly reveals "bad news" about the type of the project. Conditional on the project being bad, the signal is realized with probability $\lambda \leq 1$. If the project is good, then the signal is never realized. More precisely:

$$P(s|\theta_b) = \lambda, \qquad P(s|\theta_g) = 0.$$

I assume that the signal is publicly observed whenever it is realized.[8] Due to this publicity, "self-reporting", which refers to the event of signal realization, exogenously follows conditional on self-monitoring.

The event of no signal realization following the information acquisition, besides being informative, does not perfectly reveal the type of the project (unless $\lambda = 1$). Conditional on information acquisition, the posterior beliefs after signal realization and no signal realization

[8]This publicity assumption is without loss of any generality. All of the results would also follow in a more general framework, where the agent privately observes the signal and then decides whether or not to disclose it to the principal. This stems from the fact that, as long as the agent prefers to monitor himself, he also prefers to disclose the signal in case it is realized. Otherwise, he could simply shirk in the first place and eliminate the cost of monitoring. Of course, this property crucially depends on the signal structure governing the information acquisition.

are denoted by μ_s and μ_{ns} respectively, which satisfy:

$$\mu_s = 1,$$
$$\mu_{ns} = \frac{\mu(1-\lambda)}{1-\mu\lambda}.$$

In the ex-ante stage the expected cost of a project is equal to $\mu = \mu 1 + (1-\mu)0$. Therefore, the ex-ante expected surplus that arises from the project approval is $v - \mu$. There is no assumption imposed on the sign of this value, hence both approval and rejection can be the optimal uninformed decision. The assumptions on the parameters that are maintained throughout the entire paper are defined as follows:

Assumption 1.1. *The parameters of the model satisfies the following:*

i) $v > \mu_{ns}$.

ii) $(1-\mu\lambda)(v-\mu_{ns}) - c > \max(v-\mu, 0)$.

iii) $\delta > \frac{1}{1+\mu\lambda v}$.

The first assumption states that, from the principal's perspective, it is optimal to approve the project in case no signal realization takes place as a result of information acquisition. The second assumption states that the information acquisition is efficient, hence the problem is not a trivial one. The third assumption states that the discount factor is large enough and players are sufficiently patient.

The effort spent on self-monitoring is not observed by the principal, and this generates the moral hazard component which constitutes the main source of the agency problem. If the incentives are not provided properly, then the agent would shirk rather than monitoring the projects. By this, the agent can get rid of the cost of self-monitoring, and at the same time prevent the revelation of bad news and hence the suspension of his projects. Nonetheless, because the information acquisition is efficient, the principal wants to design

an incentive scheme to motivate the agent towards this end. Note that the verifiability of the signal plays a crucial role. It would never be feasible to induce self-monitoring under an information structure that comprises only non-verifiable soft information.

There are no monetary transfers. In many situations, regulatory agencies has limited ability to use monetary transfers, which leads them to use other tools such as future regulatory behavior for incentive provision. In this framework, the principal would be able to induce the first-best outcome under the presence of monetary transfers, by using a stationary payment scheme.[9] Such a stationary scheme, however, would be insufficient to sort out the extent to which the principal utilizes the continuation values arising from future regulatory behavior as an incentive device. In order to analyze these dynamics, one should either employ a more general framework incorporating additional aspects, or restrict the existing one. To eliminate technical difficulties and maintain tractability, I follow the latter and rule out the monetary transfers in the analysis.

There is no initial information asymmetry about the type of the project. This does not rule out the possibility of an agent having superior information about other relevant issues. For instance, as the owner of the project, the agent might be better informed about the direction in which to search for "bad news". This can be considered as the basis for agent's comparative advantage in terms of monitoring capability. This is effectively an information asymmetry, yet it is not directly related to the type of the project.

Ex-ante monitoring is the only source of information on the type of the project. There is no possibility of ex-post monitoring, and hence the realized harm is never observed. This assumption reflects the fact that, in many circumstances, the harms take place with a significant lag compared to the economic yields of the projects.

A later section of this paper analyzes a model where the principal can also monitor the project with a higher cost. In that setting, in each period prior to making the approval

[9] The first-best scheme involves delegating monitoring to the agent and employing the efficient approval decisions in every period.

decision, the principal makes a decision regarding monitoring. She either monitors the project on her own, or delegates monitoring to the agent, or completely avoids monitoring. In the current model, requesting self-monitoring from the agent is not delegation since the principal does not have an option to monitor the projects on her own. Nonetheless, for notational ease, I will use the word delegation to refer to a self-monitoring request in this section as well.

1.2.1 Actions and Preferences

For each period t, after a new project arrives, the principal first decides whether or not to delegate monitoring to the agent. If delegation occurs, the agent moves and decides decides either to shirk or to exert effort and monitor himself. Then, finally, the principal moves again and decides whether or not to approve the project. This stage game is repeated infinitely many times.

The scope of the conflict between the principal and the agent is not limited to the social costs that the projects may generate. The monitoring costs that the agent assumes is another factor contributing to the extent of the conflict. This paper studies two different specifications. Under the first specification, the principal internalizes the cost of self-monitoring; in the second one, she does not internalize it. The conflict becomes more intense under the second specification. The corresponding stage game, together with the payoffs corresponding to the first specification, is illustrated in the following figure. Note that the principal's payoffs are in expectation terms except for those terminal nodes resulting from project rejections, and signal realization. The expectation is based on the belief about the type of the project, which is either μ, if no information is acquired, or μ_{ns} if information is acquired but no signal is realized.

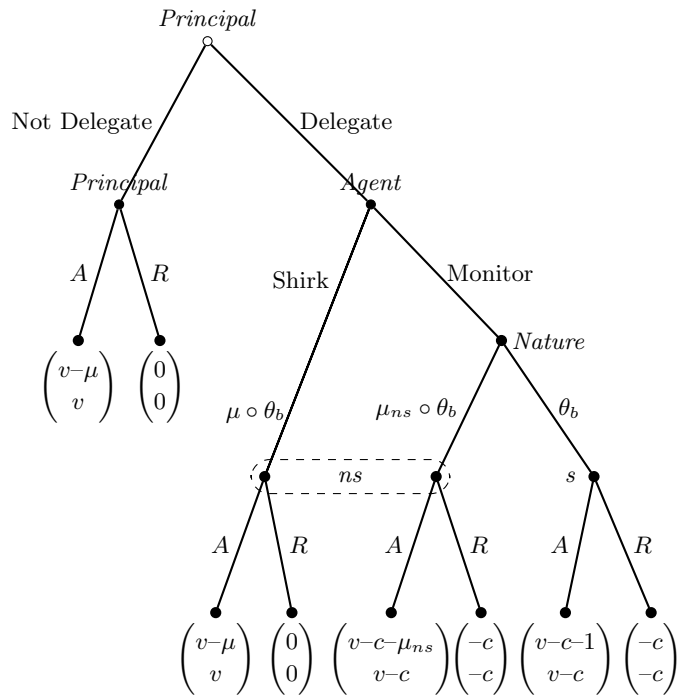

Figure 1.1: The stage game between the principal and the agent. The principal internalizes the cost of self-monitoring. All the terminal nodes, except those are reached by nature's move θ_b,s , and project rejections, are reflecting principal's expected payoffs over the determination of the type of the project. μ is the unconditional probability of project being bad, and μ_{ns} is the probability of project being bad conditional on no signal is realized after information acquisition.

1.2.2 An Alternative Interpretation

There is an alternative interpretation of the model. Suppose that there are some costly precautionary measures that can be initiated by the agent if the principal forces him to do so. From the principal's point of view, these precautions are necessary in case the project is bad since they completely eliminate the harms; otherwise, they are wasteful. The agent wants to avoid these measures regardless of the type of the project, due to the costs. Let z be the cost of these measures; then the corresponding ex-post payoffs are described in the following table:

	Precautions	No Precautions
Good	$v-z, v-z$	v, v
Bad	$v-z, v-z$	$v-1, v$

Table 1.1: Description of the corresponding payoffs at the ex-post stage for the principal and the agent respectively.

When $z = v$, the above description is equivalent to the baseline model, where forcing the precautions corresponds to the rejection of the project, and vice versa. However, the results of the paper also follow for a larger set of z values. One just needs to make sure that z is not very different from v.[10]

1.2.3 Policy

The principal is endowed with the full commitment power, and, at $t = 0$, commits to a dynamic policy that specifies delegation and approval decisions over time as a function of the public history. The public history consists of information about the realized decisions in the earlier periods as well as the self-monitoring outcomes for those periods in which the monitoring is delegated to the agent.

For a given time period t, the corresponding delegation decision and the realized outcome of the self-monitoring, if performed, are denoted together by $r_t \in \{s, ns, n\}$. When monitoring is delegated to agent, r_t will be either "s", if self-disclosure takes place; or will be "ns", if no self-disclosure takes place. If there is no delegation, then $r_t = n$. Moreover, the approval decision at time t is denoted by $d_t \in \{0, 1\}$, where 0 and 1 indicates rejection and approval respectively. Consequently, a within-period public history, at the end of the period, which is denoted by h_t, is of the following form:

$$h_t = (r_t, d_t) \in \{s, ns, n\} \times \{0, 1\},$$

[10]For instance, when z is sufficiently small, then the principal would always force these measures and avoid costly information acquisition.

for each t. At the beginning of a period t, a public history is defined as

$$h^t = (h_1, ..., h_{t-1}).$$

The initial history is $h^1 = h_0 = \emptyset$, and H^t is the set of public histories at period t.

A policy $\Gamma = \{\gamma_t, x_t\}_{t=1}^{\infty}$ is then a sequence of functions which are defined as follows:

$$\gamma_t : H^t \to [0, 1],$$

$$x_t : H^t \times \{s, ns, n\} \to [0, 1].$$

The function γ_t is the probability of delegation. Because the delegation takes place at the beginning of each period, it is a function defined over the set H^t. On the contrary, the approval decision is also conditioned on the value of r_t, hence the relevant domain for x_t is $H^t \times \{s, ns, n\}$. For each possible value of r_t, I use a separate notation for the approval rate, i.e., $x_t = (x_t^s, x_t^{ns}, x_t^n)$. Note that, x_t^s and x_t^{ns} are well-defined as long as $\gamma > 0$; similarly x_t^n is relevant when $\gamma < 1$.

For an incentive compatible policy Γ, the expected utilities of the principal and the agent are denoted by V and U respectively, and are given by:

$$U = \mathbb{E}\Big[(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\big[\underbrace{\gamma_t\{\mu\lambda x_t^s v + (1-\mu\lambda)x_t^{ns}v - c\}}_{\text{Delegation}} + \underbrace{(1-\gamma_t)x_t^n v}_{\text{No Delegation}}\big]\Big]$$

$$V = \mathbb{E}\Big[(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\big[\underbrace{\gamma_t\{\mu\lambda x_t^s(v-1) + (1-\mu\lambda)x_t^{ns}(v-\mu_{ns}) - c\}}_{\text{Delegation}} + \underbrace{(1-\gamma_t)x_t^n(v-\mu)}_{\text{No Delegation}}\big]\Big]$$

The posterior beliefs about the type of the project affect only the principal's utility. The agent does not care about the project's type. Another important point is that the self-monitoring cost is included in the principal's utility. As it was mentioned earlier, the

other case will be analyzed later on.

1.2.4 Stationary Representation

Following Spear and Srivastava (1987), I express the principal's problem within a stationary form, in which the ex-ante expected utility of the agent is the state variable. In this form, the interval $[0, v]$ is the corresponding state space as it consists of all of the possible values that the ex-ante expected utility of the agent can take. The agent's utility cannot be negative, because he can always guarantee a non-negative utility by shirking every time he is asked to monitor. On the other hand, v is the maximum that the agent can receive in a policy. The principal can grant this maximal utility to agent by approving all of the projects without requesting self-monitoring.

State variable is updated over time depending on the realized public history. Within-period decisions and the promised future continuation utilities of the agent, depend on this state variable as well as the realized outcomes of the current period. An optimal policy specifies a different continuation utility for each possible $r_t \in \{s, ns, n\}$ as in the case of the approval probabilities. More precisely, the components of the policy are defined as:

$$\begin{aligned} \gamma, x_s, x_{ns}, x_n &: [0, v] \to [0, 1], \\ U_s, U_{ns}, U_n &: [0, v] \to [0, v]. \end{aligned}$$

The delegation and the approval decisions consist, in essence, of probabilities; therefore, the functions γ, x_s, x_{ns}, x_n take their values from the unit interval. The continuation utilities, on the other hand, specify the state variable for the next period; hence, they are defined as functions from the state space to itself. For a given policy, the functions U_s, U_{ns}, x_s, and x_{ns} are relevant only for those values of $U \in [0, v]$ satisfying $\gamma(U) > 0$, whereas the functions U_n and x_n are relevant only for those satisfying $\gamma(U) < 1$.

The promised utility of the agent is calculated in ex-ante terms; hence, it will be granted to the agent only in expectation. It aggregates the flow and continuation utilities of the agent. Its transition is governed by the policy and the stochastic realizations. Starting from U, the state variable of the next period becomes $U_n(U)$, or $U_s(U)$, or $U_{ns}(U)$ with the corresponding probabilities $1-\gamma(U)$, $\gamma(U)\mu\lambda$ and $\gamma(U)(1-\mu\lambda)$ respectively.

In an optimal policy, the functions $U_n, U_s, U_{ns}, \gamma, x_n, x_s, x_{ns}$ are chosen to maximize the principal's objective function. There are two constraints that the principal needs to take into account in this problem: incentive constraint and promise keeping constraint. More precisely:

1.2.5 Principal's Problem $\mathcal{P}$

$$V(U) = \max_{\substack{\gamma, \\ x_n, x_s, x_{ns}, \\ U_n, U_s, U_{ns}}} \gamma\Big[\mu\lambda\big[(1-\delta)x_s(v-1)+\delta V(U_s)\big] + (1-\mu\lambda)\big[(1-\delta)x_{ns}(v-\mu_{ns})+\delta V(U_{ns})\big] - c\Big]$$
$$+(1-\gamma)\Big[(1-\delta)x_n(v-\mu)+\delta V(U_n)\Big]$$

subject to the (PK) and (IC) respectively:

$$U = \gamma\Big[\mu\lambda\big[(1-\delta)x_s v+\delta U s\big] + (1-\mu\lambda)\big[(1-\delta)x_{ns}v+\delta U_{ns}\big] - c\Big] + (1-\gamma)\Big[(1-\delta)x_n v+\delta U_n\Big],$$
$$\gamma\Big[\mu\lambda\big[(1-\delta)x_s v+\delta U s\big] + (1-\mu\lambda)\big[(1-\delta)x_{ns}v+\delta U_{ns}\big] - c\Big] \geq \gamma\Big[(1-\delta)x_{ns}v+\delta U_{ns}\Big].$$

The first line of the principal's objective function includes her utility contingent upon self-monitoring request; hence, it is multiplied by the probability γ. The second line, on the other hand, corresponds to the contingency of no self-monitoring; hence, it is multiplied by $1-\gamma$.

The first constraint is the promise keeping constraint. It makes sure that the agent's expected utility is equal to U, the utility level promised to him. Likewise the principal's

utility, the agent's utility also has two components depending on the principal's delegation decision.

The second constraint is the incentive constraint, which is defined to make sure that acquiring information is an optimal choice for the agent when he is asked to do so; hence, it is relevant only if $\gamma > 0$. It guarantees that the utility the agent achieves from shirking is no better than the promised utility. If he shirks, there is no self-disclosure; hence, the current approval rate and the continuation utility will be equal to x_{ns} and U_{ns} respectively.

Note that Blackwell's sufficiency conditions, i.e., monotonicity and discounting, are fulfilled. Therefore, the existence of a solution for the problem $\mathcal{P}$ is guaranteed.

1.3 Case 1: Principal Internalizes Cost

In this section, I study the optimal policy and its properties under the first preference specification, i.e., when the principal internalizes the self-monitoring costs. First, I start with a benchmark analysis.

1.3.1 Observable Information Acquisition

In this benchmark, I consider the case in which the agent's self-monitoring costs are publicly observed; hence, there is no agency problem. The natural question is whether the principal can induce the first best outcome or not? The first best outcome involves information acquisition and utilization of the optimal approval decision, i.e., approving the project if no-signal is generated and rejecting it otherwise, in every period. The principal can simply induce this first best outcome by punishing any deviation with permanent rejection of the future projects. Such an incentive scheme is sufficient to induce the proper behavior; because the agent has a positive utility from the first best outcome, whereas permanent rejection leaves him a 0 payoff. The expected utilities of the agent and the principal are denoted by

w and π respectively, and are given by:

$$w = (1 - \mu\lambda)v - c,$$
$$\pi = (1 - \mu\lambda)(v - \mu_{ns}) - c.$$

The optimal policy when the self-monitoring effort is observable induces the first best outcome, which gives the maximum possible utility to the principal. What about the optimal policy conditional on the agent receiving a certain utility $U \in [0, v]$ when the effort is observable? To answer this question, one needs to solve the same problem $\mathcal{P}$ without including the incentive constraint. This problem has a solution, and the value function arising from its solution, which I denote by V^*, is an upper-bound for the value function V.

To describe V^* and the corresponding "benchmark policy", I will point out some initial observations. First of all, the benchmark policy is stationary without loss of generality. The expected utility of the agent stays constant throughout time; and hence, one just needs to characterize the delegation and the approval decisions, which I denote by $\gamma^*, x_n^*, x_s^*, x_{ns}^*$, as functions of $U \in [0, v]$.

Second, the information must be used efficiently. In other words, whenever the agent is asked to monitor himself, the following approval decision must be efficient conditional on the content of the resulting information. This is a direct implication of the fact that the principal internalizes the self-monitoring costs. Rather than having $x_s^* > 0$ or $x_{ns}^* < 1$, the principal could adjust the probability of delegation, γ^*, without hurting the promise keeping constraint, and get strictly better off. To see this, first note that:

$$U = \gamma^*[\mu\lambda x_s^* v + (1 - \mu\lambda)x_{ns}^* v - c] + (1 - \gamma^*)x_n^* v.$$

Suppose, $x_s > 0$ to get a contradiction, then it must be true that $x_{ns} = 1$; because, otherwise, there is an immediate deviation that the principal can perform by decreasing x_s^* and increasing x_{ns}^*. Then, consider decreasing $x_s \downarrow$ by some $\epsilon > 0$; and decreasing delegation

probability to $\gamma - \zeta$. Moreover assume that the principal employs a direct approval with the remaining ζ probability. Then ζ satisfies:

$$\gamma^*[\mu\lambda x_s^* v + (1-\mu\lambda)v - c] = (\gamma^* - \zeta)[\mu\lambda(x_s^* - \epsilon)v + (1-\mu\lambda)v - c] + \zeta v.$$

Therefore:

$$\zeta = \frac{\gamma\mu\lambda v\epsilon}{\mu\lambda v(1 - x_n + \epsilon) + c}.$$

Such a deviation is strictly better for the principal, because:

$$\Big[(\gamma^*-\zeta)[\mu\lambda(x_s^*-\epsilon)(v-1)+(1-\mu\lambda)(v-\mu_{ns})-c]+\zeta(v-\mu)\Big]-\Big[\gamma^*[\mu\lambda x_s^*(v-1)+(1-\mu\lambda)(v-\mu_{ns})-c]\Big] > 0.$$

A similar contradiction follows for the case $x_{ns}^* < 1$.

Finally, upon not requesting self-monitoring, the principal either directly approves or directly rejects the projects and not chooses a randomized approval decision. Instead of choosing an $x_n^* \in (0,1)$, she could increase the probability of delegation, γ^*, and adjust the value of x_n^* by respecting the promise keeping constraint. This increases the probability of informed decision making and improves the principal's objective.

By putting these observations in an order, one can easily specify the details of the benchmark policy. When $U \in [0, w)$, the agent is supposed to get a utility that is less than the utility he would get in the first best outcome. Therefore, the principal randomizes between the first best outcome and the direct rejection without information acquisition. On the other hand, when $U \in (w, v]$, the randomization occurs between the first best outcome and the direct approval without information acquisition. In both cases, the probability of delegation γ is chosen such that the agent gets the exact utility promised to him. More precisely, an optimal policy in this benchmark is given by:

$$\gamma^*(U) = \begin{cases} \frac{U}{w} & \text{if } U \in [0, w] \\ \frac{v-U}{v-w} & \text{if } U \in (w, v] \end{cases}$$

$$x_s^* = 0, \; x_{ns}^* = 1, \; \forall U \in (0, v)$$

$$x_n^*(U) = \begin{cases} 0 & \text{if } U \in [0, w) \\ 1 & \text{if } U \in (w, v] \end{cases}$$

And the resulting value function is:

$$V^*(U) = \begin{cases} \frac{U}{w}\pi & \text{if } U \in [0, w] \\ \frac{v-U}{v-w}\pi + \frac{U-w}{v-w}(v-\mu) & \text{if } U \in (w, v] \end{cases}$$

The following figure illustrates the value function.

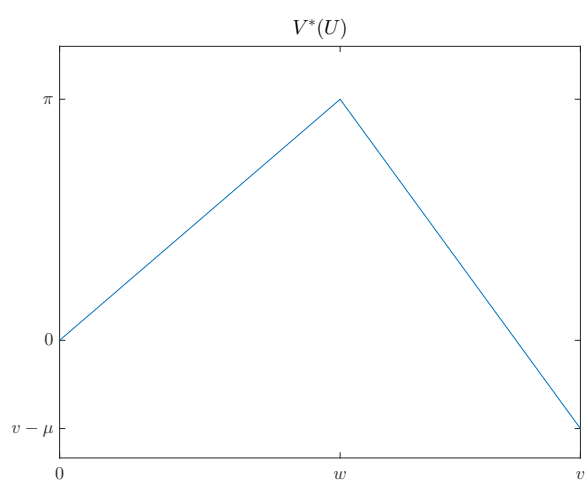

Figure 1.2: The benchmark value function V^*, which consists of an upper bound for V. The parameters used in the illustration assumes that $v < \mu$.

1.3.2 Moral Hazard and The Optimal Policy

The focus is now on the agency problem where the agent's efforts spent on self-monitoring are not observed by the principal. The interest is particularly on the characterization of the value function V, within-period decisions γ, x_n, x_s, x_{ns}, and the promised continuation utilities U_n, U_s, U_{ns}, which are defined as functions of the ex-ante expected promised utility U. The following lemma, which is proved in appendix, is a first step towards this goal.

Lemma 1.1. *The value function, V is concave and hence differentiable almost everywhere. Moreover its derivative is bounded and satisfies:*

$$1 - \frac{\mu\lambda}{\mu\lambda v + c} \leq V'(U) \leq 1 - \frac{\mu(1-\lambda)}{(1-\mu\lambda)v - c}, \ \forall U \in [0, v]. \tag{1.1}$$

Proof. See appendix A.1. □

The concavity of the value function is a direct implication of the fact that the principal can randomize between different utility levels while granting a specific promised utility to agent. Since V is concave, it is almost everywhere differentiable, which can also be proved by applying the result of Benveniste and Scheinkman (1979). The following observations are sufficient to show the bounds of the derivative of V. First, for any promised utility U, $V(U)$ cannot be larger than $V^*(U)$. Second, the values of V and V^* are equal to each other at the boundaries of the state space, i.e., at 0 and v. These observations together with the concavity require that the constant slopes of V^* over the intervals $[0, w]$ and $[w, v]$ are the upper and the lower bounds of V' respectively.

In each period, the principal first makes her delegation decision. In case she randomizes with some $\gamma \in (0, 1)$, each possible outcome of this randomization may result with different utility levels for the principal and the agent. This stems from the fact that the promised utility of the agent is given only in expectation. The principal's commitment power plays a crucial role here. She can fulfill the exact randomization that the policy specifies, and keep

her promises in expectation. Otherwise, the principal, in each period that she is supposed to randomize, could pick the outcome that she prefers most instead of following the specified randomization.

In this regard, one can represent the utility of the agent as a weighted average of two components, one for each possible outcome of the delegation decision. Precisely, let U_D and U_N be the resulting utilities of the agent after delegation and no delegation respectively. More precisely:

$$U_N = (1-\delta)x_n v + \delta U_n$$
$$U_D = \mu\lambda\big[(1-\delta)x_s v + \delta U s\big] + (1-\mu\lambda)\big[(1-\delta)x_{ns} v + \delta U_{ns}\big] - c.$$

The promise keeping constraint imposes a restriction on the choices of U_D, U_N, and γ, so that the equality $U = \gamma U_D + (1-\gamma)U_N$ must hold. When the agent is delegated with certainty, i.e when $\gamma = 1$, it must be $U = U_D$; similarly, $U = U_N$ must hold when $\gamma = 0$.

1.3.3 Conditional Representation

In what follows, I will further exploit the above-mentioned observation, and rewrite the principal's problem $\mathcal{P}$ as a decomposition of two conditional sub-programs. These sub-programs are defined conditional on the current delegation decision, and their task is to characterize the approval decision in the current period as well as the continuation utilities for the next period.

The first problem is defined conditional on principal delegating monitoring to the agent in the current period. The solution to this problem characterizes the optimal values for x_{ns}, x_s, U_{ns}, and U_s depending on the value of U_D. It incorporates the incentive constraint and a promise keeping constraint. The promise keeping constraint here is defined to make sure that the agent's conditional expected utility is equal to U_D.

The second program is defined conditional on agent not delegating monitoring to the agent in the current period. Its solution characterizes the optimal values of x_n and U_n depending on the value of U_N. There is no incentive constraint in this problem, since there is no request of self-monitoring. There is only a promise keeping constraint defined to make sure that x_n and U_n are arranged so that the agent's utility is equal to U_N.

Conditional on no delegation, the agent's utility can take all the values in the entire state space, hence the range of U_N is equal to $[0, v]$. However, this is not the case for U_D. It is defined conditional on an incentive compatible self-monitoring in the current period; hence, the agent already assumes the cost c. This means that U_D cannot be equal to v or anything sufficiently close to v; therefore, the range of U_D can only be a proper subset of the state space. The exact range of U_D will be discussed later on.

The corresponding value functions arising from these conditional programs are denoted by V_D and V_N respectively. The unconditional value function, V, is then given by:

$$V(U) = \gamma(U)V_D(U_D) + (1 - \gamma(U))V_N(U_N),$$

where $U = \gamma(U)U_D + (1 - \gamma(U))U_N$. The following diagram illustrates this conditional representation of the principal's problem.

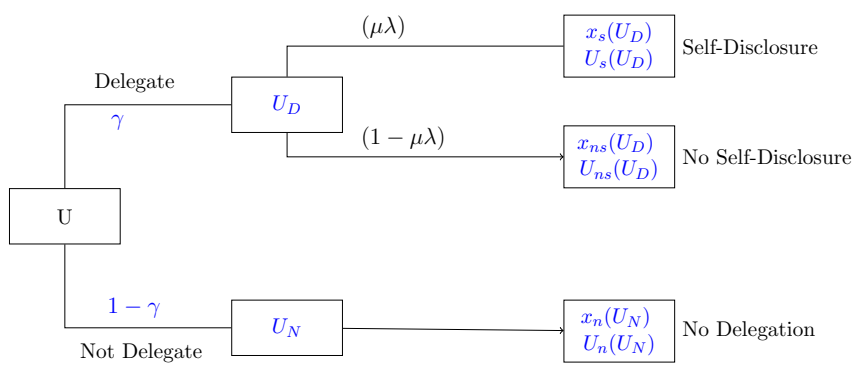

Figure 1.3: The principal's problem is equivalent to solving for the optimal values of U_D, U_N, and γ.

This new form of the principal's problem is denoted by $\mathcal{P}'$ and it is defined as follows:

$$V(U) = \max_{\gamma, U_D, U_N} \gamma V_D(U_D) + (1-\gamma)V_N(U_N),$$
$$\text{s.t.} \quad U = \gamma U_D + (1-\gamma)U_N.$$

The sub-program that is conditional on no delegation in the current period is denoted by $\mathcal{P}_N$, and is defined as follows:

$$V_N(U_N) = \max_{x_n, U_n} \left[(1-\delta)x_n(v-\mu) + \delta V(U_n)\right],$$
$$\text{s.t.} \quad U_N = (1-\delta)x_n v + \delta U_n.$$

Finally, the sub-program that is conditional on delegation in the current period is denoted as $\mathcal{P}_D$, and is defined as:

$$V_D(U_D) = \max_{\substack{x_n, x_s, x_{ns}, \\ U_n, U_s, U_{ns}}} \Big[\mu\lambda\big[(1-\delta)x_s(v-1) + \delta V(U_s)\big] + (1-\mu\lambda)\big[(1-\delta)x_{ns}(v-\mu_{ns}) + \delta V(U_{ns})\big] - c\Big]$$

subject to (PK_D) and (IC_D) respectively:

$$U_D = \mu\lambda\big[(1-\delta)x_s v + \delta U s\big] + (1-\mu\lambda)\big[(1-\delta)x_{ns}v + \delta U_{ns}\big] - c,$$
$$U_D \geq (1-\delta)x_{ns}v + \delta U_{ns}.$$

Thanks to this conditional formulation, it is possible to analyze the principal's delegation and approval decisions separately. First, I will consider the conditional problems in isolation, and solve for the corresponding optimal approval decisions. Then, I will focus on the unconditional problem $\mathcal{P}'$, and characterize the optimal delegation decision over the state space.

There is a crucial observation that I make use of during the above-mentioned process. If, for a given U, a randomized delegation decision, $\gamma \in (0,1)$, is optimal and randomization

takes place between $V_D(U_D)$ and $V_N(U_N)$, then $V(U_D) = V_D(U_D)$ and $V(U_N) = V_N(U_N)$. In other words, $\gamma = 1$ is optimal at U_D, and $\gamma = 0$ is optimal at U_N. This observation is based on the fact that the principal, in order to grant the agent his promised utility U, can always randomize between $V(U_D)$ and $V_N(U_N)$. Therefore, $V(U_D)$ cannot be strictly larger than $V_D(U_D)$, as it would contradict with the optimality.

On account of this, characterizing the state variables that are featuring $\gamma = 0$ or $\gamma = 1$ would be sufficient to pin down the optimal delegation decision. In other words, the focus should be on the subsets of the state space over which either $V(U) = V_D(U)$ or $V(U) = V_N(U)$ is satisfied. For the rest of the state space, there will be a randomized delegation decision, and the corresponding values U_D and U_N will always be a part of the subsets of $[0, v]$ satisfying $V = V_D$ and $V = V_N$ respectively.

The discussion in the sequel will follow the plan descrived above. In order to carry through the first step, I will first focus on the conditional problems in isolation.

1.3.4 The problem $\mathcal{P}_D$

An initial observation is that the incentive constraint is always binding. First of all, note that the IC can be written as a restriction on the difference between the continuation utilities U_s and U_{ns}:

$$U_s - U_{ns} \geq \frac{1-\delta}{\delta\mu\lambda}c + \frac{1-\delta}{\delta}(x_{ns} - x_s)v.$$

In case the difference between U_s and U_{ns} is larger than the value that is necessary to maintain incentive compatibility, the principal can move them closer to each other without violating the promise keeping constraint. More precisely, she can decrease U_s by ϵ and increase U_{ns} by $\frac{\mu\lambda}{1-\mu\lambda}\epsilon$. Such a modification is always feasible as long as $\epsilon > 0$ is chosen sufficiently small, since both of the constraints are respected. Moreover, it is preferred by

the principal due to the concavity of the value function V. This is because the suggested modification consists of a mean preserving contraction of the continuation utilities, hence the expectation of V for the next period becomes larger. Therefore incentive constraint is always binding, without loss of generality. Solving binding incentive constraint together with the promise keeping constraint gives:

$$U_s = \frac{U_D}{\delta} + \frac{1-\delta}{\delta\mu\lambda}c - \frac{1-\delta}{\delta}x_s v, \tag{1.2}$$

$$U_{ns} = \frac{U_D}{\delta} - \frac{1-\delta}{\delta}x_{ns} v. \tag{1.3}$$

These expressions suggest that the agent is compensated for the cost of information acquisition only after the signal is realized. To see this more clearly, one can rewrite them as follows:

$$(1-\delta)x_{ns}v + \delta U_{ns} = U_D,$$

$$(1-\delta)x_s v + \delta U_s = U_D + \tfrac{1-\delta}{\mu\lambda}c.$$

Self-reporting increases the agent's utility by a constant. Since the signal is verifiable, it also serves as a proof of the effort spent on self-monitoring. Therefore, the most efficient incentive provision scheme involves compensating the agent for the costs of monitoring only after the realization of the signal. Another important aspect is that the continuation utility in one contingency is independent of the approval rate in the other contingency. In other words, U_s is independent of the choice of x_{ns}, and U_{ns} is independent of x_s.

After figuring out the relation between approval probabilities and the continuation utilities, it is now possible to discuss the domain of value function V_D. By using the equations (1.2), and (1.3), one can see that the maximum value that U_D can take is equal to $\delta v + \frac{(1-\delta)c}{\mu\lambda}$, which can be achieved by setting x_s and U_s equal to their maximum values, 1 and v respectively. Therefore, the domain of V_D is the interval $[0, \delta v + \frac{(1-\delta)c}{\mu\lambda}]$.

To characterize the solution of the problem $\mathcal{P}_D$, one needs to use the equations (1.2), and (1.3) that govern the trade-off between the continuation utilities and the current approval rates for both contingencies, i.e., self-disclosure and no self-disclosure. The question is, to what extent the principal would like to use efficient approval decisions, i.e. $x_{ns} = 1$ and $x_s = 0$? It turns out that the approval decisions will be set as close as possible to the efficient ones. More precisely, $x_s = 0$ and $x_{ns} = 1$ as long as the resulting continuation utilities, i.e., U_s and U_{ns}, stays inside the state space $[0, v]$. This requires U_D to be in an intermediate range. When U_D is sufficiently small, setting $x_{ns} = 1$ is not feasible, since the resulting U_{ns} would be negative. For these values, the approval rate x_{ns} will be chosen such that the continuation utility U_{ns} becomes 0. By the same logic, for those values of U_D that are sufficiently large, the approval rate x_s will be chosen so that the continuation utility U_s takes its largest possible value v. The formal statement of the lemma is given by the following lemma.

Lemma 1.2. *There exists critical values $\underline{U} = (1-\delta)v$, and $\bar{U} = \delta v - \frac{(1-\delta)c}{\mu\lambda}$, such that the solution to the problem $\mathcal{P}_D$ satisfies:*

$$(x_s, x_{ns}) = \begin{cases} (0, \frac{U_D}{(1-\delta)v}) & \text{if } U_D \leq \underline{U}, \\ (0, 1) & \text{if } U_D \in (\underline{U}, \bar{U}), \\ (\frac{U_D + \frac{(1-\delta)c}{\mu\lambda} - \delta v}{(1-\delta)v}, 1) & \text{if } U_D \geq \bar{U}. \end{cases}$$

$$(U_s, U_{ns}) = \begin{cases} (\frac{U_D}{\delta} + \frac{(1-\delta)c}{\delta\mu\lambda}, 0) & \text{if } U_D \leq \underline{U}, \\ (\frac{U_D}{\delta} + \frac{(1-\delta)c}{\delta\mu\lambda}, \frac{U_D - (1-\delta)v}{\delta}) & \text{if } U_D \in (\underline{U}, \bar{U}), \\ (v, \frac{U_D - (1-\delta)v}{\delta}) & \text{if } U_D \geq \bar{U}. \end{cases}$$

Proof. See appendix A.2. □

Focusing on the approval rates in isolation, the principal prefers to increase x_{ns} and

decrease x_s as much as possible due to the efficiency concerns. However, these approval rates also alters the continuation utilities, therefore there is a non-trivial tradeoff that the principal needs to take into account. As can be seen from the equations (1.2), and (1.3), a higher x_{ns} requires a lower U_{ns}, and a smaller x_s requires a higher U_{ns}.

Lemma 1.2 proves that, even if there is a loss resulting from a lower U_{ns}, the tradeoff always favors a higher approval rate x_{ns}. Similarly, even if there is a loss resulting from higher U_s, the tradeoff always favors lower x_s. The lower and upper bounds of the derivative of the value function V, which are defined in lemma 1.1, are the main driving force behind this result. Precisely, the upper and lower bounds of V' puts a limit on the maximum loss that can arise, and this limit is always less than the gain from employing more efficient approval decisions. Note that the $\underline{U} < \bar{U}$ is always satisfied due to the restriction imposed on the discount factor δ.

1.3.5 The problem $\mathcal{P}_N$

This problem is defined conditional on no delegation in the current period. Its solution follows from straightforward arguments. The decision is mainly about how much of the promised utility, U_N, to provide the agent in the current period in the form of project approval, and how much of it to leave as a continuation utility. The amount that is left as a continuation utility will be granted to the agent starting from the next period without any restriction on the delegation decision. The optimal choice of the approval rate follows from the following maximization problem.

$$\max_{x_n} (1-\delta)x_n(v-\mu) + \delta V(\tfrac{U_N-(1-\delta)x_n v}{\delta}).$$

The shape of the value function V plays a crucial on the optimal choice of x_n and hence

on U_n. Let $I = [\underline{a}, \bar{a}]$ be an interval where the boundaries $\underline{a}$ and $\bar{a}$ satisfy:

$$\underline{a} = \inf\{U \in [0, v] \mid V'(U) \leq \frac{v-\mu}{v}\},$$

$$\bar{a} = \sup\{U \in [0, v] \mid V'(U) \geq \frac{v-\mu}{v}\}.$$

In words, I is the interval over which the derivative of the value function V is equal to $\frac{v-\mu}{v}$. This interval resides in the interior of the state space $[0, v]$. This stems from the fact that, the line that connects the points $(0, V(0))$ and $(v, V(v))$ has the slope $\frac{v-\mu}{v}$, and the graph of the value function V locates over this line.[11] Then due to the concavity, $V'(0) > \frac{v-\mu}{v} > V'(v)$, hence the interval I is in the interior of the state space $[0, v]$.[12] Then, one can conclude that the optimal choice of x_n satisfies the following:

$$x_n(U_N) = \begin{cases} 0 & \text{if } U_N \leq \delta\underline{a}, \\ \in (0,1) & \text{if } U_N \in (\delta\underline{a}, \delta\bar{a} + (1-\delta)v), \\ 1 & \text{if } U_N \geq \delta\bar{a} + (1-\delta)v. \end{cases} \tag{1.4}$$

Intuitively, for smaller values of U_N, the continuation value U_n will be in a range where the derivative of V is sufficiently large. In this region, it is better for principal to keep the continuation utility as high as possible by setting $x_n = 0$. On the contrary, for larger values of U_N, the value of V' becomes small, and increasing the continuation utility does not benefit the principal. Therefore, she keeps the continuation utility of the agent as small as possible by setting $x_n = 1$.

In the intermediate range of U_N, however, the value of x_n is set so that the continuation value lies in the interval I, and hence the value function has a derivative equal to $\frac{v-\mu}{v}$. The principal is indifferent between marginally increasing x_n and U_n. If the interval I consists

[11] The linear line can be achieved by an always feasible policy: randomizing between its extreme values. Hence it is strictly dominated, and V stays on top of this line.

[12] The interval I can be a singleton.

of a single point, then the optimal value of x_n is also singleton. Otherwise, there is a continuum of optimal values for x_n in this intermediate range.

1.3.6 The problem $\mathcal{P}'$

From now on, the conditional problems $\mathcal{P}_D$ and $\mathcal{P}_N$ will be considered together in order to characterize the solution of the unconditional problem. The task is to figure out the optimal way to decompose U into U_D and U_N together with the the optimal choice of γ. Completing these tasks will lead to the description of the value function V, which is equal to the concavification of the functions V_D and V_N.[13]

The following lemma describes the set of state variables over which the equality $V = V_N$ is satisfied. It further indicates that the corresponding approval rates at these state variables must be either 0 or 1. In other words, having an interior probability of approval, i.e. $x_n \in (0, 1)$, after avoiding self-monitoring never happens in an optimal policy.

Lemma 1.3. *There are two critical values* $0 < \underline{U}_N < \bar{U}_N < v$*, such that:*

i) $V(U) = V_N(U)$ *if and only if* $U \in [0, \underline{U}_N] \cup [\bar{U}_N, v]$

ii) V *is linear over* $[0, \underline{U}_N]$ *and* $[\bar{U}_N, v]$.

iii) The optimal approval decision x_n *satisfies:*

$$x_n(U) = \begin{cases} 0 & \text{if } U \in [0, \underline{U}_N] \\ 1 & \text{if } U \in [\bar{U}_N, v] \end{cases}$$

Proof. See appendix A.3 □

[13]The value function V depends on V_D and V_N, which in turn depend on V. In this regard, V, V_D, and V_N are the solutions of a complicated fixed point problem. However, a solution for this problem exists since the existence of the solution for the problem $\mathcal{P}$ is guaranteed.

Lemma 1.3 points out that making an uninformed decision without monitoring can be optimal only if the promised utility is close to the boundaries of the state space. It is already known that the equality $V = V_N$ holds at the extreme values of the state space, i.e., at 0 and v. By using the comparison between the V_N' and V', which can be achieved from the equation (1.4), one can show that the equality $V = V_N$ can hold only over the union $[0, \delta\underline{a}] \cup [\delta\bar{a} + (1-\delta)v, v]$. Then by using the expression (1.4), one can conclude that the approval probability x_n must be either 0 or 1 for all the promised utilities at which no-delegation is optimal. The second step of the proof shows the existence of the cutoffs $\underline{U}_N$, and $\bar{U}_N$, and also the fact that they reside in the interior of the state space.

This result is rather intuitive. When U is sufficiently low, it is not feasible to utilize a large approval rate; similarly, when U is sufficiently large, it is not feasible to employ a large rejection rate. Therefore, the principal can get only limited benefit from the information in these state variables, and the extent of this benefit is not sufficient to compensate the cost of acquisition. For this reason, she does not ask the agent to self-monitor at these promised utilities as she also cares about the costs. Therefore, no-delegation is an optimal solution, and hence $V = V_N$ holds.

The result is also informant about the shape of the value function over the lower and upper ends of the state space, as it points out that the value function V is linear over the intervals $[0, \underline{U}_N]$, and $[\bar{U}_N, v]$. Following some simple logic, one can see that the linearity is not limited to these intervals. More precisely:

$$V(U) = V_N(U) = \begin{cases} \delta V(\frac{U}{\delta}) & \text{if } U \in [0, \underline{U}_N] \\ (1-\delta)(v-\mu) + \delta V(\frac{U-(1-\delta)v}{\delta}) & \text{if } U \in [\bar{U}_N, v] \end{cases}$$

Then, by taking the derivative of both sides, one can get the following:

$$V'(U) = V'_N(U) = \begin{cases} V'(\frac{U}{\delta}) & \text{if } U \in [0, \underline{U}_N] \\ V'(\frac{U-(1-\delta)v}{\delta}) & \text{if } U \in [\bar{U}_N, v] \end{cases}$$

Since $V'(U) = V'(\frac{U}{\delta})$ for every $U \in [0, \underline{U}_N]$, and V is concave, the slope is constant over $[0, \underline{U}_N] \cup [\underline{U}_N, \frac{\underline{U}_N}{\delta}]$. Similarly, the slope of V is constant over $[\frac{\bar{U}_N-(1-\delta)v}{\delta}, \bar{U}_N] \cup [\bar{U}_N, v]$. Due to the linearity of the value function, it is without loss of generality to assume that there is a randomized delegation decision over the intervals $(\underline{U}_N, \underline{U}_D)$ and $(\bar{U}_D, \bar{U}_N)$.

The constant slopes, however, does not carry beyond the values $\frac{\underline{U}_N}{\delta}$, and $\frac{\bar{U}_N-(1-\delta)v}{\delta}$.[14] As a result, no randomized delegation decision can be optimal at these values, since they do not have any neighborhood over which the slope of V stays constant. At these values, $\gamma = 0$ can not be optimal either, because it is already known that $V_N < V$ for all of the state variables that are outside of $[0, \underline{U}_N] \cup [\bar{U}_N, v]$ form the previous lemma. As the only remaining option, the equality $V = V_D$ must hold at these values. To this respect, one shall define:

$$\underline{U}_D = \frac{\underline{U}_N}{\delta}$$
$$\bar{U}_D = \frac{\bar{U}_N - (1-\delta)v}{\delta}$$

The values $\underline{U}_D$, and $\bar{U}_D$ are the smallest and largest values of U satisfying $V(U) = V_D(U)$ without loss of any generality. The interval $(\underline{U}_D, \bar{U}_D)$ is the only remaining region where the delegation decision is yet to be described. However, it is natural to expect that $V = V_D$ and hence $\gamma = 1$ over $(\underline{U}_D, \bar{U}_D)$.

Let $U \in (\underline{U}_D, \bar{U}_D)$. From lemma 1.3, $V_N < V$, and hence $\gamma > 0$ at this U. Moreover, a randomized delegation decision cannot be optimal either. Suppose otherwise to get a contradiction, and assume that a randomization takes place between $V_D(U_D)$ and $V_N(U_N)$.

[14] If V' were to be constant on any neighborhood of $\frac{\underline{U}}{\delta}$ and $\frac{\bar{U}_N-(1-\delta)v}{\delta}$, then the equality $V = V_N$ would hold for a larger subset of $[0, v]$, which contradicts with the definition of $\underline{U}_N$ and $\bar{U}_N$.

The corresponding value of U_N must belong to either $[0, \underline{U}_N]$ or $[\bar{U}_N, v]$; without loss of generality assume the precedent. Then, the principal could rather randomize between $V_D(\underline{U}_D)$ and $V_D(U_D)$ and get strictly better off. The reason for this stems from the fact that slope of V alters at $\underline{U}_D$, hence the line connecting the values $V_D(U_D)$ and $V_D(\underline{U}_D)$ locates on top of the line connecting $V_D(U_D)$ and $V_N(U_N)$. Therefore, it is optimal to set $\gamma = 1$ and hence $V = V_D$ in this region.

The optimal delegation decision for each possible promised utility $U \in [0, v]$ is now known, and summarized in the following diagram:

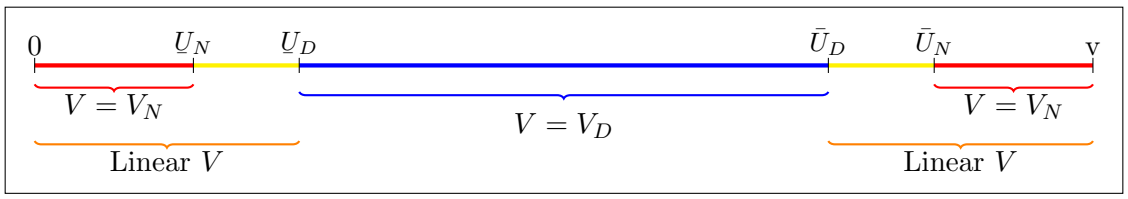

Figure 1.4: Partition of the state space depending on the optimal delegation decision.

To complete the characterization of the optimal policy, one should also describe the approval decisions. Particularly, the description of x_s and x_{ns} over the interval $[\underline{U}_D, \bar{U}_D]$, where delegation is the optimal choice, is incomplete.

The problem $\mathcal{P}_D$ is analyzed in isolation, and its solution is provided in lemma 1.2. The result asserts that the efficient decision making, i.e. $x_s = 0$, and $x_{ns} = 1$ will take place as long as the promised utility U is in between $[\underline{U}, \bar{U}]$. For the rest of the state space, making the efficient decisions is not feasible. Instead, the principal employs the decisions that are closest to the efficient among the feasible ones. However, the problem $\mathcal{P}_D$ is defined conditional on delegation in the current period regardless of its optimality. Now, the focus is on the interval $[\underline{U}_D, \bar{U}_D]$, where delegating monitoring is optimal.

The question one shall inquire at this point is, how do the values of $\underline{U}_D$, and $\bar{U}_D$ compare to the values of $\underline{U}$, and $\bar{U}$ respectively? The answer to this question is important

as it indicates whether the approval decisions taking place in an optimal policy are always efficient or not. This comparison will also be informative about the long-run outcome of the optimal policy. It turns out that $\underline{U} \geq \underline{U}_D$ and $\bar{U} \leq \bar{U}_D$. Therefore, there are some state variables over which the optimal policy employs an inefficient approval decision, i.e. $x_{ns} < 1$ or $x_s > 0$. The next lemma summarizes these findings:

Lemma 1.4. *In an optimal policy,* $\forall U \in [\underline{U}_D, \bar{U}_D]$, $V(U) = V_D(U)$, *and* $\gamma(U) = 1$. *Moreover,* $\underline{U}_D \leq \underline{U} < \bar{U} \leq \bar{U}_D$.

Proof. See appendix A.4 □

Now, the optimal delegation and approval decisions, and the transition rule over the entire state space is known. Once the policy is initiated with an initial promised utility for the agent, the rest follows immediately. The question is how to choose this initial state variable. In other words, what is the utility level that the principal promises to the agent in the beginning? To answer this question one shall look for the utility level that is maximizing the value function V.

The value function V increases over $[0, \underline{U}_D)$, as it has a positive constant slope; and decreases over $(\bar{U}_D, v]$. The initial utility, which is denoted by U^*, must be an element of the interval $[\underline{U}_D, \bar{U}_D]$. Due to the continuity, the existence of U^* is guaranteed. Moreover, I show that the value function V is strictly concave over $[\bar{U}_D, \bar{U}_D]$, hence U^* is unique. This result, and the earlier findings are summarized in the following theorem; which, at the same time, describes an optimal policy.

Proposition 1.1. *The value function* V *is strictly concave over* $[\bar{U}_D, \bar{U}_D]$, *and attains its maximum at* $U^* \in (\bar{U}_D, \bar{U}_D)$. *At* $t = 0$, *the agent is promised an expected utility of* U^*, *and delegated for self-monitoring.*

As long as the promised utility stays in the interval $(\bar{U}_D, \bar{U}_D)$, *the monitoring is delegated to agent, i.e.* $\gamma = 1$. *Moreover, the approval decisions and the transition rule for the state*

variable satisfies:

- $x_s(U) = \max\{0, \frac{U+\frac{(1-\delta)c}{\mu\lambda}-\delta v}{(1-\delta)v}\}$, *and* $x_{ns}(U) = \min\{1, \frac{U}{(1-\delta)v}\}$, $\forall U \in [\bar{U}_D, \bar{U}_D]$.
- $U_s(U) = \min\{v, \frac{U}{\delta} + \frac{(1-\delta)c}{\delta\mu\lambda}\}$, *and* $U_{ns}(U) = \max\{0, \frac{U}{\delta} - \frac{(1-\delta)v}{\delta}\}$, $\forall U \in [\bar{U}_D, \bar{U}_D]$.

When the state variable reaches to $[0, \underline{U}_D)$, *the monitoring is delegated to agent with probability* $\gamma = \frac{U}{\underline{U}_D}$. *If the realized delegation decision is:*

- *Delegation:* $(x_s, x_{ns}) = (0, \frac{\underline{U}_D}{(1-\delta)v})$, *and* $(U_s, U_{ns}) = \left(\frac{U}{\delta} + \frac{(1-\delta)c}{\delta\mu\lambda}, 0\right)$
- *No delegation:* $x_n = 0$, *and* $U_n = 0$

When the state variable reaches to $U \in (\bar{U}_D, v]$, *the monitoring is delegated to agent with probability* $\gamma = \frac{v-U}{v-\underline{U}_D}$. *If the realized delegation decision is:*

- *Delegation:* $(x_s, x_{ns}) = (\frac{U+\frac{(1-\delta)c}{\mu\lambda}-\delta v}{(1-\delta)v}, 1)$, *and* $(U_s, U_{ns}) = \left(v, \frac{U}{\delta} - \frac{(1-\delta)v}{\delta}\right)$
- *No delegation:* $x_n = 1$, *and* $U_n = v$

Proof. See appendix A.5 □

All of the components of the above theorem, except the strict concavity, are already discussed. Strict concavity of V over the interval $[\underline{U}_D, \bar{U}_D]$ follows from technical arguments that are depicted in the appendix. The following figure provides an illustration of the value functions V_D and V_N. The unconditional value function, V, is the concavification of them.

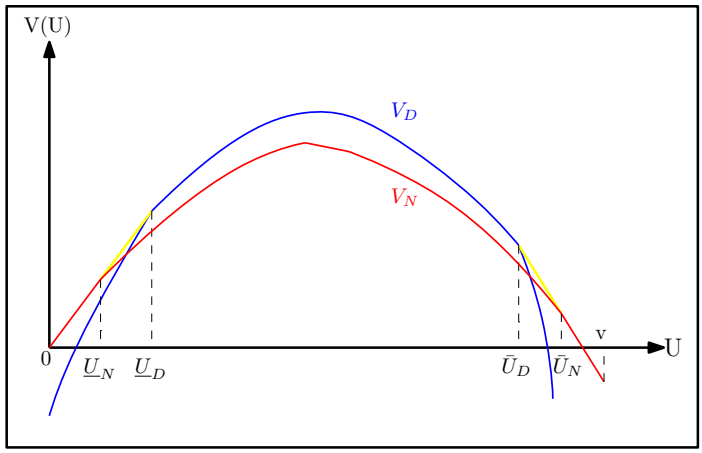

Figure 1.5: Illustration of conditional value functions V_D and V_N. The unconditional value function V is the upper envelope of V_D and V_N. The equality $V = V_N$ holds over the intervals $[0, \underline{U}_N]$ and $[\bar{U}_N, v]$. The equality $V = V_D$ holds over the interval $[\underline{U}_D, \bar{U}_D]$. For the rest of the state space V is strictly larger than V_N and V_D.

The policy described in the theorem is an optimal policy, and it is not the only one. There is a multiplicity originating from the linearity of the value function over the lower and upper ends of the state space. One can come up with another optimal policy that employs different policy specifications over these regions, where V is linear. For example, when $U \in [0, \underline{U}_N]$, it is also optimal to set $\gamma = 0$, $x_n = 0$, and $U_n = \frac{U}{\delta}$. Over the intermediate region, however, the value function is strictly concave, and there is no such multiplicity.

The specific policy given in the theorem is easy to describe and has some substantiation properties. It features a two-phase structure, in which all of the periods in which self-monitoring takes place are front-loaded. The following corollary describes this in detail.

Corollary 1.1 (Two-Phase Structure)**.** *The optimal policy depicted above consists of two consecutive phases:*

i) The initial phase, starting at $t = 0$, is where all the information acquisition takes place. The principal delegates monitoring to the agent as long as the policy stays in

this stage. Efficient decision making will be employed as long as it is available. The duration of this stage is stochastic and depends on the realized information outcomes.

ii) The terminal phase is no-delegation phase. Once it is reached there is no information acquisition anymore, and all the remaining approval decisions are uninformed. The projects are either always approved or always rejected depending on the realized outcomes during the initial phase.

As long as the delegation takes place, the agent is promised a higher continuation utility (in the form of future regulatory approval) each time he discloses bad news. His current project is less likely to get approved but the regulator promises more frequent approval in the future. If he does not disclose any signal, he is downgraded to a lower continuation utility. His current project has higher chances of approval, yet he will be given less frequent approvals in the future. The duration of this phase is stochastic; when it ends, the policy reaches a second phase in which there is no more self-monitoring.

If the policy reaches to the minimum promised utility the agent is black-listed and all of the projects will be directly rejected afterwards. On the contrary, if it reaches to its maximum, the agent is whitelisted and all the projects will be directly approved afterwards.

Such a structure is also observed in the papers Horner and Guo (2015), and (Li et al., 2015). The long run outcome consists of either a permanent rewarding or a permanent punishment depending on the earlier course of the relation. As it will be apparent shortly, this structure is not prevailing when the principal does not internalize the cost of information acquisition.

Efficiency aspects of the optimal policy are also significant. Likewise the papers referred in the previous paragraph, all the inefficiencies are back-loaded in the optimal policy that is defined in theorem 1.1. In this framework, an inefficiency arises in two consequences: first one arises if information acquisition is not requested, and the second one occurs due to

inefficient use of the acquired information. Clearly, a first type inefficiency appears only in the terminal phase of the policy. On the other hand, a second type inefficiency can occur only the last period of the initial phase. When the promised utility is in the interval $[\underline{U}_D, \underline{U})$, which is sometimes an empty set depending on the parameter values, self monitoring takes place, and the approval decision following a no self-disclosure is inefficient since $x_{ns} < 1$. Right after this event, the state variable is downgraded to its minimum, hence the ultimate stage with a permanent rejection starts. Analogous reasoning works for $(\bar{U}, \bar{U}_D]$ as well. Therefore, all the second type of inefficiencies takes place right before the ultimate stage starts. To this respect, all the inefficiencies are back-loaded.

Some of the properties mentioned above are specific to the policy depicted in the theorem. The long run outcome, on the other hand, is independent of the policy choice. It always features whitelisting or blacklisting. This stems from the fact that $\underline{U}_D \leq \underline{U} < \bar{U} \leq \bar{U}_D$.[15]The following result indicates this observation.

Corollary 1.2. *If the principal cares about the cost of monitoring, the eventual outcome features either permanent approval or permanent rejection in any optimal policy. Therefore, in the long run, the agent is either blacklisted or whitelisted with probability 1.*

1.4 Case 2: Principal Does Not Internalize Cost

Now, the principal does not internalize the agent's cost of information acquisition. In this case, the optimal policy turns out to be remarkably different in terms of its long run properties and the efficiency aspects. These discrepancies have their origins on the principal's richer capacity of providing incentives. From a technical point of view, principal can use monitoring requests purely on the grounds of punishing the agent, as she does not care about its cost. From time to time, the agent is asked to monitor himself, even

[15]If it was $\underline{U}_D > \underline{U}$ or $\bar{U} > \bar{U}_D$, then it would be possible to specify an optimal policy that does not feature blacklisting or whitelisting respectively.

though the resulting information will be completely ignored in the approval decision. Such a punishment scheme was feasible for the the previous case as well; however, as it also punishes the principal, it is not used in the optimal policy. The verifiability of the "bad news" plays a crucial role on this additional channel of incentive provision. Precisely, the signal provides a hard evidence for the efforts the agent spent on self-monitoring; therefore, the principal can make sure that the necessary punishment is executed.

The formal description of the principal's problem is barely changed, with a minor modification due to principal's shifted preferences towards monitoring costs. The objective function does not include the cost parameter, c, now. To recognize the differences, all of the variables are denoted with a tilde now. More precisely, the policy consists of the functions $\tilde{\gamma}, \tilde{x}_n, \tilde{x}_s, \tilde{x}_{ns}, \tilde{U}_n, \tilde{U}_s, \tilde{U}_{ns}$.

1.4.1 Observable Information Acquisition

The initial focus is on the benchmark where the agent's effort is observable. This will be helpful to reveal how this case is different than the previous one. The first best policy is as before, inducing the efficient outcome, i.e. monitoring and efficient decision making in each period with a stationary payment scheme. The agent's utility form this first best outcome is the same as before, i.e. equal to w; yet, the principal has a different utility since she does not care about the cost of monitoring. Her utility is denoted by $\tilde{\pi}$, and is equal to:

$$\tilde{\pi} = (1 - \mu\lambda)v - \mu(1 - \lambda).$$

The benchmark policy, which is defined conditional on the agent's ex-ante expected utility when effort is observable, is now different then the benchmark policy of the previous case. The approval decisions employed after information acquisition are not always efficient. The principal requests self-monitoring more often since she does not care about its costs.

The benchmark policy is stationary as before, hence there is no need to define the continuation utilities. The benchmark policy only consists of delegation and approval decisions, which are denoted by $\tilde{\gamma}^*, \tilde{x}_n^*, \tilde{x}_s^*, \tilde{x}_{ns}^*$.

When $U < w$, the principal, while keeping her promise, cannot employ the efficient scheme, because the agent is promised a utility that is less then the utility of the first best outcome. She can decrease the probability of self-monitoring and employ the direct rejection with positive probability instead. This was exactly what had been done in the previous case. Alternatively, she can decrease the probability of approval without decreasing the probability of delegation. Since she does not internalize the costs of self-monitoring, this alternative method turns out to be better for the principal. First, note that the probability of wrong approval, $\tilde{x}_s^*$, is always 0. Therefore, the principal only needs to choose $\tilde{\gamma}^*$ and $\tilde{x}_{ns}^*$ properly so that she keeps her promise. In doing so, her only concern is to maximize the probability of true approval, $\tilde{x}_{ns}^*$, since she does not care about the cost of self-monitoring. And, the probability of true approval is maximized when the probability of delegation is at its maximum, conditional on U being less than w. Therefore, the principal always delegates the monitoring to agent and approves the project with some probability chosen specifically to meet the promised utility U. In other words, $\tilde{x}_{ns}^*$ satisfies:

$$(1 - \mu\lambda)\tilde{x}_{ns}^* v - c = U.$$

Therefore, for every $U < w$, the benchmark policy is characterized by: $\tilde{\gamma}^* = 1$, $\tilde{x}_s^* = 0$, and

$$\tilde{x}_{ns}^* = \frac{U + c}{(1 - \mu\lambda)v}.$$

When $U > w$, the first best policy is exactly the same with the one depicted in the previous case. The promised utility of the agent is higher than the utility of the first best outcome. The optimal way to grant the promised utility to the agent requires to employ the direct approval without delegation due to the same arguments provided in the previous

section. Consequently, there is a randomization between first best outcome and the direct approval without monitoring. The value function $\tilde{V}^*$ over $[w, v]$, is the linear line combining the points $(\tilde{\pi}, w)$ and $(v, v-\mu)$. Therefore, the benchmark policy can be summarized as follows:

$$\tilde{\gamma}^*(U) = \begin{cases} 1 & \text{if } U \in [0, w) \\ \frac{U-w}{\mu\lambda v + c} & \text{if } U \in (w, v] \end{cases}$$

$$(\tilde{x}_s^*(U), \tilde{x}_{ns}(U)) = \begin{cases} \left(0, \frac{U+c}{(1-\mu\lambda)v}\right) & \text{if } U \in [0, w) \\ (0, 1) & \text{if } U \in (w, v] \end{cases}$$

$$\tilde{x}_n^*(U) = \begin{cases} \text{Irrelevant} & \text{if } U \in [0, w) \\ 1 & \text{if } U \in (w, v] \end{cases}$$

Therefore the corresponding value function $\tilde{V}^*$ is given by:

$$\tilde{V}^*(U) = \begin{cases} \frac{w-U}{w}\frac{c(v-\mu_{ns})}{v} + \frac{U}{w}\tilde{\pi} & \text{if } U \in [0, w) \\ \frac{v-U}{v-w}\tilde{\pi} + \frac{U-w}{v-w}(v-\mu) & \text{if } U \in (w, v] \end{cases}$$

The following figure illustrates the value function $\tilde{V}^*$:

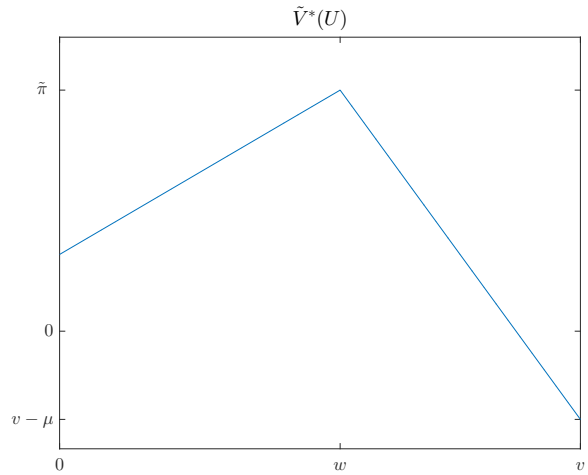

Figure 1.6: The value function $\tilde{V}^*$. The illustration assumes $v < \mu$.

1.4.2 Moral Hazard and the Optimal Policy

Now, I revert back to the agency problem in which the self-monitoring efforts are not observed by the principal. The observable effort benchmark is already informative about how the optimal policy would be different from the optimal policy given in previous case. As it will be clear shortly, the corresponding optimal delegation decision is different over the lower end of the state space, and this will significantly alter the transitional dynamics as well as the long-run outcome.

The description of the problem is barely changed. The cost is not included in the objective function; however, the incentive constraint and the promise keeping constraint remain the same. Following the exposition of the previous section, the problem is decomposed into two conditional sub-problems. The corresponding value functions are now denoted by $\tilde{V}, \tilde{V}_N$, and $\tilde{V}_N$.

For the maximal possible promised utility of the agent, permanent approval without any monitoring is still the only feasible policy, hence the optimal one. Therefore, $\tilde{V}(v) = \tilde{V}_N(v)$, and $\tilde{\gamma}(v) = 0$. On the contrary, for the minimal expected utility of the agent, permanent

rejection is not the optimal policy, unlike the previous case. The principal can do better by requesting self-monitoring. To see this, note that the incentive constraint is still binding, and the equations (1.3) and (1.2) are still valid. Then by plugging $U = 0$ into these equations, one can get:

$$\tilde{U}_{ns}(0) = -\frac{(1-\delta)v}{\delta}\tilde{x}_{ns}(0),$$
$$\tilde{U}_s(0) = \frac{1-\delta}{\delta\mu\lambda}c - \frac{(1-\delta)v}{\delta}\tilde{x}_s(0).$$

Clearly, $\tilde{x}_{ns}$ is equal to 0, since the continuation utility $\tilde{U}_{ns}(0)$ cannot be negative. On the other hand, after a self-disclosure, one possible policy is to set $\tilde{x}_s(0) = 0$ and $\tilde{U}_s(0) = \frac{1-\delta}{\delta\mu\lambda}c$.[16] Therefore:

$$\tilde{V}_D(0) \geq \delta\mu\lambda\tilde{V}(\frac{(1-\delta)c}{\delta\mu\lambda}) > 0$$

Therefore no delegation cannot be the optimal policy at the minimal promised utility, as it would require $V(0) = 0$. This situation alters the structure of the optimal policy over the lower end of the state space.

On the other hand, there is still a neighborhood of the maximal promised utility v over which no delegation is an optimal policy. In the outside of this neighborhood, $\tilde{V}_N$ is always strictly smaller than $\tilde{V}$, hence the probability of delegation is always strictly positive. In other words, not requesting self-monitoring can be an optimal policy only if the promised utility of the agent is sufficiently large. The following lemma summarizes these findings. And its proof follows the same logic used in the proof of lemma 1.3.

Lemma 1.5. *There is a critical value* $0 < \tilde{U}_N < v$*, such that:*

i) $\tilde{V}(U) = \tilde{V}_N(U)$ *if and only if* $U \in [\tilde{U}_N, v]$*, moreover* $\tilde{x}_n = 1$ *in this region.*

[16] Setting $\tilde{x}_s = 0$ is optimal at $U = 0$ as it will become clear shortly.

ii) $\tilde{V}$ is linear over $[\tilde{U}_N, v]$.

Proof. See appendix A.6 □

Using the same arguments provided in the previous section, one can show that the linearity of $\tilde{V}$ carries over to a larger interval, which is denoted by $[\tilde{U}_D, v]$ where $\tilde{U}_D = \frac{\tilde{U}_N - (1-\delta)v}{\delta}$. The slope of V changes at $\tilde{U}_D$, and the equality $\tilde{V}(\tilde{U}_D) = \tilde{V}_D(\tilde{U}_D)$ holds.[17]

Over the interval $[0, \tilde{U}_D]$, delegation is optimal for the principal. It is already known that this statement is valid for the boundaries of the interval, and the rest follows from the same idea used in lemma 1.4. The following diagram summarizes the findings so far.

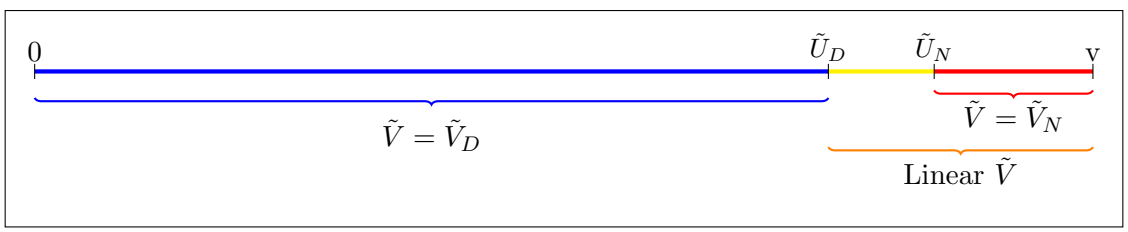

Figure 1.7: Illustration of the partition of the state space depending on the optimality of the delegation decisions, when the principal does not care about monitoring costs.

Due to the linearity of the value function over the interval $[\tilde{U}_D, v]$, there are many different possible ways to specify the optimal policy. The policy the paper focuses on assumes that whenever the promised utility reaches to this region, the principal randomizes between $\tilde{V}_N(v)$ and $\tilde{V}_D(\tilde{U}_D)$ with a properly chosen delegation probability. This completes the characterization of the optimal delegation decisions.

When it comes to the approval decisions, it is known that, if the agent is not asked to monitor himself, the project will be approved directly, hence $\tilde{x}_n$ is equal to 1. To achieve a

[17] The reasons for the latter are twofold. First, it is already known that $\tilde{V}(\tilde{U}_D) > \tilde{V}_N(\tilde{U}_D)$ from the previous result, hence $\tilde{\gamma} = 0$ cannot be optimal. Second, there is no neighborhood of $\tilde{U}_D$ over which $\tilde{V}'$ stays constant, hence a randomization cannot be optimal either.

complete characterization of the approval decisions, one also needs to describe the optimal values of $\tilde{x}_s$, and $\tilde{x}_{ns}$. For this, it is sufficient to focus on the interval $[0, \tilde{U}_D]$, where $\tilde{\gamma} = 1$ is the optimal decision.

Recall that the approval decision and the promised continuation utility are substitutes of each other, after both self-disclosure and no self-disclosure. More precisely, $\forall U \in [0, \tilde{U}_D]$:

$$(1-\delta)\tilde{x}_s v + \delta \tilde{U}_s = U + \frac{(1-\delta)c}{\mu\lambda}$$

$$(1-\delta)\tilde{x}_{ns} v + \delta \tilde{U}_{ns} = U$$

In the previous section, it was asserted that the principal employs the efficient decision making as long as the resulting continuation utilities are feasible. Bounds of the derivative of the value function was the main driving force behind this result. It was shown that the gain from increasing x_{ns} always dominates any possible loss due to the reduction on U_{ns}. Similarly, the gain from decreasing x_s always dominates any possible loss due to the amplification on U_s.

Here in this case, there is an obvious lower bound for $\tilde{V}'$, which is equal to the slope of the benchmark value function $\tilde{V}^*$ over $[w, v]$. And the logic given in the previous paragraph leads to the same conclusion for $\tilde{x}_s$. On the contrary, the slope of $\tilde{V}^*$ over the region $[0, w]$ does not necessarily constitute an upper bound for $\tilde{V}'$, since the equality $\tilde{V}^*(0) = \tilde{V}(0)$ does not hold. As a result, it is not clear whether or not the principal always prefers to set $\tilde{x}_{ns}$ as high as possible.

In principle, $\tilde{V}'$ can be sufficiently large, when U is sufficiently close to 0; and, hence, the gain from a higher continuation utility might dominate the gain from higher $\tilde{x}_{ns}$. Yet, it turns out that the trade-off is still in favor of $\tilde{x}_{ns}$. In other words, as long as it is feasible, setting $\tilde{x}_{ns} = 1$ is optimal. Otherwise $\tilde{x}_{ns}$ will be set such that the continuation utility hits its minimum.

Finally, the value function $\tilde{V}$ is strictly concave over the delegation region $[0, \tilde{U}_D]$. Therefore there is a unique state variable $\tilde{U}^* \in [0, \tilde{U}_D]$ maximizing $\tilde{V}$. And the value of $\tilde{U}_D$ is larger than $\bar{U}$, hence incorrect approval following a self-disclosure takes place with positive probability in the optimal policy. The following theorem, which constitutes the main result of this section, summarizes these findings:

Proposition 1.2. *The value function $\tilde{V}$ is strictly concave over $[0, \tilde{U}_D]$, and attains its maximum at some $\tilde{U}^* \in (0, \tilde{U}_D)$. At $t = 0$, the agent is promised the expected utility of $\tilde{U}^*$, and the monitoring is delegated to him.*

As long as the state variable stays in the interval $[0, \tilde{U}_D]$, the principal requests self-monitoring with probability 1. The approval decisions and the transition rule satisfies:

- $\tilde{x}_s(U) = \max\{0, \frac{U + \frac{(1-\delta)c}{\mu\lambda} - \delta v}{(1-\delta)v}\}$, *and* $\tilde{x}_{ns}(U) = \min\{1, \frac{U}{(1-\delta)v}\}$, $\forall U \in [0, \tilde{U}_D]$.
- $\tilde{U}_s(U) = \min\{v, \frac{U}{\delta} + \frac{(1-\delta)c}{\delta\mu\lambda}\}$, *and* $\tilde{U}_{ns}(U) = \max\{0, \frac{U}{\delta} - \frac{(1-\delta)v}{\delta}\}$, $\forall U \in [0, \tilde{U}_D]$.

When the state variable reaches to $(\tilde{U}_D, v]$, there will be a randomized delegation decision; the monitoring is delegated to agent with probability $\tilde{\gamma} = \frac{v-U}{v-\tilde{U}_D}$. If the realized randomized decision is:

- *Delegation:* $(\tilde{x}_s, \tilde{x}_{ns}) = (\frac{U + \frac{(1-\delta)c}{\mu\lambda} - \delta v}{(1-\delta)v}, 1)$, *and* $(\tilde{U}_s, \tilde{U}_{ns}) = (v, \frac{U}{\delta} - \frac{(1-\delta)v}{\delta})$
- *No delegation:* $\tilde{x}_n = 1$, *and* $\tilde{U}_n = v$

Proof. See appendix A.7. □

The dynamic properties of this policy are remarkably different now. Particularly, there is no permanent rejection state, and hence the agent is never blacklisted unlike the previous case. Instead, there is a state of probation, which occurs when the agent's promised utility reaches to its minimum. At this state, the project will be rejected for sure, yet the agent is still asked to monitor himself. This probation state keeps repeating itself until a self disclosure takes place. After a self-disclosure the agent is promoted to a higher promised

utility, and get out of probation for one period. Leaving the probation state today does not rule out the possibility of facing it again in the future.

Eventually, a terminal phase will be reached where there is no information acquisition anymore and all the projects are directly approved. This happens when the agent's promised utility reaches its maximum. The following corollary summarizes these implications of the optimal policy.

Corollary 1.3. *The optimal policy has a two-phase structure.*

i) The initial phase is where the agent is always asked to monitor himself. Acquired information is not always used for the current approval decision. The policy reaches to a probationary state with positive probability during this phase. In this probationary state, the content of the information will be completely ignored for the current approval decision. This probation state keeps repeating itself until a self-disclosure takes place. The duration of this stage is stochastic and depends on the outcomes of the self-monitoring.

ii) The terminal phase is no-delegation phase. Once it is reached there is no information acquisition anymore, and all the remaining approval decisions are uninformed. Projects are directly approved in this phase. This phases starts when the promised utility of the agent reaches its maximum.

The distribution of inefficiencies over time constitutes an important divergence from the previous case, and from the existing papers. The inefficiencies are not entirely back-loaded now. Especially, the second type of inefficiencies arise in a stochastic nonconsecutive order until the second phase of the policy is reached. To this respect, the efficiency will be gained and lost throughout time.

The eventual outcome described above is not specific to the policy that I focus. Since the inequality $\bar{U} \geq \tilde{U}_D$ is satisfied, the permanent approval is the eventual outcome of all

optimal policies. The following corollary points out this issue.

Corollary 1.4. *In any optimal policy, there is a unique long-run outcome featuring permanent approval of the projects without any request of self-monitoring. Therefore, in the long-run, the agent is whitelisted with probability 1.*

1.5 Limited Commitment

So far,I assumed that the principal has full commitment power in that she can commit to any incentive compatible policy. In this section I relax this assumption and consider a case where the principal's commitment power is limited. More precisely, I assume that she cannot commit to any policy with a negative continuation value. Precisely, at the beginning of each period, she needs to have a non-negative value in expectation. She still has within period commitment power and can fulfill any specified within period randomization. In other words, the limited commitment structure I impose does not rule out the possibility of having a negative realized value for some periods. For expositional convenience, I will call the model with full commitment power as the baseline model and denote the corresponding value function by V_B afterwards. Clearly, when the expected social cost of the projects is less then their economic benefits, i.e. $\mu \leq v$, an optimal policy for the baseline model is also optimal in this limited commitment environment. This stems from the fact that the continuation valuation of the principal never becomes negative in this case, i.e. $V_B(U) \geq 0$ for every $U \in [0, v]$.

On the contrary, when the expected cost is higher then the economic benefits, i.e. $\mu > v$, baseline policy can not be an optimal policy. Principal's continuation value becomes negative with positive probability, as in the outcome of whitelisting. She cannot promise the maximal utility to agent in this limited commitment environment. As a result, the optimal policy differs from the one of baseline model. From now on, the analysis will be based on the case $\mu > v$.

For a given history, the continuation of an optimal policy must also be optimal condition on the agent's expected utility. Moreover it maintains a non-negative value for the principal by definition. Therefore, it is still possible to represent principal's problem within a stationary form. The distinction is that the state space is endogenous and will be a proper subset of $[0, v]$. In describing this endogenous state space, the most prominent property one shall look for is the non-negativity of the corresponding value function for each value inside the state space. Indeed, the state space will be the maximum subset of $[0, v]$ that is satisfying this property. This is due to the fact that an optimal policy conditional on a specific state space is also a feasible policy under any larger state space. Hence enlarging state space without hurting the non-negativity can only improve the principal's policy.

In this regard, it is not hard to conclude that the state space is an interval and contains the utility level 0. Then, it is sufficient to find out the maximal utility U_{max} that the principal can promise to the agent without violating the limited commitment constraint. This would complete the characterization of the endogenous state space $[0, U_{max}]$. However, one needs to be sure about the existence of such a value. To this end, I define some auxiliary objects. First of all, for a given value $W < v$, it is known that the Bellman equation that is defined over $[0, W]$ is guaranteed to have a solution since Blackwell sufficiency conditions are satisfied. Let V_W be the corresponding value function arising from the solution of Bellman equation.

Lemma 1.6. *When principal has a limited commitment power, there exists a utility level U_{max} which is equal to the maximum utility that the principal can promise the agent in an optimal policy. Its value is given by:*

$$U_{max} = \sup\{W \mid V_W(W) \geq 0\}.$$

Moreover, the value function arising from an optimal policy satisfies:

$$V = V_{U_{max}}, \quad \textit{and} \quad V(U_{max}) = 0.$$

The value of U_{max} strictly decreases with μ, conditional on having an optimal policy inducing self-monitoring.

Proof. See appendix A.8 □

The proof of this result, first indicates that the existence of U_{max} is guaranteed. Then it is argued that $V_W(W)$ is continuous, and hence the resulting value function always takes non-negative values over the interval $[0, U_{\max}]$. The fact that the value of the principal at the maximal promised utility is equal to 0 follows from the continuity of $V_W(W)$. Finally, the proof points out an important observation in order to conclude the monotonicity of U_{max} with respect to μ. Incentive compatibility of an optimal policy is independent of the prior belief. Therefore an optimal policy at the maximal utility for some prior is also incentive compatible and brings a higher value to principal for any other smaller prior.

Then by using the arguments provided in the earlier sections one can conclude the following result.

Proposition 1.3. *When principal has a limited commitment power and the expected social cost of a project is larger than its value, whitelisting never appears in an equilibrium. Conditional on having an informative optimal policy,*

- *If principal internalizes the cost of self-monitoring, then the agent gets blacklisted eventually in the optimal contract.*
- *If principal does not internalize the cost of self-monitoring, then the optimal policy never reaches to a stable outcome and fluctuates over time.*

1.6 Concluding Remarks

This paper studies a regulatory system that incorporates self-monitoring. More precisely, it explores the behavior of regulators in environments where there is a significant uncertainty about the activities that the regulated agent carries on. The regulator, in an efficient regulatory regime, would like to use the information of the agent, who is superior in acquiring information. In order to incentivize the agent to acquire unfavorable information about his own activities, the regulator uses continuation values arising from future regulatory behavior as an incentive device.

I show that, when the regulator has full commitment power, self-monitoring can only be sustained in an initial phase of the optimal policy. During this phase, the agent is promised a higher continuation utility (in the form of future regulatory approval) each time he discloses "bad news"; otherwise, he is downgraded to a lower continuation utility in order to encourage him to acquire information. The eventual outcome crucially depends on the regulator's preferences over the cost of self-monitoring. If she internalizes this cost, both whitelisting and blacklisting are possible long-run outcomes; otherwise, whitelisting is the only long run outcome.

When the regulator has limited commitment power in that she can only commit to policies with non-negative continuation values, the results are remarkably different. When the expected social costs of the projects are larger than their economic yields, the policy does not feature whitelisting anymore. Furthermore, it is possible to sustain self-monitoring over the long-run in this case.

Chapter 2

Product Upgrades and Posted Prices

2.1 Introduction

The literature on durable good monopolies assumes a population of forward looking buyers with heterogeneous valuations for a unique product that a monopolist sells over time. Buyers, that are strategically timing their purchase decisions, have unit demand for the product, and hence leave the market after purchasing. In this general framework, there are two counterbalancing factors governing the pricing decision of the monopolist. First, to be able to sell the product to the agents with lower valuations, the monopolist must decrease the price of the good over time. Second, customers with high valuations might delay their purchases since they predict that the monopolist will decrease the price over time. Hence, decreasing prices is beneficial as it allows the firm to capture the surplus from the demand of the agents with lower valuations, but at the same time it is costly as it causes a delay in the purchases.

The pioneering paper of the literature, Stokey (1979), shows that, if the firm can commit to a price path before starting to sell, then the optimal price path is constant and equal to the static monopoly price. Consequently, all the purchases take place at the beginning of

the sales and so there is no delay. This result is significant as it asserts that the time is not used to discriminate over buyers with different valuations, which would have occurred with a decreasing price path and delayed purchases.

In this paper, we turn our attention to the optimal pricing problem when a new version of the existing good is expected to arrive at some point in the future. In contrast to one of the main assumptions of the existing literature, durable goods, for many real life examples, do not persistently stay in the market. Rather, newer and better functioning versions are taking place of the older ones over time. Technology companies such as Apple, Intel, Samsung, and Microsoft are good examples for this. This situation does not hurt the durability of the product that is replaced as customers can still use it after the new version is launched. However it alters the consumer's preferences as the newer product might offer more benefits. In such an environment, the structure of the buyers' incentives would be different than the ones in the classical framework. In particular, a buyer does not necessarily leave the market after purchasing a version of the good. He may rather prefer to stay to purchase the newer version as well when it is launched. Or he may abandon to purchase the current version to purchase the newer one. That is to say that the price path of a version of the good not only affects its own sales, but also the sales of the other versions.

We consider a monopolist which is selling two consecutive versions of the same good, with a restriction that it only sells the most current version at a given period. We assume that it has full commitment power, and can commit to a price path for both versions of the good at the beginning. The monopolist is initially endowed with the first version and the upgrade will take place over time. The timing of the upgrade is not a choice variable in this paper, rather we take it as exogenously given.

The analysis is divided into two parts depending on the specification of the arrival process of the new version of the good. In the first part, the arrival time is stochastic and follows a Poisson process. The optimal price path is shown to be constant for both versions of the good (at different levels) and the price level for the second version of the

good is independent from the realized arrival time. Consequently, there are no delays in any purchases. Any purchase of the first version occurs immediately at the beginning, and any purchase of the second version occurs as soon as it arrives into the market. This result is in line with Stokey (1979) as it also suggests that the time is not used as a tool to discriminate over buyers with different valuations. Stationarity is the main reason for this result to carry over to our setting.

In the second part, we considered the case in which the arrival of the newer version occurs at a commonly known certain date. This non-stationary environment comes with a cost of intractability, and to overcome this, we assume a binary type space for buyers' valuations. It is shown that, in this case, a decreasing price path is possible for some parameters and this gives rise to delayed purchases. That is to say that the time might be used as a discriminatory tool when the arrival of the new version occurs at a commonly known certain date.

Unlike the existing literature on durable good pricing with full commitment power, in our environment, a sale mechanism with posted prices is not an optimal mechanism. Strictly speaking, the anonymous structure of the posted prices (The sales of the new version of the good cannot be conditioned on the first version sales under the regime of posted prices.), comes at a cost for the monopolist. The optimal selling mechanism, on the other hand, requires the monopolist to bundle both versions of the good and selling them only together. More precisely, in this optimal sales mechanism there exists a group of consumers that are purchasing both versions of the good, and the rest of the consumers does not purchase anything. The resulting allocation of this policy cannot be implemented with posted price path which is anonymous by definition.

2.1.1 Literature Review

Our paper mainly fits into the literature of durable good pricing with full commitment power. The classic reference is Stokey (1979), which shows that the optimal price path is constant and equal to the static monopoly price.[1] Therefore all the buyers either purchase immediately or leave without any purchase, hence the time is not used by monopolist to discriminate over buyers with different valuations.

In our paper, there are both favoring and contradicting results with this mainstream result. We show that, under some circumstances, the optimal posted price path may be decreasing over time. Some other papers have also shown that the optimal posted price path, under full commitment power, may fluctuate over time. In particular, Board (2008) shows that, in a setting with stochastic population, if there is demand heterogeneity over the incoming population of buyers, then the optimal price path might fluctuate over time. This leads some buyers to delay their purchases, hence time is an effective discriminatory tool in such a situation. Another important paper showing a fluctuating price path is Garrett (2012) in which the flow valuations of buyers are stochastic due to the private circumstances. The environment is stationary, as neither the value distributions of the entering buyers nor the stochastic process governing the valuations are time dependent, yet the optimal price path is fluctuating over time. Like our paper, in Garrett (2012), the posted prices is not an optimal selling mechanism. The optimal mechanism involves selling option contracts to purchase the good at future dates. However, the source of the inefficiency regarding the posted prices is different than the one in this paper.

In addition, our paper is also related to the literature on dynamic auctions.[2] A common feature of this literature is that there is a certain time period T, until when the seller must sell his multi-unit indivisible goods. The buyers are entering into and leaving from

[1] Actually optimal price path is not unique, there are infinitely many price paths resulting with the same allocation. The important thing is that the initial price level is equal to the static monopoly price and the price level never falls below that.

[2] Bergemann and Said (2011) is a good survey about dynamic auctions.

the market over time. Pai and Vohra (2008), also Board and Skrzypacz (2010) are good examples. They first define the optimal allocations, which follows a simple index rule, and then show how to implement this via monetary transfers. The main difference of our paper from this literature is that, in ours, the number of goods that the seller can sell is not limited whereas in dynamic auction literature the seller is endowed with a certain number of goods.

There are couple of other topics incorporating some features that are conceptually related to ours. The literature of planned obsolescence is one of them. The monopolist that is selling a single good in a dynamic setting strategically arranges the lifetime of the durable good. There are benefits from producing goods with a shorter lifetime as it leads consumers to repurchase again. Bulow (1986) is a good example of this literature. There is an old literature investigating the effect of vintage capital on the aggregate growth of the economy. The basic story of these papers is that, in each period, the machines of production are being improved because of technological progress. And firms are deciding how much to invest in replacing old machines with the newer ones. An example of this literature is Benhabib and Rustichini (1991). Finally, there is a literature regarding the $R\&D$ planning of durable good monopolies, see for example Swan (1970), Fishman and Rob (2000). The main concern of these papers is to understand the product development decisions of the firms given a non-strategic buyers side, which is represented with a demand function on the quality of the durable good.

2.2 Model

Time, $t \in [0, \infty)$, is continuous and r is the common discount factor. There is a monopoly selling a durable good, an initial version of which exists in the marketplace at the beginning of the time (t=0), and a newer version will eventually take the place of the existing one. The process governing the arrival of the newer version will be described later. The firm

sells only the most current version of the good at a given time; in other words, when the new version arrives the firm is no longer able to sell the earlier version. The cost of the production is normalized to 0 for both versions of the good.

On the demand side, there is a unit mass of buyers that are heterogeneous in terms of their valuations. They can consume at most one unit of the good at a time, and all of them are having a higher value for the newer version of the good. Moreover, we assume that the ratio of valuations for two versions of the good is same for all buyers. In particular, there is a constant $\beta > 1$ such that, a type x buyer has a flow utility x and and a flow utility βx from consuming the first and the second versions of the good respectively. Buyers' types follows a continuously differentiable distribution function $F(x)$ over the unit interval $[0, 1]$. $F(x)$ has full support and a corresponding continuous density function$f(x)$. The buyers are strategically deciding the time of their purchase(s), and also which version(s) of the good to buy. The good is durable and so a buyer may use it forever after purchasing. However, since the flow utility of the newer version is higher, one may want to replace the old one with the newer one. Therefore, buyers do not necessarily leave the market upon purchase, unlike the existing literature. On the other hand, version-wise strategic delays of purchases may appear.Precisely, a buyer might prefer to wait for the arrival of the new version rather than buying the current version of the good.

The monopolist commits to a price path for both versions of the good at $t = 0$. The price path is consisting of a price level for the first version of the good for each time until the arrival of the new version and also a price level for the new version of the good for each time after its arrival contingent on its arrival time. Note that a posted price that is defined in this way indirectly implies anonymity: the monopolist has to charge the same price for every buyer that is purchasing at the same. In other words, the possibility of conditioning the new version sales to the buyers' ownership status of the old version is ruled out. This puts a restriction on the monopolist and hence the resulting optimal posted prices will not be the optimal mechanism. Nevertheless, as a benchmark, we also analyzed the case

without anonymity. The arrival process of the new version is modeled in two different ways:

2.2.1 Stochastic Arrival

The second version of the good arrives stochastically with a Poisson arrival process at rate λ, the realized arrival time is denoted by T. The price path that the firm commits at $t = 0$ is contingent on T. More precisely, it is of the form: $\left(\{p_t\}_{t\in[0,\infty)}, \left\{\{p_t^T\}_{t\in[0,\infty)}\right\}_{T\in[0,\infty)}\right)$. The first term is the single price path for the first good, since the arrival is stochastic it is defined over $t \in [0, \infty)$. The second term is the price path for the newer version. Note that there is a different price path for every possible arrival time. In particular, conditional on the arrival time T, the term p_t^T is the price level of the second version of the good at $T + t$, i.e. t period after the arrival T.

Each buyer decides whether and when to purchase the first version of the good, also whether and when to purchase the second version of the good for each possible arrival time T. More precisely, buyer x's decisions are of the form: $\left(t_x, \{t_{x,T}\}_{T\in(0,\infty)}\right)$. The term t_x is the purchase time of the first version of the good; hence if realized T is less than t_x, then it means that the buyer does not buy the first version. Therefore, if $t_x = \infty$, then it means that the buyer never purchases the first version of the good. The term $t_{x,T}$, on the other hand, specifies how long after the realized arrival time T, buyer x purchases the second version. Hence, the corresponding calendar time of the purchase is $T + t_{x,T}$. Again if $t_{x,T} = \infty$ then the agent does not purchase the second version if the arrival occurs at T.

The utility type x buyer, denoted by $U(x)$:

$$
\begin{aligned}
U(x) = &\int_0^{t_x} e^{-\lambda T}\lambda\left(\int_{T+t_{x,T}}^{\infty} e^{-rt}\beta x dt - e^{-r(T+t_{x,T})}p_{t_{x,T}}^T\right) dT \\
&+ \int_{t_x}^{\infty} e^{-\lambda T}\lambda\left(\int_{t_x}^{T+t_{x,T}} e^{-rt}x dt + \int_{T+t_{x,T}}^{\infty} e^{-rt}\beta x dt \ - e^{-rt_x}p_{t_x} - e^{-r(T+t_{x,T})}p_{t_{x,T}}^T\right) dT
\end{aligned}
$$

The first line captures the contingencies in which arrival occurs before t_x. The utility in these cases depends only on the timing of the purchase of the newer version. For each arrival time $T \in [0, t_x]$, the corresponding utility is the expression inside the parenthesis. Then, after weighting them with probability of arrival at T(i.e with $e^{-\lambda T}\lambda$), we integrate it to get the expectation. The second line accounts for the arrivals occurring after the purchase of the first version, where the expression inside the parenthesis represents the utility corresponding to a specific arrival time $T \in [t_x, \infty)$. In each contingency the agent acquires a flow utility x until the purchase of the new version, and βx afterwards. We also discount the payments and integrate them after weighting with the probability of arrival. Note that, if $t_x = \infty$, i.e the agent never purchases the first generation of the good, then the second line is irrelevant; similarly if $t_x = 0$ then the first line is irrelevant.

The profit of the firm, denoted by Π:

$$\Pi = \int_0^1 e^{-\lambda t_x} e^{-r t_x} p_{t_x} f(x) dx + \int_0^\infty e^{-\lambda T} \lambda \left(\int_0^1 e^{-r(T+t_{x,T})} p^T_{t_{x,T}} f(x) dx \right) dT.$$

First and second terms are the corresponding profits from the sales of first and second versions of the good respectively. For the first term, the discounted payment of each type of buyer is integrated over the type space. To discount the payment of type x buyer (i.e p_{t_x}), we multiply it by e^{-rt_x} and also by the probability of the event that the arrival does not occur until t_x, which is $e^{-\lambda t_x}$. For the second term, the inner integral is the level of profit resulting from a specific arrival time T; and the outside integral takes their expectations over possible arrival times.

2.2.2 Deterministic Arrival

In the second part of the paper, we assume that the arrival occurs at a certain time period T, which is commonly known. In this case, the monopolist commits to a single price path: $\{p_t\}_{t \in [0,\infty)}$, where p_t is the price level of the first(second) version of the good at time t if

$t < T$ ($t \geq T$). To the sequel, we characterize the optimal posted prices for both of these arrival processes.

2.3 Optimal Posted Prices: Stochastic Arrival

Before delving into the main concern of the paper, we consider some benchmarks to develop a better understanding of the general framework. First we illustrate the canonical durable good monopoly pricing problem, in which there is no product upgrade. The second benchmark considers the case where there is a product upgrade but the monopolist is not restricted to use posted prices. It can rather use any selling mechanism including the non-anonymous ones.

2.3.1 Benchmark I: Canonical Durable Good Monopoly

This benchmark is analyzed in Stokey (1979).[3] There is only one version of the durable good staying in the market forever. It is a special case of our model in several directions. For example, we can get this canonical model from ours by assuming that $\lambda = 0$, i.e. the newer version of the good never arrives; or by assuming $\beta = 1$, i.e. there is no distinction between the first and the second versions of the good for buyers. The monopolist chooses a unique price path $\{p_t\}_{t\in[0,\infty)}$, and agents decide the timing of their purchases t_x. Corresponding utility of the agent x is:

$$U(x) = \int_{t_x}^{\infty} e^{-rt} x dt - e^{-rt_x} p_{t_x} = \ e^{-rt_x}(\frac{x}{r} - p_{t_x}), \tag{2.1}$$

[3]Here unlike the analysis presented in Stokey (1979) we follow the general mechanism design approach. We first characterize the incentive constraints and then rewrite the firm's problem as an optimal allocation problem.

and the profit of the firm is:

$$\Pi = \int_0^1 e^{-rt_x} p_{t_x} f(x) dx.$$

Since the monopolist has full commitment power, his problem is a mechanism design problem. Thanks to the revelation principle we can restrict attention to the direct mechanisms. In particular, the firm asks agents to report their types, and then decides their allocations, i.e. a purchase time t_x, and a payment level p_x in an incentive compatible way. The payment for the agents that are purchasing the good at the same time must be the same, otherwise truthful reporting would not be incentive compatibility. Therefore, for each allocation time there is a corresponding payment level, i.e. we can denote the payments by p_t. The following Lemma, illustrating the nature of the incentive constraints, is an adapted version of the fundamental IC Lemma corresponding to the durable good pricing framework.

Lemma 2.1. *The direct mechanism is incentive compatible iff:*

$$i)\ t_x \text{ is non-increasing with } x.$$
$$ii)\ U(x) = U(0) + \frac{1}{r}\int_0^x e^{-rt_{\tilde{x}}} d\tilde{x}$$

Proof. See appendix B.1. □

Lemma 2.1 states that a higher type will not purchase the good at a later time than a lower type. It also asserts that the derivative of $U(x)$ is proportional to the effective discount ($\frac{1}{r}e^{-rt_x}$). Since these conditions are both necessary and sufficient for incentive compatibility, the monopolist's problem can be written as:

$$\max_{\{p_t\}_{t\in[0,\infty)},\{t_x\}_{x\in[0,1]}} \int_0^1 e^{-rt_x} p_{t_x} f(x) dx \quad s.t \ \cdot \ t_x \text{ is non-increasing with x.}$$
$$\cdot \ U(x) = U(0) + \frac{1}{r}\int_0^x e^{-rt_{\tilde{x}}} d\tilde{x} \ \ \forall x$$

We can further simplify the above problem, and get rid of the price terms in it. To this respect, by using equation (2.1) and Lemma 2.1 we get:

$$e^{-rt_x}p_{t_x} = e^{-rt_x}\frac{x}{r} - \frac{1}{r}\int_0^x e^{-rt_{\tilde{x}}}d\tilde{x}.^4$$

Therefore, the profit of the firm is equal to:

$$\Pi = \int_0^1 e^{-rt_x}p_{t_x}f(x)dx = \int_0^1 \big(^{-rt_x}\frac{x}{r} - \frac{1}{r}\int_0^x e^{-rt_{\tilde{x}}}d\tilde{x}\big)f(x)dx.$$

After integrating it by parts we get the new form of the problem as:

$$\max_{\{t_x\}_{x\in[0,1]}} \quad \frac{1}{r}\int_0^1 e^{-rt_x}\left(x - \frac{1-F(x)}{f(x)}\right)f(x)dx \quad \text{s.t} \;\; t_x \text{ is non-increasing.} \tag{2.2}$$

Note that, this problem consists of only the terms t_x's. Therefore, its solution gives us the optimal allocations, and then by using the incentive constraints we get the corresponding price path inducing the optimal allocations. The term $(x - \frac{1-F(x)}{f(x)})$ is referred as the virtual value of type x, and as can be seen the monopolist's problem boils down to maximization of the integral of the discounted virtual valuations. The optimal solution of the above problem is of the following form:

$$t_x = \begin{cases} 0 & \text{if } x \in [x^\star, 1] \\ \infty & otherwise \end{cases}$$

In words, there is a threshold type $x^\star$ such that all the buyers above this threshold are acquiring the good immediately, and the rest of the buyers are never getting the good. If the virtual valuation function, $x - \frac{1-F(x)}{f(x)}$, is increasing in x, then the threshold $x^\star$ would be the minimum value of x at which the virtual value function is equal to 0. This is rather intuitive since the monopolist's problem boils down to maximization of the integral of discounted virtual valuations. For a given type x, if the virtual value is positive(negative) then $t_x =$

[4] In optimal mechanism $U(0) = 0$, hence we can omit it.

$0(t_x = \infty)$ and this does not violate the monotonicity constraint. If virtual valuation function is non-monotonic, the optimal allocations follows a cutoff rule with immediate allocations as well. But in this case the monopolist will choose $x^\star$ in such a way that the integral of the virtual valuation function above $x^\star$ is maximized. This does not mean that all the buyers of a type higher than $x^\star$ have a positive virtual valuation though. An important thing to note here is that this allocation rule is exactly the same with the optimal allocation rule of the static monopoly.

Now, we need to find an optimal price path inducing this optimal allocation rule. Since $U(0) = 0$ in an optimal mechanism, $U(x^\star) = e^{-rt_{x^\star}}(\frac{x^\star}{r} - p_{t_{x^\star}}) = U(0) + \frac{1}{r}\int_0^{x^\star} e^{-rt_{\tilde{x}}} d\tilde{x} = 0$, which requires $p_{t_{x^\star}} = p_0 = \frac{x^\star}{r}$. Therefore an optimal price path to implement the optimal allocation is a constant price path at level $\frac{x^\star}{r}$. The importance of this price level is that the buyer with type $x^\star$ is indifferent between purchasing and not.[5] The significance of this result is that, for the dynamic setting, the optimal price is constant and equal to the static monopoly price. Therefore the monopolist does not use the time to discriminate over the buyers with heterogeneous valuations.

2.3.2 Benchmark II: Product Upgrade without Anonymous Prices

In this benchmark, the arrival process is a stationary stochastic process, but the monopolist is not restricted to use posted prices as the selling mechanism. More precisely, the monopolist can choose a mechanism in which he can condition the second version sales on the sales of the first version. The direct mechanism, in this case, contains a joint allocation rule for both versions of the goods and a corresponding payment for each type of buyer. The payment rule is denoted by $P(x)$ for a given reported type x and is to be paid at time $t = 0$.[6] Allocations for type x buyer are t_x, and $\{t_{x,T}\}_{T\in[0,\infty)}$, which are defined in the same

[5] Even though the optimal allocation is unique, there are infinitely many price paths that can implement it. The important thing is to fix the initial price level to $\frac{x^\star}{r}$ and always keep it above or equal to the initial level.

[6] Any dynamic payment rule with a present value equal to $P(x)$ would be an equivalent to this payment rule. Hence there is no loss of generality here.

way as before. Then the utility of the agent x, $U(x)$, can be written as:

$$U(x) = Q(x)x - P(x).$$

where

$$Q(x) = \int_0^{t_x} e^{-\lambda T}\lambda\, e^{-r(T+t_{x,T})}dT \; + \int_{t_x}^{\infty} e^{-\lambda T}\lambda\Big(\int_{t_x}^{T+t_{x,T}} e^{-rt}dt + \; \beta\int_{T+t_{x,T}}^{\infty} e^{-rt}dt\Big)dT$$

The term, $Q(x)$, can be considered as the total allocation that is resulting from the allocations of both versions of the good, and it can only take values from $[0, \frac{r+\beta\lambda}{r(r+\lambda)}]$. Its maximum value is acquired by arranging $t_x = 0$ and $t_{x,T} = 0\ \forall T$ (immediate allocation of both versions), and its minimum value is acquired by arranging $t_x = \infty$ and $t_{x,T} = \infty\ \forall T$ (no allocation of any versions of the good). Following the same steps as in the previous benchmark we get:

Lemma 2.2. *The direct mechanism without anonymity restriction is incentive compatible if and only if:*

$$i)\ Q(x) \textit{ is non-decreasing}$$
$$ii)\ U(x) = U(0) + \int_0^x Q(\tilde{x})d\tilde{x}$$

Proof. Follows exactly the same steps with the proof of lemma 2.1. □

Then, the monopolist's problem can be written as:

$$\max_{\{Q(x)\}_{x\in[0,1]}} \frac{1}{r}\int_0^1 Q(x)\left(x - \frac{1-F(x)}{f(x)}\right)f(x)dx \quad s.t \quad Q(x) \text{ is non-decreasing with x.}$$

The solution will be analogous to the one of the previous benchmark. In particular, there exists a threshold $x^\star$, which is equal to the threshold that is defined in the first benchmark, such that for all the buyers above(below) this threshold the value of $Q(x)$ is maxi-

mized(minimized). More precisely, the optimal allocation rule is:

$$i) t_x = t_{x,T} = 0 \quad \forall x \in [x^\star, 1]$$

$$ii) t_x = t_{x,T} = \infty \quad \forall x \in [0, x^\star)$$

In the optimal mechanism there is no buyer acquiring only one version of the good. In other words, the monopolist is bundling two generations of the good and selling them only together. This allocation rule resembles some selling strategies that we observe in real life. For instance, some companies, like Microsoft, offer discounts to their customers in case they have an old version and want to upgrade to a newer one. Here we see an extreme version of this policy in the sense that the price of the second version for those who already own the first version of the good is equal to zero.

The payment scheme inducing this allocation requires all the agents in $[x^\star, 1]$ to pay the same amount since all of them have the same $Q(x)$. This payment is equal to $x^\star \frac{r+\beta\lambda}{r(r+\lambda)}$, which leaves the marginal agent $x^\star$ indifferent between purchasing and not. This mechanism is the optimal selling mechanism. However, it is impossible to implement this allocation rule by using posted prices. To see this, suppose there is a contingent price path that can implement it anonymously. Then, under these prices, marginal return from the second version purchase for agent $x^\star$ is larger than or equal to the price at the corresponding time period. But since his marginal benefit is equal to $(\beta-1)x^\star$, the agent $x^\star-\epsilon$, for ϵ sufficiently small, would prefer to purchase the second version as well. Hence we get a contradiction.

2.3.3 Sales With Posted Prices: Anonymity

Here, the focus is on the characterization of the optimal posted prices, which is anonymous by definition. In this case, buyer x that is owning the first version of the good and the buyer $(\beta-1)x$ that is not owning the first version of the good will have the same marginal benefit from the newer version of the good. Therefore a mechanism corresponding to the posted

prices must treat these buyers in the same manner for the allocation of the second version of the good. Therefore a direct mechanism will have some further constraints in this case. For this reason, we rather use a different approach, which we call as "two-step mechanism", in which the allocations of each version of the good follows independent reporting stages. First, we define the following concept:

Definition 2.1. *The Effective type of buyer x at realized arrival time T, is equal to his marginal flow utility from the purchase of the second version of the good. Particularly, it is equal to βx if he does not own the first version and it is equal to $(\beta - 1)x$ if he does.*

The buyers can use at most one unit of the good at a time. Therefore a buyer, owning both versions of the good, uses only the current one as it gives more flow utility. Alternatively, we can think of these two versions of the goods as if they are two separate goods where both can be consumed at the same time, yet the flow utility of the second good is equal to the effective type that we described above. From this point on we exploit this interpretation in our analysis, as it simplifies the exposition. The two-step mechanism is defined as:

Definition 2.2. *The two-step mechanism is a mechanism in which buyers are asked to report twice. First, at $t = 0$, buyers are asked to report their types. Then the allocations and payments for the first version of the good are decided according to the first step reports. Second, at the realized arrival time T, buyers are asked to report their effective types and the allocations and the payments of the second version of the good are decided according to the second step reports independent from the first step reports.*

Finding the optimal mechanism among this class of mechanisms will give us the optimal posted prices that the monopolist can commit. The mechanism structure here is different than a direct mechanism, hence we slightly modify the notation specified earlier. For the second step allocations, contingent on the realized arrival time T, the amount of time after

which the effective type x purchases the second version of the good is denoted by t_x^T and so the corresponding purchase time is $T + t_x^T$. [7] For the first step allocations the relevant information is the initial type hence we keep the old notation, where t_x is the purchase time for type x.

The utility of an agent is has two parts, one for each step of the mechanism. Starting with the second step, contingent on the arrival time T, discounted expected utility of the effective type x calculated at T, denoted by V_x^T:

$$V_x^T = \int_{t_x^T}^{\infty} e^{-rt} x dt - e^{-rt_x^T} p_{t_x^T}^T \;=\; e^{-rt_x^T}\left(\frac{x}{r} - p_{t_x^T}^T\right).$$

For the first step, the expected utility of buyer of type x, calculated at t=0, denoted by V_x:

$$V_x = e^{-\lambda t_x}\int_{t_x}^{\infty} e^{-rt} x dt - e^{-(r+\lambda)t_x} p_{t_x} = e^{-(r+\lambda)t_x}\left(\frac{x}{r} - p_{t_x}\right).$$

Note that this expression reflects the alternative interpretation mentioned earlier. In particular, first version of the good is used forever after the purchase. Total utility of buyer x, which is denoted by $U(x)$, can be written as:

$$U(x) = V_x + \int_0^{t_x} e^{-(r+\lambda)T}\lambda V_{\beta x}^T dT + \int_{t_x}^{\infty} e^{-(r+\lambda)T}\lambda V_{(\beta-1)x}^T dT.$$

There is a crucial point that we better to stress out here: depending on the realized arrival time the effective type and hence the resulting second step utility of the buyer changes. If the arrival occurs before (after) t_x, then the effective type of the agent x is βx ($(\beta-1)x$) and the corresponding second step utility is $V_{\beta x}^T$ ($V_{(\beta-1)x}^T$).

The profit of the firms , Π, has also two components:

$$\Pi = \int_0^1 e^{-rt_x} e^{-\lambda t_x} p_{t_x} f(x) dx \;+\; \int_0^{\infty} e^{-(r+\lambda)T}\lambda \Pi^T dT. \tag{2.3}$$

[7] Note that the previous notation was $t_{x,T}$, for type x. We now take the effective types as our basis rather than the initial type.

The first and second terms are accounting for the expected profits from the first and second steps of the mechanism respectively. The term Π^T is the expected profit (calculated at time T) conditional on the arrival time T, and it is equal to:

$$\Pi^T = \int_0^1 e^{-rt_x^T} p_{t_x^T}^T f_T(x)dx.$$

The density $f_T(.)$ is the distribution of the effective types at the realized arrival time T and it is depending on the allocation time t_x as well as the realized T. To this respect, the monopolist by arranging the allocations of the first step of the mechanism, can also alter the distribution of the buyers' marginal flow benefits from the newer version of the good.

To characterize the incentive constraints of the buyers, we need to consider both reporting stages separately. Since buyers are forward looking, while reporting in the initial stage they will internalize the effect of their report on the second stage of the mechanism. Therefore, we start our characterization of incentive constraints from the second stage.

Lemma 2.3. *The second step of the mechanism is incentive compatible if and only if,* $\forall T$

$$i)\ t_x^T \text{ is non-incraeasing with } x$$
$$ii)\ V_x^T = V_0^T + \frac{1}{r}\int_0^x e^{-rt_{\tilde{x}}^T} d\tilde{x}.$$

Proof. Follows from the same arguments with the proof of Lemma 2.1. □

Assumption 2.1. $\lambda \leq \frac{r}{\beta-1}$.

Now we deal with the the incentive constraints of the first stage reports. Any deviation from truthful reporting at this stage will also alter the optimal behavior in the second stage as it changes the allocation time and hence the corresponding effective types.

Lemma 2.4. *A two-step mechanism with an incentive compatible second stage, is also*

incentive compatible at the first stage only if:

i) t_x *is non-increasing with* x

ii) $V_x \;=\; V_0 \;+\frac{1}{r}\int_0^x e^{-(r+\lambda)t_{\tilde{x}}}d(\tilde{x}) \;-\lambda\int_0^x e^{-(r+\lambda)t_{\tilde{x}}}\frac{\partial t_{\tilde{x}}}{\partial x}\Big(V^{t_{\tilde{x}}}_{\beta\tilde{x}} - V^{t_{\tilde{x}}}_{(\beta-1)\tilde{x}}\Big)d\tilde{x}$

Proof. See appendix B.2. □

There are four conditions in total, which are defined in lemma 2.3 and lemma 2.4, that the optimal two step mechanism needs to satisfy. Two conditions given in Lemma 2.3 are necessary and sufficient for the second step incentive compatibility, whereas two conditions given in Lemma 2.4 are just necessary for the first step incentive compatibility.[8] To the sequel, we define the problem of the monopolist by only taking these four conditions into account. As the latter two conditions are not sufficient for the first stage incentive compatibility, the solution of our problem does not necessarily be the optimal solution that we are looking for. However, the solution of our problem is shown to be incentive compatible, therefore it is also the optimal solution that we are looking for. By using lemma 2.3, and the fact that the monopolists sets $V_0^T = 0$ in an optimal mechanism, we get:

$$e^{-rt_x^T}p^T_{t_x^T} \;=\; e^{-rt_x^T}\frac{x}{r} - \frac{1}{r}\int_0^x e^{-rt_{\tilde{x}}^T}d\tilde{x}.$$

Integrating this by parts gives us:

$$\Pi^T \;=\; \int_0^1 e^{-rt_x^T}p^T_{t_x^T}\; f_t(x)dx \;=\; \frac{1}{r}\int_0^1 e^{-rt_x^T}(x - \frac{1-F_T(x)}{f_T(x)})f_T(x)dx.$$

By lemma 2.4, and the fact that $V_0 = 0$ in an optimal mechanism, we get:

$$e^{-(r+\lambda)t_x}p_{t_x} = e^{-(r+\lambda)t_x}\frac{x}{r} - \frac{1}{r}\int_0^x e^{-(r+\lambda)t_{\tilde{x}}}d(\tilde{x}) \;+\lambda\int_0^x e^{-(r+\lambda)t_{\tilde{x}}}\frac{\partial t_{\tilde{x}}}{\partial x}\Big(V^{t_{\tilde{x}}}_{\beta\tilde{x}} - V^{t_{\tilde{x}}}_{(\beta-1)\tilde{x}}\Big)d\tilde{x}.$$

[8] Our conjecture is that they are also sufficient but since we do not need the sufficiency in the general result we did not show it formally.

Integrating this by parts gives us:

$$\begin{aligned}\int_0^1 e^{-rt_x} e^{-\lambda t_x} p_{t_x} f(x)dx &= \frac{1}{r}\int_0^1 e^{-(r+\lambda)t_x}\Big(x - \frac{1-F(x)}{f(x)}\Big)f(x)dx \\ &\quad + \lambda \int_0^1 e^{-(r+\lambda)t_x}\frac{\partial t_x}{\partial x}(1-F(x))\Big(V_{\beta x}^{t_x} - V_{(\beta-1)x}^{t_x}\Big)dx\end{aligned}$$

Therefore the monopolist's optimization problem is:

$$\begin{aligned}\max_{\{t_x\}_{x\in[0,1]},\{\{t_x^T\}_{x\in[0,1]}\}_{T>0}} \quad & \frac{1}{r}\int_0^1 e^{-(r+\lambda)t_x}\Big(x - \frac{1-F(x)}{f(x)}\Big)f(x)dx \\ & + \lambda \int_0^1 e^{-(r+\lambda)t_x}\frac{\partial t_x}{\partial x}(1-F(x))\Big(V_{\beta x}^{t_x} - V_{(\beta-1)x}^{t_x}\Big)dx \\ & + \frac{\lambda}{r}\int_0^\infty e^{-(r+\lambda)T}\Big(\int_0^1 e^{-rt_x^T}\Big(x - \frac{1-F_T(x)}{f_T(x)}\Big)f_T(x)dx\Big)dT \\ \textit{subject to} \quad & \cdot\ t_x \textit{ is non-increasing in } x \\ & \cdot\ t_x^T \textit{ is non-increasing in } x, \quad \forall T \in [0,\infty)\end{aligned} \tag{2.4}$$

We know that $f_T(.)$ is a function of $\{t_x\}_{x\in[0,1]}$ for each realized arrival time T. Therefore, while finding the optimal mechanism, one must take this indirect effect into account. The first line of the objective function is the sum of the discounted virtual valuations corresponding to sales of the first version of the good, and the third line is the analogous of it corresponding to the second version of the good with an important distinction that the virtual valuations are now based on the effective types and their distributions for every possible arrival time T. The second line, which is always non-positive due to the monotonicity of t_x on x, can be interpreted as the cost of the inter-versional incentives for the monopolist. A buyer, while purchasing the first version of the good, considers the possibility of the arrival of the newer version of the good at the very moment. If he decides to purchase, his effective type and hence his marginal willingness to pay for the newer version will change. These inter-versional incentives is reflected at the term $(V_{\beta x}^{t_x} - V_{(\beta-1)x}^{t_x})$ appearing in the second line of above program.

Note that, when $\lambda \to 0$ the second and the third terms of objective function in (2.4) is equal to 0, and hence the problem (2.4) is equivalent to the problem (2.2). This is rather intuitive, because when λ approaches to 0 there is no upgrade, and hence we get back to the canonical model. Proposition 2.1 characterizes the optimal allocation rule of the monopolist problem.

Proposition 2.1. *The optimal solution to (2.4) consists of two cutoff values x_1, and x_2 such that:*

$$t_x = \begin{cases} 0 & x \geq x_1 \\ \infty & x < x_1 \end{cases} \qquad t_x^T = \begin{cases} 0 & x \geq x_2 \\ \infty & x < x_2 \end{cases} \quad \forall T.$$

Proof. See appendix B.3. □

To prove this result, we write an auxiliary optimization problem in which the second line of the objective function is omitted. Then we show that the solution to this auxiliary problem is also the solution to the problem 2.4. This follows from the fact that the second line of the objective function is always non-positive due to the monotonicity of t_x, and its value is maximized, i.e. equal to 0, at the optimal solution of the auxiliary problem.

The important thing to note here is that introducing product upgrades into durable good markets does not alter the main result of the canonical model, in the sense that the monopolist still does not use time to discriminate over buyers. Even though the type of buyers that are purchasing the good is different than the canonical durable good model, we still have immediate allocations for both versions of the good. The optimal allocation may incorporate different scenarios in terms of the version(s) of the goods that each type of buyer is acquiring. The distribution function $f(x)$ and the parameters β, λ, and r, are crucial to decide the optimal values of x_1, and x_2, and hence on the individual allocations of each buyer. However, here we do not precisely characterize the values of x_1 and x_2 for a given set of parameters.

The monopolist can induce this allocation rule by setting a constant price level for

both versions p_1, and p_2, since allocations are immediate.[9] Note that the price of the newer version of the good is independent from the arrival time T. and is given by $p_t^T = p_2 = \frac{x_2}{r}$ $\forall T, t$. At this price level, the buyer with the effective type x_2 is indifferent between purchasing and not purchasing the second version, so that $V_{x_2}^T = 0$. On the other hand, p_1 is the price level at which the agent of type x_1 is indifferent on his purchase decision of the first version. In particular, if $(\beta - 1)x_1 \geq x_2$, then he is indifferent between purchasing both versions of the good and purchasing only the second version. And if $(\beta - 1)x_1 < x_2$, then he is indifferent between purchasing only the first version and purchasing only the second version.

Remark 2.1. *As it is mentioned earlier the solution of the optimization problem 2.4 does not necessarily be the optimal solution of the monopolist's problem, since Lemma 2.4 is an "only if " statement. However, the optimal allocation and the price path inducing this optimal allocation is obviously incentive compatible, therefore it is also the solution that we are interested in.*

2.4 Deterministic Arrival

Now, the arrival of the newer version of the good occurs at a certain time T which is commonly known from all of the participants of the market. We will show that, in this non-stationary environment, the monopolist, depending on the values of the parameters, might use time to discriminate over buyers with heterogeneous valuations.

To show the possibility of a decreasing optimal price path, we simplify our model by assuming a binary type space for buyers with types H (High) and L (Low), where $H > L$.[10]

[9] Version-wise constant price path is not the only price path to implement the optimal allocation rule. In particular, any non-decreasing contingent price path, satisfying $p_0 = p_1$, and $p_0^T = p_2$, would also induce the optimal allocation rule. A decreasing price path would only be optimal if the optimal allocation were to occur throughout time.

[10] To omit the discount factor r that appears due to the integration of the flow utilities, say the types are h, l, and we have $H = \frac{h}{r}$, and $L = \frac{l}{r}$.

The buyers are still a continuum with a unit mass and the measure of the H-type buyers is equal to $\mu \in (0,1)$ while the measure of the L-type buyers is $1-\mu$. The flow utility acquired from the second version of the good is still β times the flow utility acquired from the first version of the good for both type of buyers.

To focus on the price path of the first version of the good, we further assume that the price path for the second version is constant,and hence any purchase of the second version occurs only at the arrival time T.[11] The monopolist commits to a price path: $\{p_t\}_{t\in[0,T]}$, where p_t is the price level of the first version at $t \in [0,T)$, and p_T is the price for the second version good. The following assumption guarantees that the utility from purchasing only the first good at $t=0$ is higher than the utility from purchasing only the second version at time T.

Assumption 2.2. $\beta e^{-rT} < 1$.

To understand the incentive of the buyers, consider an arbitrary price path $\{p_t\}_{t\in[0,T]}$ and say that a buyer finds it optimal to purchase the first version of the good at a time $t \in [0,T)$ (he may or may not purchase the second version). Then it would also be an optimal decision for this buyer to purchase the first version of the good at $t \in [0,T)$ in an environment where there is only the first version of the good with the corresponding price path: $\{p_t\}_{t\in[0,T)}$. To see this, suppose that there exists another time period $\tilde{t} \in [0,T)$ that strictly dominates purchasing at t. But for this to be correct, this buyer must purchase the second version of the good as well, otherwise we get a direct contradiction. Then by revealed preferences of type-X buyer:

$$e^{-r\tilde{t}}(X-p_{\tilde{t}}) > e^{-rt}(X-p_t)$$
$$X(e^{-rt}+(\beta-1)e^{-rT}) - e^{-rt}p_t - e^{-rT}P_T \geq X(e^{-r\tilde{t}}+(\beta-1)e^{-rT}) - e^{-r\tilde{t}}p_{\tilde{t}} - e^{-rT}P_T.$$

However the inequalities above are contradicting with each other, hence our claim is correct.

[11] We can rather think of this as a restriction so that the markets close down after T.

This observation is crucial for the next lemma, which is showing that there exists a critical time period, $t^\star < T$, such that before this $t^\star$ the purchasing decisions for the first version of the good is monotonic with respect to the buyers' type. More precisely, if L-type buyers purchase the first version of the good at a time $t < t^{star}$, then H-type buyers also purchase the first version of the good and their purchase time is not later than t. On the contrary, this monotonicity does not carry out to the purchases occurring after $t^\star$. This is because of the fact that the arrival of the second version of the good becomes closer as time goes on, and the incentive to wait for the newer version of the good becomes strengthened, and these strengthened incentives is stronger for H-type buyers if the arrival time is sufficiently close.

Lemma 2.5. *For a given price path* $\{p_t\}_{t\in[0,T]}$,

i If the L-type buyers purchase both versions of the good then the H-type buyers would also purchase both versions of the good.

ii There exists a critical time period $t^\star$, *that is defined by* $e^{-rt^\star} = \beta e^{-rT}$, *such that if the L-type buyers purchase the first version of the good at a time* $t < t^\star$, *then H-type buyers also purchase the first version of the good and their purchase time is not later than* t.[12]

Proof. See appendix B.4. □

Now we can use this partial monotonicity result given in lemma 2.5 to show that in an optimal price path the monopolist should allocate the first version of the good to H-type buyers immediately, if the assumption 2.2 is satisfied. If the monopolist is not allocating the first version of the good to H-type buyers at $t = 0$, then it must be the case that they are only purchasing the second version of the good. Then L-type buyers are either only purchasing the first version of the good after $t^\star$, or only purchasing the second version of the good, and both of these allocations are dominated when assumption 2.2 is satisfied.

[12]The existence of this $t^\star$ is guaranteed by the assumption 2.2

Lemma 2.6. *In an optimal posted price mechanism, H-type buyers purchase the first version of the good immediately at* $t = 0$

Proof. See appendix B.5 □

The optimal posted prices depends on the values of H and L. In particular, the relation between $(\beta - 1)H$ and βL is crucial. When $(\beta - 1)H \geq \beta L$ ($(\beta - 1)H \leq \beta L$) the H-type buyer owning the first version of the good gets more (less) additional utility from the second version purchase compared to a L-type buyer that does not own the first version of the good. From now on we will consider the case $(\beta - 1)H > \beta L$. The analysis of the other case follows from similar arguments.

Proposition 2.2 lists all of the possible optimal price paths that the monopolist can commit. All of the price paths listed is an optimal one for some subset of parameter values. As usual, the price path inducing the optimal allocation is not unique. However, we say that the prices are constant for the first version of the good as long as all of the purchases occurs at time $t = 0$. In this case time is not used to discriminate over buyers for the sale of the first version of the good. On the other hand, if the optimal allocations for the first version of the good occurs at different time periods for different type of buyers, then the monopolist is using time to discriminate over buyers. In this case, the price path inducing the optimal allocation is decreasing over time. The following proposition shows that for some subset of the parameter space, it is possible to get a decreasing optimal price path contrary to the case with stationary stochastic arrival.

Proposition 2.2. *Suppose* $(\beta - 1)H > \beta L$ *and the assumption 2.2 is satisfied. Then the optimal posted prices and the corresponding purchase decisions of each type of buyer is one of the following. Moreover, each policy is an optimal policy for a non-empty subset of the parameter space*

1) $p_t = H \ \forall t \in [0, T)$*, and* $p_T > (\beta - 1)H$*. Only H-type buyers purchase the first version*

of the good, and no one purchases the second version.

2) $p_t = L \ \forall t \in [0, T)$, *and* $p_T = (\beta - 1)H$. *Both type of buyers purchase the first version of the good at* $t = 0$, *and only H-type buyers purchase the second version.*

3) $p_t = \begin{cases} (1 - e^{-rT})H & \forall t \in [0, \bar{t}) \\ L & \forall t \in [\bar{t}, T) \end{cases}$ *and* $p_T = (\beta - 1)H$ *where* $\bar{t}$ *satisfies* $e^{-r\bar{t}} = e^{-rT}\frac{H}{H-L}$. *Both H-type and L-type buyers purchase the first version of the good at times* $t = 0$ *and* $t = \bar{t}$ *respectively. And only H-type buyers purchase the second version.*

4) $p_t = (1 - e^{-rT})H \ \forall t \in [0, T)$, *and* $p_T = \beta L$. *Only H-type buyers purchase the first version of the good, and both types purchase the second version.*

5) $p_t = (1 - e^{-rT})L \ \forall t \in [0, T)$, *and* $p_T = (\beta - 1)L$. *Both types purchase the first version at* $t = 0$ *and they also purchase the second version of the good.*

It is easy to calculate the corresponding profit of each policy for the monopolist. Then we can see that each of these policies is the optimal one for some values of the parameters of the model.[13] In other words, for each policy there is a subset of parameters, which are also satisfying the condition $(\beta - 1)H > \beta L$ and assumption 2.2, such that the policy is optimal. We are particularly interested on the third policy, because it displays a decreasing price path for the first version of the good. In particular, the purchases of the first version of the good occur throughout time and hence the corresponding price path implementing this allocation must be decreasing.

The non-stationary environment, resulting from a certain arrival time for product upgrades, strengthens the ability of the monopolist to sort out the buyers with lower valuations. More precisely, unless the monopolist prefers to omit the second version sales by charging a sufficiently high price p_T, the H-type buyers that are purchasing the first version of the

[13] The parameters are μ, r, T, H, L.

good at $t = 0$ has an additional option: purchasing only the second version of the good. Therefore even for the case in which the monopolist allocates only the H-type buyers at $t = 0$, H-type buyers have a positive utility, if the sales of the second version of the good is not omitted by the monopolist. This in turn introduces the possibility of allocating the first version of the good to the L-type buyers after $t = 0$ without hurting the incentives of the H-type players on their purchases of the first version of the good. [14] The way that we defined time period $\bar{t}$ given in the third policy of proposition 2.2 exploits this possibility. In particular selling the first version of the good to L-type of buyers at period $t = \bar{t}$ with a price equal to their maximum willingness to pay does not hurt the incentives of the H-type players. And $\bar{t}$ is the earliest among such time periods. The proof of proposition 2.2, given in the appendix, follows a backward analysis. We first define the optimal sales of the first version of the good for each value of p_T, and then we optimize with respect to p_T. Since the main concern of this section is to show the possibility of a decreasing optimal price path, the cumbersome details of the possible anonymous optimal sale mechanism are left to the appendix and discussed in the proof of the proposition.

2.5 Conclusion

The optimal pricing problem of a durable good monopolist is analyzed. An upgraded and a superior version of the durable good arrives and replaces the existing version of the good. The main assumption is that the sales are anonymous, so that the seller cannot condition the sale of the second version of the good on the sales of the first version. When the arrival of the upgraded version follows a stationary stochastic process, the optimal price path is shown to be constant for both versions of the good, hence all the purchases occurs immediately, i.e. the monopolist does not use the time to discriminate over buyers. On the contrary, when

[14] This is not possible when there is no product upgrade; because in the canonical model, if the optimal allocation rule allocates the good only to the H-type of buyers, then the optimal price path inducing this allocation would leave 0 utility to the H-type buyers. Hence allocating the good to the L-type of buyers at a lower price would hurt the incentives of the H-type buyers.

the arrival occurs at a commonly known certain time period, it is shown that, depending on the parameters of the model, the optimal price path might decrease over time,and hence delayed purchases might occur. Hence the time might be a useful discriminatory tool for the monopolist that is endowed with full commitment power.

For both cases the optimal selling mechanism, without the restriction of anonymity, requires the monopolist to bundle both versions of the good and to sell them only together. The corresponding allocations in this case cannot be implemented by posted prices which is anonymous by definition.

Chapter 3

Man vs. Machine: When is automation inferior to human labor?

This chapter is co-authored with Pinar Yildirim.

3.1 Introduction

The World Bank estimates that about 60% of jobs will be automated in the near future (Frey and Osborne, 2017). The reasons for this expectation are relatively straightforward. Machines, robots, and artificial intelligence technologies generally provide reliable and consistent output, and do it at a lower cost compared to their human counterparts. Not surprisingly, we are seeing increasing levels of adoption of automated technologies, and recent studies document that these technologies can substitute human labor in many consumer facing environments (Acemoglu and Restrepo, 2017).

In this study, we provide a counterargument for a manager's adoption of automation by showing that automated systems can result in less desirable outcomes compared to systems that are operated by only humans. While gains from automation are indisputable, there

are also clear reasons for why automation may not always result in the best outcome. Specifically, we are concerned with two consequences of automation: loss of employees' ability to closely monitor each other's effort and the principal's diminishing ability to detect shirking behavior.

Automation may negatively influence the output in a workplace due to a host of factors. Some of these may be behavioral reasons such as reduced morale, fear of losing employment, feeling powerless, and loss of social interactions which lead to fewer learning opportunities for the employees. In this paper, we abstract away from these behavioral explanations. Our model instead focuses on the consequences of automation from a team perspective. We consider firms where the output is created by the combined effort of agents working in teams. When employees work in teams, they get a chance to closely monitor each other's effort. Monitoring of effort, however, is a problem for most managers since continuously auditing and checking up on workers are too costly. When the employee pay incentives are set by the principal in a way that they have an incentive to monitor their team member's behavior, the principal can take advantage of this to induce both individuals to exert high effort. On the other hand, when one member of the team is replaced with automated machinery, the principal's monitoring capacity diminishes, because the human labor remains unmonitored. While the machine effort (or input) is always high and its cost is always lower than its human counterpart, in organizations with automation, to incentivize the employee to exert high effort, the principal has to offer a lot more. As a result, in some environments, automation becomes less desirable compared to all-human labor.

We develop a model based on Che and Yoo (2001) and study three possible incentive schemes which can be offered to an employee: contracts that reward workers when their peers also exert effort (joint performance evaluation or JPE in short), contracts that reward workers when they perform better than their peers (relative performance evaluation or RPE in short) and contracts that reward employees individually (independent performance evaluation, or IPE). These represent the wage contracts we observe in real life, and cover

the incentives based on both collaboration and on competition between employees in teams.

We model teams where a single member is replaced by automation and compare the output of this team to that of a team with two human employees.[1] We compare teams with and without automation and allow the principal to consider a set of payment schemes which reward an employee relative to, or jointly with his peer. While automation always leads to deterministic high effort and cost savings, human-machine teams lose their strategic component because a machine cannot retaliate against an employee that it is teaming with, and cannot impose any peer-sanctions. We argue that this lack of strategic behavior can reduce a principal's ability of benefiting from the interaction between the agents and can make human-teams more preferable over automated teams.

Our findings demonstrate three key insights. First, we show that adoption of automation to substitute human labor is not always preferable, even when automation implies consistent high effort and low cost. This is because while teams of humans have the ability to monitor each other's behavior and apply peer sanctions in case one deviates from the optimal path of play, machines cannot act in the same strategic way. We show that the human-machine teams can be more costly because the principal has to offer higher incentives to motivate the human member of the team. Second, we show that adoption of automation changes the wage contracts preferred by the managers in the marketplace. While both JPE or RPE are preferred as an incentive scheme to induce effort from human employees, adoption of automation makes JPE the less preferred incentive relative to RPE.

The rest of the paper is organized as follows. In Section 3.2 we provide a brief summary of the literature on team incentives, automation, and human-machine teams. In Section 3.3, we provide our model and lay out some key insights. In Section 3.4, we provide a summary of the optimal regime. We follow up with a discussion of extensions in Section 3.5 to the model to show robustness of the findings. Finally, in Section 3.6, we conclude with takeaways for

[1] We focus on teams of two for simplicity, however, the results can be extended to teams with higher number of members.

managers and policy makers.

3.2 Literature Review

Our study relates to the long tradition of study on employee contracts and team motivation. Earlier works in economics focused on the comparison of incentive schemes in teams. In a well-cited paper, Che and Yoo (2001) study the principal's wage setting problem when two agents are working together and the principal can only imperfectly monitor their effort. In this setting, the authors study both a static and dynamic game environment, and show that (1) in a static setting, RPE is the optimal incentive scheme to induce effort while (2) in a repeated interaction environment, JPE or IPE could also be preferred. Our results are largely in comparison to this earlier paper. Unlike the previous paper, we focus on the strategic choice of adopting automation and how this changes which incentives are preferred by the principal.

Others have also contrasted the RPE and JPE incentive schemes and pointed out to their desirable and undesirable properties as wage contracts. JPE can promote cooperation since an agent is rewarded only if his peers perform well (Holmstrom, 1982). But it also creates an environment open to free riding (Alchian and Demsetz, 1972; Holmstrom, 1982). Che and Yoo (2001) show that when agents interact repeatedly, JPE can become the preferred incentive.

Holmstrom (1982) argues that RPE helps the principal to reward an agent based on his effort rather than luck - i.e., the exogenous shocks that influence his output partly independent of his effort. Further, it is argued that in rank-order RPE incentives, agents are completely insulated from the risk of common negative shocks (Lazear and Rosen, 1981; Nalebuff and Stiglitz, 1983). Carmichael (1983) argued that RPE may dominate IPE, even in the absence of common shocks. The issues with the RPE scheme is that it is conducive to collusion (Mookherjee, 1984) and sabotage as well as discouraging cooperation (Lazear,

1989).

More recently, scholars in both economics and marketing started to focus on automation, robots, and artificial intelligence (Moriarty and Swartz, 1989; Venkatraman, 1994; Brynjolfsson and McAfee, 2012). A significant majority of these studies focused on whether new technologies and automation eliminate the jobs or create new ones. Studies support both sides of the argument. Acemoglu and Restrepo (2016) argue that automation will replace some jobs while creating others, and the rate of production can exceed that of destruction. Bessen (2016) investigated the relationship between technology and occupations and looks at occupations since 1980 to explore whether adoption of computers is related to job losses. Bessen (2016) argues that occupations using computers grow faster, even for the highly routine and mid-wage occupations. Acemoglu and Restrepo (2017), in a follow up empirical study demonstrated that automation resulted in the loss of about 670,000 manufacturing jobs between 1990 and 2007. Sachs et al. (2015) studied the impact of robotization and found that a rise in robotic productivity is more likely to lower the welfare of young workers and future generations when the saving rate is low.

In marketing, scholars have studied the impact of technology adoption and automation in consumer service environments (Speier and Venkatesh, 2002; Jayachandran et al., 2005). However, these studies do not focus on how to optimally integrate automation to teams. This is how our study contributes the literature. We demonstrate the conditions under which automation is preferable and we also demonstrate the payment incentive that is best suited to maximize the output of employees.

Finally, our paper also relates to the literature on collusion and renegotiation-proofness in repeated games. First introduced by Farrell (1983), a renegotiation-proof equilibrium, is a sub-game perfect equilibrium in which there is no continuation play that is Pareto dominated by another sub-game perfect equilibrium. Van Damme (1989) demonstrates that, in a repeated prisoners' dilemma game, it is possible to sustain every feasible and

individually rational payoff outcome by means of a renegotiation-proof equilibrium.[2] In contrast to this literature, we follow Che and Yoo (2001), and impose restrictions to rule out collusive-outcomes only on the equilibrium path. Despite this difference, we show that, the resulting optimal scheme induces the desired outcome also by means of a renegotiation-proof equilibrium as in the literature.

3.3 Model

Consider a firm which will make a decision on whether or not to initiate partial substitution of labor with automated systems. This choice makes up two possible production regimes. The first regime is without automation, and the production is purely based on labor force. The second regime, on the other hand, incorporates automation, and the production is based on labor as well as automated machinery. In order to examine this choice, we study a dynamic framework within a discrete time setting, and set δ as the common time discounting rate. In what follows, we will describe these production regimes in more detail.

3.3.1 No-Automation Regime

The framework is borrowed from Che and Yoo (2001). Suppose a firm hires two identical agents to perform a task in each period.[3] Each agent makes a binary effort choice in each period. The agent can either "work" by setting his effort level to $e = 1$, or can "shirk" by setting it to $e = 0$. There is a discomfort, or cost, of high level effort, $e = 1$, which is denoted by $c > 0$. The cost of shirking is normalized to 0.

The agents' effort choices are not directly observed by the principal, instead she observes two distinct informative signals about the effort exerted by each agent. The signals, similar to the effort choices, are binary and can be either "favorable" or "unfavorable",

[2] See also Farrell and Maskin (1989).

[3] We will refer to the human employees as agents or workers interchangeably throughout the document.

i.e., $s = 1$, or $s = 0$. The signal of an agent is independent of the effort choice of the other, however, there is an underlying common component influencing the realizations of both signals. This common component may, for instance, indicate the state of the economy or a particular industry which influences both workers in a similar way. More precisely, there is an aggregate shock which can either be "good" or "bad" with probabilities σ and $1 - \sigma$ respectively, where $\sigma < 1$. If the shock is good, then the signals of both agents are favorable regardless of their effort choice. If the shock is bad, then the signal depends on the effort choice of a worker. In particular, if he chooses to put effort ($e = 1$), then his performance signal will be favorable with probability p_1, and if he chooses to shirk ($e = 0$) then his performance signal will be favorable with probability p_0, where $1 > p_1 > p_0 > 0$. Put differently, when the aggregate shock is bad, the probability that an agent will receive a more favorable evaluation is higher if he chooses to put in effort. The joint distribution of the signals is summarized in the Table 3.1.

		Signal Pairs			
		(1,1)	(1,0)	(0,1)	(0,0)
Effort Pairs	(1,1)	$\sigma + (1-\sigma)p_1^2$	$(1-\sigma)p_1(1-p_1)$	$(1-\sigma)(1-p_1)p_1$	$(1-\sigma)(1-p_1)^2$
	(1,0)	$\sigma + (1-\sigma)p_1p_0$	$(1-\sigma)p_1(1-p_0)$	$(1-\sigma)(1-p_1)p_0$	$(1-\sigma)(1-p_1)(1-p_0)$
	(0,1)	$\sigma + (1-\sigma)p_0p_1$	$(1-\sigma)p_0(1-p_1)$	$(1-\sigma)(1-p_0)p_1$	$(1-\sigma)(1-p_0)(1-p_1)$
	(0,0)	$\sigma + (1-\sigma)p_0^2$	$(1-\sigma)p_0(1-p_0)$	$(1-\sigma)(1-p_0)p_0$	$(1-\sigma)(1-p_0)^2$

Table 3.1: The joint distribution of the signals conditional on the effort profile of the agents.

The workers observe the effort choice of each other as they work closely. This introduces a capacity of mutual monitoring that the principal can exploit in order to mitigate the agency problem. We will demonstrate that this mutual monitoring is a key characteristic in a principal's preference for all-human teams, particularly when her capacity to detect shirking is lower.

3.3.2 Automation Regime:

In this regime, we consider the case of partial substitution of labor with automation. We assume that the principal hires one employee and replaces the other with a machine. We also assume that the machine introduces some efficiency to the principal: it operates at a capacity that is equivalent to the high level effort of the agent but at a lower operating cost, which is assumed to be αc, where $\alpha \in (0, 1)$. Thus, substituting one of the employees with machinery has two direct benefits: reduction in operating costs and the scope of the moral hazard problem since there is only one agent now with an unobservable effort choice. Nevertheless, this replacement exterminates the interaction among the agents; hence diminishes the potential benefit that the principal may receive due to the agents' mutual monitoring capacity. The latter accounts for the indirect cost of partial substitution of human labor with machines. The trade-off between the direct and indirect effects comprise the core of the discussion of this paper.

Similar to the no-automation regime, the employee's effort choice is not directly observed by the principal. She observes a binary performance signal, which has the same distribution as before. In this case there is only one signal since there is only one agent.

3.4 Optimal Incentive Contract

Throughout this section, we maintain the assumption that the high level effort (i.e., work) is sufficiently valuable for the principal such that her central focus is to find an optimal incentive scheme inducing high effort, $e = 1$, in every period. We restrict our attention to constant incentive schemes where the payment that an agent receives is time-invariant and depends on the performance measure(s) realized within the current period. It is also assumed that the agents have limited liability and the payments are non-negative after each possible contingency.

We first analyze the optimal production in all-human (i.e., human-human) and human-machine teams individually. Next we compare the expected cost of performing the task in both cases to determine the optimal scheme.

3.4.1 Optimal Incentive Scheme Under Automation

We start the analysis with investigating the optimal incentive for a human-machine team. With partial substitution of labor, the principal's problem is to set the incentive scheme to induce high effort from the employee. An incentive scheme, in this case, is a two dimensional vector specifying the amount of pay the agent receives after each possible performance measure. More precisely, $\mathbf{w} = (w_1, w_0)$ is the wage scheme where w_1 and w_0 represent the amount paid to agent if his performance signals are $s = 1$, and $s = 0$, respectively. In each period, conditional on the wage scheme set by the principal, the agent determines the level of effort he puts in. To induce high effort from the agent, the principal has to take the following incentive constraint into account while setting the wages:

$$\underbrace{(\sigma + (1-\sigma)p_1)w_1 + (1-\sigma)(1-p_1)w_0 - c}_{\text{Utility from working}} \geq \underbrace{(\sigma + (1-\sigma)p_0)w_1 + (1-\sigma)(1-p_0)w_0}_{\text{Utility from shirking}}.$$

The left hand side of the inequality is the agent's utility from working, and the right side is his utility from shirking. After some simplification, this constraint boils down to the following expression:

$$(1-\sigma)(p_1 - p_0)(w_1 - w_0) \geq c. \tag{$\mathcal{IC}$}$$

Then the principal's problem is the following:

$$\min_{w_1, w_0} \quad [\sigma + (1-\sigma)p_1]w_1 + (1-\sigma)(1-p_1)w_0 \quad \text{s.t. } \mathcal{IC} \tag{$\mathcal{P}$}$$

It is straightforward to see that under the optimal incentive scheme, the agent receives

a positive payment only after a favorable performance signal, therefore $w_0 = 0$. Moreover, the incentive constraint is binding, therefore we can derive the optimal value of w_1. The following proposition provides the optimal incentive scheme when the task is performed by a human-machine team.

Proposition 3.1. *The optimal incentive scheme that induces the agent to work in every period is given by* $\boldsymbol{w} = (\hat{c}, 0)$*, where* $\hat{c} = \frac{c}{(1-\sigma)(p_1-p_0)}$.

The expression for the optimal value of wage w_1 conveys that when agents work in a more favorable economy (as σ increases) or when the returns from putting a high effort under a bad shock decreases (as the difference between p_1 and p_0 gets smaller) it becomes harder for the principle to detect an agent when he shirks. Therefore, the pay needed to incentivize him to exert effort becomes larger.

3.4.2 Optimal Incentive Scheme Under No-Automation

Our discussion in this section follows the model and the derivations in Che and Yoo (2001), however, complementing the earlier work, we provide the complete description of the optimal incentive schemes here. This extends the analysis provided in the earlier work.

With all-human teams, there are two agents employed and they observe the effort choice of each other since they work closely. The principal has to provide incentives for both to motivate them to work. To this end, she can condition the payment of one agent on the realized performance of the other. More precisely, she can set the payment scheme as a four dimensional vector $\mathbf{w} = (w_{11}, w_{10}, w_{01}, w_{00})$ where $w_{s_i s_j}$ is the amount paid to agent i when his signal is s_i and his teammate's signal is s_j.[4] The structure of wage scheme, $\mathbf{w}$, plays a crucial role in shaping the inherent features of the team interaction. This in turn allows the principal to manipulate this interaction in his benefit and to mitigate the overall

[4]Note that this formulation restricts the wage structure to be symmetric for the agents. This restriction, however, is innocuous given that the agents are identical.

agency problem.

Similar to Che and Yoo (2001), we characterize the wage schemes based on how an agent's payment is determined depending on the other agent's performance. An agent can be rewarded or punished when his teammate performs well. More specifically, following Che and Yoo (2001), if a wage scheme rewards an agent for a good performance of his teammate, we refer to this scheme as the "joint performance evaluation" (JPE, in short). JPE satisfies

$$(w_{11}, w_{01}) >> (w_{10}, w_{00}),$$

implying that $w_{11} \geq w_{10}$, $w_{01} \geq w_{00}$, and at least one of the constraints has to be strict. So an agent is paid more when his partner has a favorable evaluation.

On the contrary, if a wage scheme punishes an agent for the good performance of his teammate, we refer to it as "relative performance evaluation" (RPE, in short). RPE satisfies

$$(w_{10}, w_{00}) >> (w_{11}, w_{01}).$$

High level effort is sufficiently valuable for the principal. Thus, she wants to choose an incentive scheme that induces effort pair $(e_1, e_2) = (1, 1)$ in every period as a sub-game perfect equilibrium outcome. Given such an incentive scheme, however, there may exist multiple sub-game perfect equilibria with different outcomes. In such a circumstance, the game may proceed with an outcome other than the one the principal prefers. In this respect, the principal needs to choose an incentive scheme so that the repeated play of (work, work) is the best outcome, that can be supported by a sub-game perfect equilibrium, in terms of maximizing the agents' total payoffs. This restriction allows us to draw a unique prediction regarding the outcome of the agents' repeated interaction. Moreover, it also allows us to address and eliminate the possibility of a collusion between the agents. This restriction rules out such a possibility at any period in time on the path of play.[5] As in Che and

[5]This differs from the standard formulation of collusion and renegotiation in the literature, see for instance

Yoo (2001), we call such an equilibrium a "team equilibrium" throughout the rest of the analysis.

Based on what we presented until now, for a given wage scheme $\mathbf{w}$, and the effort profile (e_i, e_j), the expected payment that agent i receives is denoted by $\pi_i(e_i, e_j, \mathbf{w})$:

$$\begin{aligned}\pi_i(e_i, e_j, \mathbf{w}) =& [\sigma + (1-\sigma)p_{e_i}p_{e_j}]w_{11} \\ &+(1-\sigma)[p_{e_i}(1-p_{e_j})w_{10} + (1-p_{e_i})p_{e_j}w_{01} + (1-p_{e_i})(1-p_{e_j})w_{00}]\end{aligned}$$

3.4.3 Static Environment

We first analyze the problem considering a static, one period version of the game. In the absence of repeated interaction between the agents, the principal cannot fully benefit from their mutual monitoring capacity. This stems from the fact that the agents cannot use peer sanctions against each other as the team interaction takes place only once.

First, note that, for a given wage scheme, $\mathbf{w}$, the effort profile $(e_1, e_2) = (1, 1)$ is a Nash equilibrium if and only if the payoff from working exceeds the payoff from shirking:

$$\pi(1, 1, \mathbf{w}) - c \geq \pi(0, 1, \mathbf{w}) \tag{$\mathcal{IC}_\mathcal{S}$}$$

This incentive constraint constitutes a necessary condition for the team equilibrium. In this respect, the following problem comprises a relaxed version of the principal's problem in the static setting.

$$\min_{\mathbf{w}} \ \pi(1, 1, \mathbf{w}) \ \text{ s.t. } \mathcal{IC}_\mathcal{S}. \tag{$\mathcal{P}_\mathcal{S}$}$$

The principal wants to find an optimal incentive scheme that supports both agents working (i.e., (work, work)) as a team equilibrium outcome. Under the solution to the above problem,

Van Damme (1989). The effects of this difference on the optimal incentive scheme will be discussed later on in the paper.

however, (work, work) is only guaranteed to be a Nash equilibrium. Therefore the solution of $\mathcal{P}_{\mathcal{S}}$ does not necessarily coincide with the solution of the principal's problem. Yet, it turns out that (work, work) is the unique equilibrium under the incentive scheme $\mathbf{w}^s$ that solves $\mathcal{P}_{\mathcal{S}}$. Therefore, $\mathbf{w}^s$ is the optimal incentive scheme.

As depicted in the following statement, the optimal incentive scheme displays an extreme form of RPE. More precisely, the agent receives a positive payment only when his performance measure is more favorable compared to the other agent. In all the other remaining contingencies, the amount of payment he receives is zero.

Proposition 3.2. *The optimal incentive scheme in the static setting displays an extreme form of RPE, and is given by* $\boldsymbol{w}^s = (0, w^s_{10}, 0, 0)$*, where* $w^s_{10} = \frac{\hat{c}}{(1-p_1)}$.

Proof. See appendix C.1. □

Proposition 3.2 states that when agents' performance measures have a common component, a relatively more favorable signal is a stronger indication of an agent's high level effort. Therefore, paying the agents only in such a circumstance is the cheapest way to provide incentives for inducing high level effort. When the agents' interaction takes place only once, the optimal incentive scheme requires to pay each agent only when he receives a relatively better performance measure. This effectively puts the agents into a competition, in which the amount of payment is set so that each agent is indifferent in his effort choice, conditional on the other agent exerting high level effort. When we move into the dynamic setting, we will see how this incentive scheme is vulnerable to collusion.

3.4.4 Dynamic Environment

We next study the dynamic setting, where the agents repeatedly interact with each other over an infinite time horizon. Repeated interaction introduces the threat of peer sanction that was not present in the static version of the game. It also implies that the possibility

of collusion is now a more important concern for the principal. There is an opportunity of mutual monitoring as the agents observe the actions of each other. This allows them to sustain outcomes favoring mutual benefit by means of using credible threats. We will see that, in a dynamic environment, RPE incentive schemes are more vulnerable to collusion in comparison to the static setting, and JPE may become the optimal way of incentive provision depending on the parameter range.

A strategy for an agent is a mapping from the set of histories to the set of actions at every period. A relevant history includes the realized performance measures as well as the effort choices of the agents in the earlier periods. The principal is interested in choosing an optimal incentive scheme inducing high level of effort in every period as a team equilibrium outcome. The following lemma asserts that, in an optimal incentive scheme, an agent receives a positive payment only if he receives a favorable performance signal. In other words, the optimal values of w_{01}, and w_{00} is equal to 0. As a result, the principal's problem boils down to finding out the values of w_{11}, and w_{10}.

Lemma 3.1. *In a human-human team, an optimal incentive scheme always satisfies* $w_{10} = w_{00} = 0$.

Proof. See appendix C.5. □

Paying a positive amount to an agent when he has a poor performance evaluation makes it more difficult to induce high effort. The lemma presents this intuition. Thanks to this result, we now know that an optimal incentive scheme has to be either JPE, RPE, or IPE, depending on how the values of w_{11}, and w_{10} compare to each other. [6] If $w_{11} > w_{10}$, then $\mathbf{w}$ is a JPE; and if $w_{11} < w_{10}$, then $\mathbf{w}$ is a RPE. In the case when $w_{11} = w_{10}$, the incentive scheme is an IPE.

In order to describe the optimal values of w_{11} and w_{10}, first, consider the following

[6] Note that these three categories do not cover the whole space of possible incentive schemes. For instance, an incentive scheme satisfying $w_{11} > w_{10}$, and $w_{01} < w_{00}$ is neither JPE nor RPE nor IPE.

inequality which comprises a necessary condition.

$$\pi(1,1,\mathbf{w}) - c \geq (1-\delta)\pi(0,1,\mathbf{w}) + \delta \min\{\pi(0,0,\mathbf{w}), \pi(0,1,\mathbf{w})\} \qquad (\mathcal{IC}_{\mathcal{D}})$$

LHS of the inequality is the resulting payoff of each agent from repeated joint work. If this value is lower than the RHS of the inequality, then the agents would receive a higher utility from any deviation. As a result, it would be impossible to sustain repeated joint work as a sub-game perfect equilibrium outcome. This constraint does not constitute a sufficient condition for a sub-game perfect equilibrium since the implied continuation play following a deviation does not necessarily have to be self-enforcing. Therefore it cannot be a sufficient condition for a team equilibrium either. In this regard, the following problem, which takes only this constraint into account, is just a relaxed version of the principal's problem.

$$\min_{w_{11},w_{10}} \ \pi(1,1,\mathbf{w}) \ \text{ s.t. } \mathcal{IC}_{\mathcal{D}}. \qquad (\mathcal{P}_{\mathcal{D}})$$

Solving $\mathcal{P}_{\mathcal{D}}$ is an important step towards the characterization of the optimal incentive scheme in an all-human team. The following lemma demonstrates the solution of the problem $\mathcal{P}_{\mathcal{D}}$.

Lemma 3.2. *Define a level of discount factor* $\hat{\delta}(\sigma) = \frac{\sigma}{(1-\sigma)p_1 p_0}$. *Then*
(i) If $\delta > \hat{\delta}(\sigma)$, *then a JPE scheme* $\boldsymbol{w}^j = (w_{11}^j, 0, 0, 0)$ *with* $w_{11}^j = \frac{\hat{c}}{p_1 + \delta p_0}$ *solves* $\mathcal{P}_{\mathcal{D}}$.
(ii) If $\delta \leq \hat{\delta}(\sigma)$, *then the solution to the static problem* ($\boldsymbol{w}^s$) *also solves the problem* $\mathcal{P}_D$ *optimally. In this case, the solution is an RPE scheme.*

Proof. See appendix C.2. □

Lemma 3.2 states that the solution of the problem $\mathcal{P}_{\mathcal{D}}$ displays either an extreme form of JPE or an extreme form of RPE. For this solution to coincide with the optimal solution of the principal's problem, it must induce both agents to choose high effort in every period, as a team equilibrium outcome.

The JPE incentive scheme, $\mathbf{w}^j$ possesses all the required properties. More precisely,

when the pay scheme is $\mathbf{w}^j$, the repeated joint work is a sub-game perfect equilibrium outcome. Furthermore, $\mathbf{w}^j$ does not create any collusive outcome. To see this, first note that, when the pay scheme is $\mathbf{w}^j$, the interaction between the agents is equivalent to a Prisoner's Dilemma game in which (shirk, shirk) is the unique stage game Nash equilibrium. Agents' payoff increases with the performance of their teammates, i.e, $\pi(0,1,\mathbf{w}^j) > \pi(0,0,\mathbf{w}^j)$, and $\mathcal{IC}_{\mathcal{D}}$ boils down to:

$$\pi(1,1,\mathbf{w}^j) - c \geq \delta\pi(0,1,\mathbf{w}^j) + (1-\delta)\pi(0,0,\mathbf{w}^j).$$

Therefore, initiating (shirk, shirk)$^\infty$ can deter unilateral deviations from the desired outcome, (work, work)$^\infty$. Such a punishment is self-enforcing, since repetition of (shirk, shirk) is self-enforcing. As a result, (work, work)$^\infty$ can be sustained as a sub-game perfect equilibrium outcome. Moreover, the effort pair (work, work) maximizes the agents' total payoff. Precisely,

$$\underbrace{2\pi(1,1,\mathbf{w}^j) - 2c}_{\text{Total payoff from (work, work)}} \geq \underbrace{\pi(1,0,\mathbf{w}^j) - c + \pi(0,1,\mathbf{w}^j)}_{\text{Total payoff from (work, shirk)}},$$

and

$$\underbrace{2\pi(1,1,\mathbf{w}^j) - 2c}_{\text{Total payoff from (work, work)}} \geq \underbrace{2\pi(0,0,\mathbf{w}^j)}_{\text{Total payoff from (work, work)}}.$$

Thus, $\mathbf{w}^j$ induces (work, work)$^\infty$ as a team equilibrium outcome, and hence is the optimal incentive scheme for principal, whenever it is a solution for the problem $\mathcal{P}_{\mathcal{D}}$. Under this incentive scheme, what happens as a matter of course is the following. The principal benefits from the agents' mutual monitoring capacity by asking them to impose peer sanctions to each other in case they observe any shirking behavior. The agents have enough incentive to do this since they do not want their teammates to shirk as it hurts them as well.

Such a scheme becomes optimal for sufficiently large values of the discount factor. This is rather intuitive. When the agents are more patient, the repeated play of (shirk, shirk) becomes a more severe punishment. The agents are more willing to exert high level effort in

order to avoid the punishments imposed by their teammates, and hence the amount of the payment, w_{11}^j, that is needed to convince them to work is lower. As a result, it is natural to have a JPE scheme as an optimal pay scheme when the agents are sufficiently patient.

The fact, that a JPE incentive scheme may be optimal, comprises the main divergence between the static and the dynamic settings. When agents repeatedly interact with each other over an infinite time horizon, the principal can convince them to monitor and discipline each other by using peer sanctions through the help of JPE.

In contrast to the JPE scheme $\mathbf{w}^j$, the RPE scheme $\mathbf{w}^s$ does not possess all the desired properties. In particular, it is susceptible to collusion hence cannot be the optimal pay scheme of the dynamic setting. It supports (work, work)$^\infty$ as a sub-game perfect equilibrium since (work, work) is the unique stage game Nash equilibrium. However, the agents can sustain better payoff outcomes by coordinating over the strategy profiles that include unilateral or bilateral deviations from (work, work) on the path of play. For instance, consider a strategy profile with an outcome in which the agents alternate between (work, shirk), and (shirk, work). The corresponding expected payoffs of the agents from this play is equal to:

$$\frac{1}{1+\delta}(\pi(1,0,\mathbf{w}^s)-c)+\frac{\delta}{1+\delta}\pi(0,1,\mathbf{w}^s),$$

and

$$\frac{\delta}{1+\delta}(\pi(1,0,\mathbf{w}^s)-c)+\frac{1}{1+\delta}\pi(0,1,\mathbf{w}^s)$$

for the first and the second agents respectively. Note that each of these are larger than the payoff $\pi(1,1,\mathbf{w}^s)-c$. Moreover, this outcome can be supported in a sub-game perfect equilibrium in which the deviations trigger the repeated play of (work, work), which is self-enforcing. Conditional on this punishment, even the agent shirking in the current period, who has a stronger incentive to deviate, does not want to deviate. More precisely:

$$\frac{\delta}{1+\delta}(\pi(1,0,\mathbf{w}^s)-c)+\frac{1}{1+\delta}\pi(0,1,\mathbf{w}^s)\geq\pi(1,1,\mathbf{w}^s)-c.$$

Consequently, under $\mathbf{w}^s$, collusion is a real concern, hence it cannot be an optimal incentive scheme. The principal has to account for such possibilities of collusion and preclude them all.

The solution to the problem $\mathcal{P}_D$ characterizes the optimal incentive scheme for the case in which the principal is able to select the equilibrium in her interest. Yet, she can not observe the agents' effort choices, hence she has to incentivize them properly in order to make sure that they do not collude. This concern becomes particularly relevant when $\mathbf{w}^s$ is the solution, and overall $\mathcal{P}_D$ provides only a partial characterization.

Then, what is the optimal incentive scheme when $\mathbf{w}^s$ is the solution of the problem $\mathcal{P}_{\mathcal{D}}$? In order to answer this question, we focus on the the principal's problem by restricting her choice into the set of incentive schemes satisfying $w_{11} \leq w_{10}$. We already know her optimal choice under the restriction $w_{11} \geq w_{10}$.[7] By this approach, we reach a characterization of the optimal incentive scheme in two different subsets, the union of which covers the entire space of incentive schemes. After completing this step, finding the global optimal is just a matter of a simple comparison.

First, note that the constraint $\mathcal{IC}_{\mathcal{D}}$ is still a necessary condition, and is equivalent to $\mathcal{IC}_{\mathcal{S}}$ when $w_{11} \leq w_{10}$. On top of this constraint, the principal has to account for additional one(s) in grounds of preventing collusion. This is a difficult task as there are plenty of potential outcomes that the agents can collude upon. We, however, focus on a specific outcome, "alternated unilateral shirking", in which the agents alternate between (shirk, work) and (work, shirk) over time. We define some necessary conditions to prevent this outcome from being a possible collusion threat. This constraint then will turn out to be also sufficient in preventing all the collusive outcomes.

In alternated unilateral shirking, each agent provides a favor to his teammate on exchange of receiving another in the next period. Precisely, shirking in a period while his

[7] Among the set of incentive schemes satisfying, $w_{11} \geq w_{10}$, the optimal one is either $\mathbf{w}^j$, or $\mathbf{w}^I = (\hat{c}, \hat{c}, 0, 0)$. The latter is the optimal IPE, in which the agents' compensations are independent from each other, hence it is trivially collusion-proof.

teammate is working hurts the agent and favors his teammate. The principal has to choose a payment scheme such that either this outcome cannot be supported as a sub-game perfect equilibrium outcome, or the corresponding total utility of the agents is not any better than the one resulting from repeated joint work. The latter translates into the following condition:

$$\pi(1,0,\mathbf{w}) - c + \pi(0,1,\mathbf{w}) \leq 2(\pi(1,1,\mathbf{w}) - c). \tag{$\mathcal{IC}''_{\mathcal{R}}$}$$

On the other hand, to prevent this from being a sub-game perfect equilibrium outcome, the pay scheme has to be arranged in a way that at least one of the agents prefers to deviate from the corresponding path of play. We assume that the deviations trigger a continuation play of (work, work)$^{\infty}$, which is self-enforcing as inducing it in a self-enforcing manner is the principal's main objective.[8] Then the conditions we are looking for translate into the following constraints for the shirking and the working agents respectively:

$$\frac{1}{1+\delta}\pi(0,1,\mathbf{w}) + \frac{\delta}{1+\delta}(\pi(1,0,\mathbf{w}) - c) < \pi(1,1,\mathbf{w}) - c. \tag{$\mathcal{IC}_{\mathcal{R}}$}$$

$$\frac{1}{1+\delta}(\pi(1,0,\mathbf{w}) - c) + \frac{\delta}{1+\delta}\pi(0,1,\mathbf{w}) < (1-\delta)\pi(0,0,\mathbf{w}) + \delta(\pi(1,1,\mathbf{w}) - c) \tag{$\mathcal{IC}'_{\mathcal{R}}$}$$

Overall, the incentive scheme has to satisfy at least one of the the constraints $\mathcal{IC}_{\mathcal{R}}$, $\mathcal{IC}'_{\mathcal{R}}$, and $\mathcal{IC}''_{\mathcal{R}}$. To this respect, we can define a comprehensive constraint that encompasses all of these.

Definition 3.1. *An incentive scheme satisfies is said to satisfy $\widehat{\mathcal{IC}_{\mathcal{R}}}$ if and only if it satisfies at least one of $\mathcal{IC}_{\mathcal{R}}$, $\mathcal{IC}'_{\mathcal{R}}$, and $\mathcal{IC}''_{\mathcal{R}}$.*

Then we can define a relaxed version of the principal's problem, when her choice is

[8] Note that, we want alternated unilateral shirk to not be a sub-game perfect equilibrium outcome. But we consider just a specific strategy profile in which the deviations from this outcome lead to repeated play of (work,work). Precluding this strategy profile from being an SPE, does not necessarily be sufficient for ruling out all other possible SPE that in which the corresponding outcome is alternated unilateral shirking. Nevertheless, it turns out to be sufficient.

restricted to satisfy $w_{11} \leq w_{10}$, as follows:

$$\min_{w_{11}, w_{10}} \quad \pi(1, 1, \mathbf{w}) \quad \text{s.t.} \quad \mathcal{IC}_{\mathcal{S}},\ \widehat{\mathcal{IC}_{\mathcal{R}}},\ w_{11} \leq w_{10}. \tag{$\mathcal{P}_{\mathcal{R}}$}$$

This is just a relaxed version of the principal's problem, as it only takes a specific form of collusion into account. The next result provides a characterization of the solution to this problem. Moreover, it asserts that the constraints that are taken into account in the problem $\mathcal{P}_{\mathcal{R}}$ are sufficient to preclude all the collusive outcomes when the incentive scheme satisfies $w_{11} \leq w_{10}$.

Lemma 3.3. *The solution to the problem $\mathcal{P}_{\mathcal{R}}$ must be one of the following.*

- $\boldsymbol{w}^r = (0, w_{10}^r, 0, 0)$, *where* $w_{10}^r = \frac{\hat{c}}{1-(1+\delta)p_1}$.[9]
- $\boldsymbol{w}^I = (\hat{c}, \hat{c}, 0, 0)$.

Moreover, when the incentive scheme is restricted to satisfy $w_{11} \leq w_{10}$, the solution of the problem $\mathcal{P}_{\mathcal{R}}$ coincides with the solution of the principal's problem.[10]

Proof. See appendix C.3. □

Remarkably, solving the problem $\mathcal{P}_{\mathcal{R}}$ gives us the optimal incentive scheme among the ones satisfying $w_{11} \leq w_{10}$. In other words, the resulting solution of $\mathcal{P}_{\mathcal{R}}$ is always collusion-proof. $\mathcal{P}_{\mathcal{R}}$ accounts for a specific form of collusion, prevention of which turns out to be sufficient for precluding all the others. Collusion-proofness is an obvious property of $\mathbf{w}^I$ as the agents are getting paid depending on their own performance measures only. In $\mathbf{w}^r$,

[9] Note that, when $(1+\delta)p_1 \geq 1$, w_{10}^r is not well defined. In such a circumstance, $\mathbf{w}^I$ is the solution of the problem $\mathcal{P}_{\mathcal{R}}$.

[10] When $\mathbf{w}^r$ solves the problem $\mathcal{P}_{\mathcal{R}}$, the constraint $\mathcal{IC}_{\mathcal{R}}$ is binding. However, notice that $\mathcal{IC}_{\mathcal{R}}$ is based on a strict inequality. As a result, the actual value of w_{10}^r is infinitesimally larger than $\frac{\hat{c}}{1-(1+\delta)p_1}$. To this regard, one can either discretize the choice set and let principal to choose the minimum value larger than $\frac{\hat{c}}{1-(1+\delta)p_1}$ for w_{10}^r. In case we allow for arbitrarily fine choices, the principal's problem does not always have a solution, in the sense that there does not exist an incentive scheme maximizing her objective. However for a given $\epsilon > 0$, we can find an incentive scheme in which the principal can get a payoff in the ϵ-neighborhood of his value.

which is the other possible solution of $\mathcal{P}_{\mathcal{R}}$, we have an RPE scheme. The payment the agents receive after receiving a relatively better performance signal, w_{10}^r, is set sufficiently large so that the agents do not want to collude.The optimal RPE of the static setting, $\mathbf{w}^s$, on the other hand, did not have this property as the agent could always collude by following alternated unilateral shirking. In this regard, the difference between the payments w_{10}^r and w_{10}^s, can be thought as the cost of precluding collusion in RPE schemes for the principal.

Under $\mathbf{w}^r$, the corresponding team equilibrium features repetition of (work, work) on and off the equilibrium path. This comprises a sub-game perfect equilibrium, because (work, work) is an equilibrium of the stage game. The agents always prefer to exert high effort and compete with each other in order to receive a better performance signal.

Based on these discussions, we now know that the optimal incentive scheme has to be either $\mathbf{w}^j$, or $\mathbf{w}^r$, or $\mathbf{w}^I$. The principal either puts agents into a competition via a collusion-proof RPE pay scheme, and require a relatively better performance signal for a positive payment. Or, she may choose a JPE scheme, and take advantage of the mutual monitoring capacity between the agents by asking them to impose peer sanctions in case they observe a shirking behavior. Finally, she can also choose to incentivize the agents separately by selecting the optimal IPE scheme, which would exterminate all the possible within-team interaction. The following theorem illustrates how the parameter space is partitioned into three sub-spaces depending on the corresponding optimal incentive schemes.

Proposition 3.3. *There are two critical values of the discount factor* $\delta_j = \frac{\sigma(1-p_1)}{(\sigma+(1-\sigma)p_1)p_0}$, *and* $\delta_r = \frac{\sigma(1-p_1)}{(\sigma+(1-\sigma)p_1)p_1}$, *such that for a human-human production team the optimal incentive scheme is given by:*

$$\boldsymbol{w} = \begin{cases} \boldsymbol{w}^j & \text{if } \delta > \delta_j \\ \boldsymbol{w}^I & \text{if } \delta \in (\delta_r, \delta_j] \\ \boldsymbol{w}^r & \text{if } \delta \leq \delta_r. \end{cases}$$

As all the incentive schemes induces the desired path of play as a team equilibrium out-

come, the statement of the above proposition immediately follows from a simple comparison between the expected costs of the incentive schemes $\mathbf{w}^j$, $\mathbf{w}^I$, and $\mathbf{w}^r$. By comparing this result with lemma 3.2, one can see how the agents' ability of collusion alters the principal's optimal decision. When the agents cannot collude, the principal can select the equilibrium in her best interest. As a result, the optimal incentive scheme is either $\mathbf{w}^j$, or $\mathbf{w}^s$ as it is depicted in Lemma 3.2. When the agents can collude, however, the cost of motivating agents with an RPE scheme becomes larger, and $\mathbf{w}^s$ cannot be optimal anymore. This expands the region in which the $\mathbf{w}^j$ is optimal as it can also be seen from the fact that $\delta_j < \hat{\delta}$. The optimal collusion-proof RPE scheme, $\mathbf{w}^j$, is sufficiently costly for the principal, so that for some set of parameters using the IPE scheme $\mathbf{w}^I$ becomes optimal.

As δ gets large, it becomes easier to motivate agents with a JPE scheme. This stems from the reason that the peer sanctions that the agents can impose each other become more deterrent so that the amount of payment that is needed to motivate agents to work is smaller. For smaller values of δ, on the other hand, we are getting closer to the static setting as the value of future diminishes. The possibility of collusion becomes a less important concern, and the gap between the optimal RPE schemes of the static and the dynamic settings is reduced. As a result, $\mathbf{w}^r$ is optimal.

3.4.5 Collusion and Renegotiation-proofness

Agents' ability to collude leads the principal to look for collusion-proof incentive schemes for all-human teams. In defining collusion-proofness, we follow Che and Yoo (2001), and require incentive schemes to not to give rise to a sub-game perfect equilibrium outcome that brings a higher total payoff to agents in comparison to the repeated play of joint work. In other words, throughout the interaction that they involve, the agents do not want to renegotiate and coordinate over some other equilibrium outcome. This is a bit different from the conventional definition of collusion-proofness (or renegotiation-proofness) that exists in the literature, for instance Farrell (1983), and Van Damme (1989). Our

definition restricts the strategy profile only on the equilibrium path, while the conventional definition puts restrictions both on and off the path of play. Despite this difference, under the optimal incentive schemes we propose, there exist other outcome equivalent equilibria satisfying the conventional renegotiation-proofness criteria.

We know that, the optimal scheme in an all-human team has to be one of $\mathbf{w}^j$, $\mathbf{w}^r$, or $\mathbf{w}^I$. For $\mathbf{w}^r$, repetition of the unique stage Nash equilibrium, (work,work), comprises the unique sub-game perfect equilibrium. This strategy profile already satisfies the renegotiation-proofness criteria. When the incentive scheme is $\mathbf{w}^I$, the agents' compensations are independent from the effort choice of each other. In this respect, collusion is not a relevant concern, and hence the renegotiation-proofness criteria is automatically satisfied.

When it comes to the JPE scheme $\mathbf{w}^j$, the suggested strategy profile is not renegotiation-proof as it requires to punish all the deviations with a repeated play of (shirk, shirk). The team interaction is equivalent to a prisoner's dilemma game in which exerting the high level effort is the cooperative action. Thus, action pair (shirk, shirk) is Pareto dominated, and having it, even off the path of play, violates the conventional renegotiation-proofness criteria that has been proposed in the literature. Nevertheless, by following the logic suggested by Van Damme (1989), one can see that the outcome (work,work)$^\infty$ can also be supported by means of another equilibrium satisfying this criteria. Precisely, consider the strategy profile in which the agents start with exerting high level effort and continue to do so as long as there is no deviation; if the agent i deviates at time t, then starting from period $t+1$, the other agent shirks until player i works. As soon as the agent i works, the play reverts back and the agents continue to exert high level effort afterwards. Given this strategy profile, conditional on agent i deviating at time t, he prefers to exert high effort level at time $t+1$ so that the game reverts back to repeated joint work at $t+2$. This stems from the fact that:

$$(1-\delta)(\pi(1,0,\mathbf{w}^j)-c)+\delta\pi(1,1,\mathbf{w}^j-c)\geq\pi(0,0,\mathbf{w}^j).$$

Also, the agents do not want to deviate from exerting high level effort, since:

$$\pi(1,1,\mathbf{w}^j - c) \geq \ (1-\delta)\pi(0,1,\mathbf{w}^j) + (1-\delta)\delta(\pi(1,0,\mathbf{w}^j) - c) + \delta^2(\pi(1,1,\mathbf{w}^j) - c).$$

Moreover, none of the continuation plays arising on and off the equilibrium path are Pareto dominated. Therefore, this strategy profile comprises a renegotiation-proof sub-game perfect equilibrium with the desired outcome.

3.4.6 The Optimal Regime

Finally, we can talk about the principal's decision on whether or not to initiate the partial replacement of the labor with automated machinery. This decision is an immediate implication of the previous findings. One just needs to compare the expected costs of the corresponding optimal production plans in all-human and human-machine teams to reach the conclusion. Obviously the principal never prefers to keep an all-human team in production when IPE is optimal. This stems from the fact that when the principal compensates the agents based on their individual performance measures, there is no team interaction, and she cannot benefit from their mutual monitoring capacity. In this case, using automated machinery and partial elimination of agency costs is the best option for her.

Using an all-human team becomes valuable if the principal can exploit the interaction between the employees with an incentive scheme that falls within the category of JPE and RPE. With a JPE scheme the principal directly benefits from the agents' mutual monitoring capacity as she pays them only when they both receive a favorable performance measure. This leads them to monitor each other and impose peer sanctions to deter shirking. Or she can use RPE, and put the agents into a race by paying them only when their performance measures are relatively better compared to each other. The next proposition demonstrates how the principal's optimal team choice varies across the parameter space.

Proposition 3.4. *There are two critical values $\underline{\delta} < \bar{\delta}$ such that, the principal initiates a*

partial substitution of labor force with automated machinery if and only if $\delta \in [\underline{\delta}, \bar{\delta}]$*. When* $\delta < \underline{\delta}$*, and* $\delta > \bar{\delta}$ *she prefers human-human teams for production and uses RPE and JPE incentive schemes for compensation, respectively.*

Proof. See appendix C.4. □

Proposition 3.4 points out that automation is optimal only over a portion of the parameter space. The all-human regime can be preferred over the automation regime, and it is likely depending on the parameter values that the principal utilizes both the JPE and the RPE schemes. Human-machine teams become particularly valuable for an intermediate range of discount factors. In this range, the principal has a limited benefit from employing a human-human team. This is mainly due to two reasons. The advantage of using the optimal JPE scheme is higher for large values of the discount factor as it has a lower expected cost. On the other hand, using a collusion-proof RPE scheme is more advantageous for low values of the discount factor, since then collusion is a weaker concern, and the cost of precluding it is lower.

Figure 3.1 demonstrates how the optimal choice of the regimes and incentive schemes change as σ and δ changes. As the discount rate increases, for low σ, JPE becomes more preferable over automation. For low discount rates, as σ increases, the principal's ability to detect shirking agents narrows down. In this case, having a relatively better performance measure than the other member of the team is a very strong indication of high level effort for an agent. So in this range, we find that setting a collusion-free RPE incentive with all-human teams is preferred over automation.

The critical values defining the region of automation, $\underline{\delta}$ and $\bar{\delta}$, also depend on the other parameter values. For instance, when α gets smaller, the cost of operating machine becomes smaller, and as a result, the region featuring automation expands.

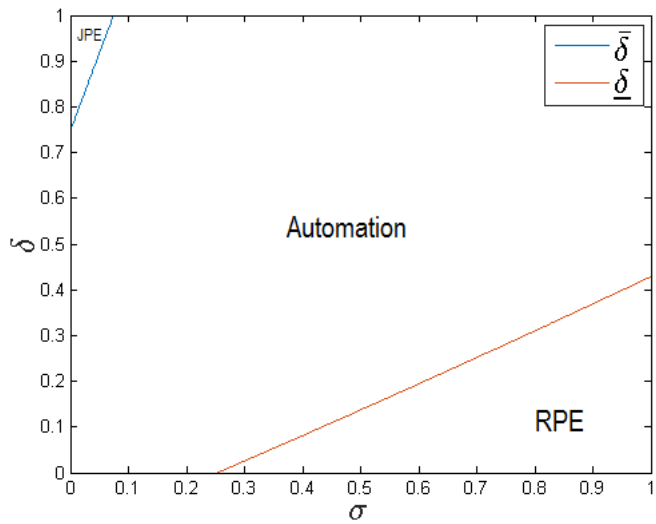

Figure 3.1: Optimal Regime and Incentive Scheme vs. σ and δ. The figure is produced under $p_1 = 0.7$, $p_0 = 0.3$, $\alpha = 0.9$.

3.4.7 Human teams survive under costless automation

Finally, we discuss a special case in which the automation is costless, i.e., $\alpha = 0$. As technology progresses, we expect that automation will become more effective because of its decreasing cost. It is reasonable to question whether the automation will be the only preferred regime as α approaches to zero. We find that this is *not* the case, as illustrated in Figure 3.2. This is a striking result as it suggests that even if automation was costless, principal may still prefer human labor over adopting it under some parameter regions.

Figure 3.2 shows that when α is equal to 0, the value of $\bar{\delta}$ is equal to 1. As a result, the region featuring a human-human team together with a JPE incentive scheme disappears. This result is not specific to the parameter values used in Figure 3.2. A human-human team together with the JPE incentive scheme $\mathbf{w}^j$ can never be better than a human-machine team when $\alpha = 0$. In other words, when $\alpha = 0$, the expected cost of $\mathbf{w}^j$ is larger than the expected cost of inducing agent to work in a human-machine team, i.e., the inequality

$$2(\sigma + (1-\sigma)p_1^2)\frac{\hat{c}}{p_1 + \delta p_0} > (\sigma + (1-\sigma)p_1)\hat{c}$$

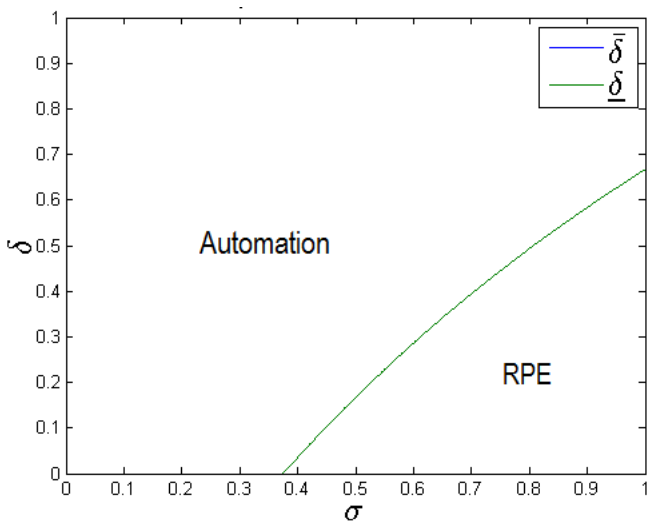

Figure 3.2: Costless Automation ($\alpha = 0$) and Optimal Regime. The figure is produced under $p_1 = 0.6$, $p_0 = 0.4$, $\alpha = 0$.

holds, regardless of the values of the parameters p_1, p_0, δ, and σ. In this respect, the idea that the JPE schemes are particularly valuable in exploiting agents' mutual monitoring abilities, which is one of the main conclusions of Che and Yoo (2001), is not likely to survive the technological progress towards reaching costless automation.

The situation is different for RPE. As it is illustrated in the figure, there is always a region featuring human-human team together with an RPE scheme even when the cost of automated machinery is 0. For this to happen, the value of σ has to be sufficiently large. In such a circumstance, the probability of aggregate shock, and hence the probability of receiving a favorable performance signal is sufficiently large. A good performance signal in itself is not a very strong indication of high level effort. As a result, the principal has a low ability to detect shirking agents. This increases the expected cost of human-machine teams as the principal has to pay the agent every time he receives a good performance signal in order to convince him to work. This means that the agent is getting paid with an excessive probability. The principal can instead hire a human-only team, and receive two signals, and strengthen her ability to detect shirking agents. This allows her to better identify the source of a favorable performance signal. She can adapt a collusion-proof RPE scheme, and

pay the agents only after they have a better performance signal than their teammates. This allows her to avoid the excessive probability at which the agents receive a positive payment.

3.5 Extensions

One of the most crucial components of our discussions so far is the fact that the partial substitution of labor with automated machinery, besides altering the structure of production, changes the extent of available information for principal. Machines, in contrast to their human counterparts, does not have a performance measure, and partial substitution of labor force eliminates a signal that carries information about the aggregate stance of the economy as well.

In this section, we study a model in which the production, in every period, results in a stochastic outcome with a distribution purely depending on the effort choices. The realization of this stochastic outcome comprises the only source of information available to principal. In this regard, the automation does not have a direct effect on available information.

We show that our main result follows here as well, i.e., a human-human team may be preferable to a human-machine team even though machine has a lower cost compared to human effort. As it was in the main model, within-team interaction and the principal's ability to exploit this interaction in her best interest via the choice of compensation scheme, comprises the main driving force behind this result. In this regard, this section can be considered as a robustness check with the main conclusion that our key insight is robust to model selection.

3.5.1 Heterogeneous Effort Decision

The employees determine their effort levels from a continuum, i.e., $e_i \in [0,1]$, $\forall i \in \{0,1\}$. The cost of effort is convex, and is assumed to be a quadratic function of the effort level, i.e., $c_i(e) = \frac{c}{2}e^2$, $\forall i \in \{0,1\}$. If the principal decides to adapt automation, she also decides at which capacity to operate the machine. This is modeled as a choice of effort from a continuum, i.e., $e_m \in [0,1]$, with the cost function $c_m(e_m) = \frac{\alpha c}{2}e_m^2$, where $\alpha \in [0,1]$. The cost of operating machinery is assumed by the principal.

We assume that the effort choices of the agents remain unobservable to principal. A stochastic outcome is realized in each period, and its distribution is determined by the joint effort of the employees, or the employee and the machine depending on the production regime. The principal can evaluate the agents' performance based on this stochastic outcome. Formally, this production outcome, in each period, results in a "success" or a "failure". In an all-human team, the probability of achieving success is equal to $f(e_1, e_2) = pe_1e_2$ for a given effort profile (e_1, e_2). Similarly, in a human-machine system, the probability of success is given by $f(e, e_m) = pee_m$.

The functional form that we assume for the probability of success has some important implications that are worth mentioning. First, there is a complementarity between the effort choices of each team member. More precisely, the marginal return from an agent's effort increases in the effort of his teammate. Second, effort of both agents are crucial in production. In other words, the probability of success is equal to 0 in case one of the agents exerts 0 effort, regardless of the other agent's choice.

The values of success and the failure for the principal are normalized to 1 and 0, respectively, and she chooses the production regime and the compensation schemes in order to maximize her expected profit. In contrast to our main model, here in this section, the principal also decides what effort level to induce the agents to choose, rather than directly inducing the maximal effort level. Before delving into the agency problem, we focus on the

first best scheme which provides a useful benchmark for the general analysis.

3.5.2 First Best Benchmark

Assume that there is a social planner who can set the level of effort for each agent. This social planner is interested in efficiency and would like to maximize total surplus. Then, the corresponding problem is given by:

$$\max_{e_1,e_2} \; pe_1e_2 - \frac{c}{2}e_1^2 - \frac{c}{2}e_2^2.$$

Due to increasing marginal returns and symmetric cost structure, the solution to this problem features either maximal or minimal effort for both agents. The ratio between the productivity and cost parameters, $\kappa \equiv \frac{p}{c}$, is crucial for the solution. A higher κ indicates a higher effective productivity, hence requires both agents to work at the maximal capacity. Similarly, a lower κ indicates a lower effective productivity and hence requires both agents to not work at all. The critical value of κ at which this switch takes place is equal to 1. Therefore, the solution to the first best problem, (e_1^*, e_2^*), is given by the following:

$$(e_1^*, e_2^*) = \begin{cases} (1,1) & \text{if } \kappa > 1 \\ (0,0) & \text{if } \kappa \leq 1. \end{cases}$$

In the rest of the analysis we assume that $\kappa > 1$ so that we have a non-trivial problem.

When it comes to the social planner's problem for a human-machine team, the situation is a bit different due to the asymmetric cost structure. The problem is the following:

$$\max_{e,e_m} \; pee_m - \frac{c}{2}e^2 - \frac{\alpha c}{2}e_m^2.$$

From the first order conditions, one can easily see that as long as the production goes on,

the social planner always set $e_m = 1$, and $e = \min\{\kappa, 1\}$.[11] Comparing the corresponding profit with the one from shutting down the production, 0, gives us the first best solution:

$$(e^*, e_m^*) = \begin{cases} (0,0) & \text{if } \kappa < \sqrt{\alpha} \\ (\kappa, 1) & \text{if } \kappa \in [\sqrt{\alpha}, 1] \\ (1,1) & \text{if } \kappa > 1. \end{cases}$$

3.5.3 Human-Machine Team

Now, we focus on the agency problem starting with human-machine teams. The principal specifies an amount w to pay the agent in case of a success, and also chooses at what level to operate the machine, i.e, e_m. Given the values of w and e_m, the agent decides his effort level e. Proposition 3.5 characterizes the optimal production plan in this regime.

Proposition 3.5. *In a human-machine team, the optimal production plan, i.e., the wage, the capacity at which the machine operates, and the induced effort level of the agent, is characterized by:*

$$(e, e_m) = \begin{cases} (0,0) & \textit{if } \kappa < \sqrt{2\alpha} \\ (\frac{\kappa}{2}, 1) & \textit{if } \kappa \in [\sqrt{2\alpha}, 2] \\ (1,1) & \textit{if } \kappa > 2 \end{cases} \qquad w = \begin{cases} \frac{1}{2} & \textit{if } \kappa \leq 2 \\ \frac{1}{\kappa} & \textit{if } \kappa > 2 \end{cases}$$

Proof. See appendix C.6. □

This statement indicates that, when the effective productivity, κ, falls below a certain threshold, i.e., $\sqrt{2\alpha}$, which is larger than the threshold of the first best benchmark due to the agency problem, the principal shuts down the production. Otherwise, she has a positive

[11] The reasons are twofold. First, e_m cannot be strictly lower then e as it is cheaper. Second, we cannot have an interior solution for both e_m, and e. To see the latter note that if some $e, e_m \in (0,1)$ brings a positive social surplus, then multiplying them with $1 + \epsilon$ increases the surplus.

profit and production continues. Likewise the first best benchmark, it is always optimal to operate the machine at its maximal capacity conditional on positive production. The induced effort from the agent, however, is not always at its maximal level due to the agency problem as well as the asymmetric cost structure. Since the principal does not observe the effort choice of the employee, she has to leave a positive rent to motivate him to work. This rent increases with the induced effort level. When κ is large, the effort is productive enough to compensate this rent, hence the principal induces the maximal effort. As κ decreases, the effort becomes less productive, and the principal is less willing to pay rent, hence she induces a lower effort level. Finally, as the value of κ becomes sufficiently small, the productivity is not large enough to compensate the cost of production and the agency anymore, and the principal shuts down the production.

3.5.4 Human-Human Team

Let the production be governed by an all-human team. The principal sets a wage w to pay each of the agents if success is achieved as the output. Given the value of w, the strategy profile that the agents follow comprises a team equilibrium, where the team equilibrium is defined as it was in the benchmark model. To this respect, the principal, by her choice of w, decides what effort levels to induce agents to exert. The next result establishes an important step towards the goal of characterizing the optimal payment scheme and the resulting team equilibrium.

Lemma 3.4. *For a given value of the wage w, a team equilibrium outcome has to be in one of the following forms:*

i) Agents choose $(e_1, e_2) = (1, 1)$ in all periods.

ii) Agents choose $(e_1, e_2) = (0, 0)$ in all periods.

Proof. See appendix C.7. □

This statement indicates that the principal either shuts down the production or specifies a wage level that induces both agents to choose their maximal effort level as a team equilibrium outcome. It is not possible to have a team equilibrium outcome in which the agents choose an interior effort level. The intuition behind this based on the fact that, if the agents were to choose an interior effort level at some period, they could do better by jointly increasing their efforts at the same rate. One shall refer to appendix for more formal discussion.

In a team equilibrium inducing the outcome $(1,1)^{\infty}$, deviations trigger a continuation play in which the agents choose their minimal efforts forever. Such a punishment is self-enforcing as the effort profile $(e_1, e_2) = (0,0)$ is always a stage game Nash equilibrium, hence can be a part of the team equilibrium.[12] This is the most extreme punishment that the agents can impose to each other as it brings the minimum possible continuation utility to both.

Let $\mathbf{e_i}(w, e_j)$ be the best response of agent i, conditional on wage w and the effort choice of his teammate e_j, which follows from the following problem:

$$\max_{e_i} pe_ie_jw - \frac{c}{2}e_i^2.$$

The first order condition lead to $\mathbf{e_i}(w, e_j) = min\{1, \kappa we_j\}$. Due to the symmetry we can further suppress the notation, and use $\mathbf{e}(w, e)$ instead. In an optimal production plan that keeps the production open, the following condition must be satisfied:

$$wp - \frac{c}{2} \geq (1-\delta)[wp\mathbf{e}(w,1) - \frac{c}{2}(\mathbf{e}(w,1))^2].$$

This incentive constraint ensures that the agents do not want to deviate from exerting maximal effort. The left hand side of the inequality is the expected utility of an agent on the

[12]Such a punishment violates the renegotiation-proofness criteria proposed in the literature. In contrast to our benchmark model, here it is not possible to find another equilibrium inducing the same outcome under the same pay scheme without violating this criteria. This stems from the fact that it is impossible to deter deviations without hurting Pareto efficiency in this framework with the functional form we assume.

equilibrium path, whereas the right hind side is his expected payoff resulting from deviation. In the current period, the agent deviates to his best response, and receives a positive payoff. In the remaining periods, he gets punished and receives a 0 payoff. If the best response effort level is equal to 1, then this incentive constraint is automatically satisfied. Therefore, for this condition to be non-redundant, we must have $\mathbf{e}(w, 1) = \min\{1, \kappa w\} = \kappa w$, and the above incentive constraint boils down to:

$$wp - \frac{c}{2} \geq (1-\delta)\frac{w^2p^2}{2c}.$$

Then, the principal's problem can be written as follows:

$$\max_w \ (1-2w)p \quad \text{s.t.} \ \ wp - \frac{c}{2} \geq (1-\delta)\frac{w^2p^2}{2c}$$

The objective function is equal to $(1-2w)p$, because both agents will get paid w if a success takes place. From this problem, it is evident that the principal chooses the minimum wage level satisfying the incentive constraint. If this wage level brings her a positive profit, she keeps the production running and induces the maximal effort from the agents; otherwise she closes the production. The next proposition formalizes this intuition.

Proposition 3.6. *The optimal production plan, i.e., the optimal wage and the induced effort levels of agents, in a human-human team is given by:*

$$(e_1, e_2) = \begin{cases} (0,0) & \textit{if } \kappa \leq \frac{2}{1+\sqrt{\delta}} \\ (1,1) & \textit{if } \kappa > \frac{2}{1+\sqrt{\delta}} \end{cases} \qquad w = \begin{cases} 0 & \textit{if } \kappa \leq \frac{2}{1+\sqrt{\delta}} \\ \frac{1}{\kappa(1+\sqrt{\delta})} & \textit{if } \kappa > \frac{2}{1+\sqrt{\delta}} \end{cases}$$

Proof. See appendix C.8. □

From Lemma 3.4, we already knew that the principal either shuts down the production or induces both agents to choose their maximal effort levels as a team equilibrium outcome. Proposition 3.6 characterizes the critical value of the effective productivity below which

shutting down the production is the optimal choice for principal. The value of this threshold decreases with the value of δ. Intuitively this stems from the fact that as the agents become more patient, the punishment that they impose each other becomes more deterrent. As a result, principal can convince the agents to exert maximal effort with a lower payment, hence she is willing to keep the production open even for lower levels of κ.

Now, we know the optimal production plans for both human-human and human-machine teams. Thus, we can figure out the corresponding profit levels, which we denote by π_{hh}, and π_{hm} respectively.

$$\pi_{hm} = \begin{cases} 0 & \text{if } \kappa < \sqrt{2\alpha} \\ \frac{p\kappa}{4} - \frac{\alpha c}{2} & \text{if } \kappa \in [\sqrt{2\alpha}, 2] \\ p - \frac{(2+\alpha)c}{2} & \text{if } \kappa > 2 \end{cases} \qquad \pi_{hh} = \begin{cases} 0 & \text{if } \kappa \leq \frac{2}{1+\sqrt{\delta}} \\ p - \frac{2c}{(1+\sqrt{\delta})} & \text{if } \kappa > \frac{2}{1+\sqrt{\delta}} \end{cases}$$

It is evident from these expressions that an all-human team may generate a higher profit for the principal depending on the parameter values. To illustrate this, suppose that the effective productivity is sufficiently large, i.e., $\kappa > 2$, so that the principal optimally induces the maximal effort level from the agents regardless of whether or not she initiates partial automation. In this case, if the discount factor is greater than $\frac{(2-\alpha)^2}{(2+\alpha)^2}$, the principal prefers a human-human team over a human-machine team. The agents are sufficiently patient, as a result, the punishment that the agents impose on each other in case of a deviation, exerting minimal effort afterwards, is more deterrent. Consequently, the principal can take advantage of the mutual monitoring capacity between the agents at a larger scale. Another important thing to point out is that as the cost parameter of operating machinery increases, the cutoff value of δ at which the above switch takes place becomes smaller. Machinery is more costly, as a result, the principal is more willing to have an all-human team in production.

3.6 Conclusions

In this paper, we ask a simple question: "Is it possible that partial subtitution of employees with automation results in inferior outcomes for principals?" We find that the answer is yes. Specifically, we extend our investigation to question how when some of the employees in an industrial or service environment are replaced by machines, the interaction between the members of a system changes. We show that replacing human labor with machines can, in some cases, be suboptimal because it can make the remaining agents more expensive.

More specifically, we demonstrate in this paper that automation has both desirable and undesirable properties. Most examples of automation introduces efficiencies in cost while reducing the uncertainty in the quality of output. Moreover, as we argue in the extensions to our model, automation can complement human output. Many examples where human employees are partially replaced by their machine counterparts exist in consumer facing environments. For instance, call center employees are replaced by automated answering systems, cash registerers are replaced by self-check out counters, ticket salespersons are replaced by kiosks, bank tellers are replaced by ATMs (Autor et al., 2002). Partial substitution of labor with machines might result in lower productivity of the remaining workers for behavioral reasons such as the fear of loss of employment in the future, boredom, losing interest in jobs due to lower social interactions, or the inability to learn from peers. All of these constitute as the undesirable effects of adopting automation at work.

In this study, we took a different approach to studying the negative impact of automation and demonstrated a reason for lower productivity under automation. Our explanation is simple yet powerful: in the absence of human team members, the principle's capacity to monitor agents' effort diminishes. This is because all-human teams have the ability to monitor each other's effort choices more closely compared to the principal can monitor their work. This close monitoring results in them imposing peer sanctions on each other when a party deviates from the optimal path of play.

In specific, we show that adopting a regime that includes partial automation does two things. The type of incentive scheme chosen by the principle (rewarding performance of agents relative to each other or jointly) depends on the automation regime adopted. We show that human teams can be incentivized based on rewarding cooperation or competition in output, conditional on the values of the discounting and common environmental shock. Importantly, considering various payment schemes, there is a parameter region under which automation is better, but also there are regions where all-human machines are more preferable. When we contrast our results to the earlier literature, we are not only looking at the outcome of automation (Acemoglu and Restrepo, 2017), but we consider the strategic decision of adopting automation. Our study, to our knowledge is the first to argue that while automation is expected to substitute labor, this is not unconditionally true the despite the benefits from reduced costs lower uncertainty in output. It is also the first to study automation in the context of teams.

When the discount rates and the degree of dependency in exogenous shocks to productivity between the workers are low, principals prefer to introduce joint performance evaluation (JPE) or automated systems over a relative performance evaluation (RPE). At high discount rates and low likelihood of a good common shock (low σ), the peer sanction by a team member is an effective disciplining device to punish the shirking member, and the manager can induce both workers to work through JPE. In this case, as members will only receive a wage when they both work, they can effectively monitor each other. If the discount rates decreased, this makes the punishment less costly, and reduces the effectiveness of JPE. Similarly, if the common shock is higher, the probability that an agent will receive a favorable outcome when he shirks is higher. So under automation, the principal's ability to detect shirking is lower. She prefers RPE over automation when the common shock is sufficiently high, because an RPE scheme without collusion motivates agents to choose to work.

An important second insight coming from our research is that automation is likely to

change the degree to which existing incentive schemes are used. The research by Che and Yoo (2001) suggests that, in the δ - σ space, JPE is more often preferred than RPE.Indeed, many firms hire individuals with contracts the employee is rewarded based on the performance of others in the same team or department. In a world where automation is an option for managers, JPE is preferred less often than RPE. In fact, if automation becomes significantly cheaper than human work, JPE is the incentive that will cease to be effective. While automation replaces some human labor, within the remaining all-human teams, RPE is more common than JPE. Broadly, this implies that as technology evolves and the cost of automation becomes cheaper, the remaining human systems are more likely to be paid based on a relative evaluation scheme. While we do not expect human teams to be obsolete in the near term, we expect that the payment schemes of employees in automated environments will be more competitive rather than collaborative. These can both be desirable and undesirable for the future of human labor.

As automated systems are gradually replacing human labor in manufacturing and service environments, managers are naturally concerned about the morale of the remaining employees. The World Bank estimates that about %60 of jobs will be automated in the near future (Frey and Osborne, 2017).

So what could the principals do to prepare for automation? First automation is likely to bring significant cost savings compared to human operators. But we show that principals should not expect to adopt automation under all conditions. In fact, automation can be more expensive compared to human labor when the principal has the ability to detect employee shirking or when the agents in a team have incentives to monitor each other's behavior.

Our findings help managers to think about how to set employee contracts to increase the efficiency of their systems. For policy makers or social planners, the concerns span beyond higher efficiency in workplaces. It is important that humans obtain and maintain employment and do not lose a significant income to maintain quality of life. While our

findings show that automation is indeed set to replace jobs, this may not necessarily be the outcome in all environments. Particularly when humans value their future earnings highly, and in economies where individuals are more likely to observe favorable outcomes, automation may be less preferred by the principles. This statement, however, should be taken with a grain of salt since out study only applies to environments with partial substitution. When full substitution is possible, then a manager could always yield higher benefits. Full automation in all work environments or may happen in the not so near future. Until then, there are reasons to believe that human teams will not become obsolete.

For future research, we leave many interesting questions. Many service robots in practice are installed in consumer facing environments (ATM machines, automatized call centers, recommendation and information kiosks in stores). Moreover, robots, machines and smart technologies are also helping consumers to go through their days, sometimes partially replacing human work (e.g., Roomba replacing cleaning services, Alexa replacing personal shopping assistants, etc.). As consumers are becoming more dependent on these new technologies, this opens up a whole new area of research for marketers.

In marketing, for example, retail sales personnel are beginning to be replaced by machines that provide information about products, check availability of products in the inventory, or make recommendations based on what the consumer is searching for. One relevant question is how to set the salesforce incentives when they are working in such partly automated retail environments so that they still engage in informing and educating consumers about the products and not simply rely on the consumers to obtain that information from the machines.

Appendix A

Appendix to Chapter 1

A.1 Proof of Lemma 1.1

The principal can always randomize over different values of U with a restriction that the expectation resulting from the randomization is exactly equal to the promised utility. This immediately requires the concavity of the value function V, and then the almost everywhere differentiability directly follows.[1]

To see the bounds indicated in expression (1.1), first note that $V(0) = V^*(0) = 0$ and $V(v) = V^*(v) = v - \mu$. We also know that $V(U) \leq V^*(U)$ for every $U \in [0, v]$, since the value function V^* is defined by an optimization problem with a smaller set of constraints compared to the one of the value function V. Moreover, V' is decreasing over the state space as it is a concave function. Therefore its derivative cannot be larger than the slope of V^* at 0 and cannot be smaller than the slope of V^* at v. But we know that the value function V^* is piece-wise linear, and its slope is $1 - \frac{\mu\lambda}{\mu\lambda v + c}$ over the interval $[0, (1-\mu\lambda)v - c)$, and equal to $1 - \frac{\mu(1-\lambda)}{(1-\mu\lambda)v - c}$ over the interval $((1-\mu\lambda)v - c, v]$. This concludes our proof.

[1] One can also apply the result of Benveniste and Scheinkman (1979), to show the differentiability of the value function. To see how their result can be applied in this context see for example (Horner and Guo, 2015).

A.2 Proof of Lemma 1.2

The continuation utilities are given in the equations 1.2 and 1.3. By utilizing these expressions, it will be shown that the approval probabilities x_s and x_{ns} will be set equal to their efficient levels, 0 and 1 respectively, as long as this does not violate the promise keeping constraint.

To this end, assume that $x_{ns}(U_D) < 1$ for some value of U_D. Then consider the following deviation that is acquired by increasing x_{ns} by ϵ and decreasing U_{ns} by $\frac{(1-\delta)v}{\delta}\epsilon$ for a sufficiently small ϵ. Note that this change is respecting the constraint PK. In consequence, the principal's utility increases by:

$$\Delta = (1-\mu\lambda)\left[(1-\delta)\epsilon(v - \frac{\mu(1-\lambda)}{1-\mu\lambda}) + \delta\left(V(U_{ns}) - V(U_{ns} - \frac{(1-\delta)v}{\delta}\epsilon)\right)\right]$$

Moreover, from lemma 1.1, it is known that $V' \leq 1 - \frac{\mu(1-\lambda)}{(1-\mu\lambda)v-c}$. Then by using the fundamental theorem of calculus, one can get:[2]

$$\Delta \geq (1-\delta)\epsilon(1-\mu\lambda)\left[v - \frac{\mu(1-\lambda)}{1-\mu\lambda}v - \frac{\mu(1-\lambda)v}{(1-\mu\lambda)v-c}\right] > 0.$$

Therefore this deviation, in case if it is feasible, strictly benefits the principal. The suggested deviation would not be feasible if $U_{ns} = 0$, which occurs when $U_D < (1-\delta)v$. Therefore, when $U < \underline{U} = (1-\delta)v$ the approval rate x_{ns} will be set such that the continuation utility $U_{ns} = 0$, otherwise $x_{ns} = 1$. This proves the lemma for x_{ns} and U_{ns}.

Now suppose that $x_s > 0$ for some value of U. Then consider decreasing x_s by ϵ and hence increasing U_s by $\frac{(1-\delta)v}{\delta}\epsilon$ for a sufficiently small ϵ. This results with a change Δ that

[2]The concavity of the value function V is sufficient for utilizing the fundamental theorem.

is equal to:

$$\Delta = \mu\lambda \left[-(1-\delta)\epsilon(v-1) + \delta \left(V(U_s) - V(U_s + \frac{(1-\delta)v}{\delta}\epsilon) \right) \right]$$

Again from lemma 1.1, we know that $V' \geq 1 - \frac{\mu\lambda}{\mu\lambda v + c}$, therefore:

$$\Delta \geq (1-\delta)\epsilon\mu\lambda \left[1 - v + v - \frac{\mu\lambda v}{\mu\lambda v + c} \right] > 0.$$

Therefore this deviation, as long as it is feasible, strictly benefits the principal. The suggested deviation would not be feasible if $U_s = v$, which occurs when $U_D > \delta v - \frac{(1-\delta)c}{\mu\lambda}$. Therefore when $U > \bar{U} = \delta v - \frac{(1-\delta)c}{\mu\lambda}$ the approval rate x_s will be set such that the continuation utility $U_s = v$, otherwise $x_s = 0$. This proves the lemma for x_s and U_s.

A.3 Proof of Lemma 1.3

First remember that $V(0) = V_N(0)$ and $V(v) = V_N(v)$. Moreover, from the optimal choice of x_n which is indicated in (1.4), it is known that:

$$V_N'(U) = \begin{cases} V'(\frac{U}{\delta}) & \text{if } U \leq \delta\underline{a} \\ \frac{v-\mu}{v} & \text{if } U \in (\delta\underline{a}, \delta\bar{a} + (1-\delta)v) \\ V'(\frac{U-(1-\delta)v}{\delta}) & \text{if } U \geq \delta\bar{a} + (1-\delta)v \end{cases}$$

Then by the concavity, the comparison of V' and V'_N satisfies:

$$V'_N(U) \leq V'(U) \text{ if } U \in [0, \delta\underline{a}]$$
$$V'_N(U) < V'(U) \text{ if } U \in (\delta\underline{a}, \underline{a})$$
$$V'_N(U) = V'(U) \text{ if } U \in [\underline{a}, \bar{a}]$$
$$V'_N(U) > V'(U) \text{ if } U \in (\bar{a}, \delta\bar{a} + (1-\delta)v)$$
$$V'_N(U) \geq V'(U) \text{ if } U \in [\delta\bar{a} + (1-\delta)v, v]$$

But this requires that:

$$V_N(U) \leq V(U) \text{ if } U \in [0, \delta\underline{a}]$$
$$V_N(U) < V(U) \text{ if } U \in (\delta\underline{a}, \delta\bar{a} + (1-\delta)v)$$
$$V_N(U) \leq V(U) \text{ if } U \in [\delta\bar{a} + (1-\delta)v, v]$$

Otherwise it would not be possible to have $V(0) = V_N(0)$ and $V(v) = V_N(v)$. Therefore, the equality $V = V_N$ can only arise in $[0, \delta\underline{a}] \cup [\delta\bar{a} + (1-\delta)v, v]$, and this in turn implies that $x_n \in \{0, 1\}$ in the optimal policy.

Focusing on the interval $[0, \delta\underline{a}]$ first, from the solution of the problem $\mathcal{P}_N$, it is known that $x_n = 0$ and $V'_N \leq V'$. Then due to the fact $V(0) = V_N(0)$, it immediately follows that there must exists a critical value $\underline{U}_N$, satisfying the following:

$$\underline{U}_N = \sup\{U \in [0, v] \mid V = V_N \text{ over } [0, U]\}$$

Since $V_N < V$ over the interval $(\delta\underline{a}, \delta\bar{a} + (1-\delta)v)$, we must have $U_N \leq \delta\underline{a}$. In principle U_N might be 0 so that $[0, U_N] = \{0\}$. However, we will show that $\underline{U}_N > 0$.

Suppose $\underline{U}_N = 0$ to get a contradiction, and hence $V_N(U) < V(U)$ for each U sufficiently close to 0. However, when U is sufficiently close to 0, the solution to the problem $\mathcal{P}_N$ implies

that $x_s = 0$, $x_{ns} = \frac{U}{(1-\delta)v}$, $U_{ns} = 0$, and $U_s = \frac{U}{\delta} + \frac{(1-\delta)c}{\delta\mu\lambda}$, therefore:

$$V_D(U) = (1-\delta)\left[(1-\mu\lambda)\frac{U}{(1-\delta)v}(v - \frac{\mu(1-\lambda)}{1-\mu\lambda}) - c\right] + \delta\mu\lambda V(\tfrac{U}{\delta} + \tfrac{(1-\delta)c}{\delta\mu\lambda})$$

Then from the upper bound of V', it is known that:

$$V(\tfrac{U}{\delta} + \tfrac{(1-\delta)c}{\delta\mu\lambda}) \leq \left(\frac{U}{\delta} + \tfrac{(1-\delta)c}{\delta\mu\lambda}\right)\left(1 - \frac{\mu(1-\lambda)}{(1-\mu\lambda)v - c}\right).$$

Thus, $V_D(U) < 0$ for sufficiently small U. Which suggests that $V_D(U) < V(U)$, as well as $V_N(U) < V(U)$, when U goes to 0.

Consequently, there must be a randomization, i.e. $\gamma \in (0,1)$, for the values of U that are close enough to 0. This in turn, requires that V is linear in this region, with a slope m. Then in this region, $V_N(U) = \delta V(\frac{U}{\delta}) = \delta m \frac{U}{\delta} = V(U)$, and this gives us a contradiction as we assumed that $V_N < V$. Therefore $U_N > 0$.

To show the linearity of V, and hence V_N, over $[0, U_N]$, note that $V'(U) = V_N'(U) = V'(\frac{U}{\delta})$ in this region. However, since V is concave, the derivative is weakly decreasing. Hence the derivative must be constant, i.e the value function is linear.

The proof for the other end of the state space follows from analogous arguments.

A.4 Proof of Lemma 1.4

Refer to the main text for the proof of the first part, i.e. $V(U) = V_D(U)$, and $\gamma(U) = 1$ for every $U \in [\underline{U}_D, \bar{U}_D]$.

To show the second part of the result, i.e the inequality $\underline{U}_D \leq \underline{U} < \bar{U} \leq \bar{U}_D$, I focus on the comparison between $\underline{U}_D$ and $\underline{U}$. To get a contradiction, suppose that the inequality

$\underline{U}_D > \underline{U}$ holds, and hence:

$$V(\underline{U}_D) = (1-\delta)\left[(1-\mu\lambda)v - \mu(1-\lambda) - c\right] + \delta\left[\mu\lambda V(\tfrac{\underline{U}_D}{\delta} + \tfrac{(1-\delta)c}{\delta\mu\lambda}) + (1-\mu\lambda)V(\tfrac{\underline{U}_D-(1-\delta)v}{\delta})\right]$$
$$= m\underline{U}_D$$

where m is the constant slope of V over $[0, \underline{U}_D]$. Then, consider some $U = \underline{U}_D - \epsilon$ for a sufficiently small $\epsilon > 0$ satisfying $U > \underline{U}$. Clearly:

$$V_D(U) = (1-\delta)\left[(1-\mu\lambda)v - \mu(1-\lambda) - c\right] + \delta\left[\mu\lambda V(\tfrac{U}{\delta} + \tfrac{(1-\delta)c}{\delta\mu\lambda}) + (1-\mu\lambda)V(\tfrac{U-(1-\delta)v}{\delta})\right]$$
$$> V(\underline{U}_D) - m\epsilon.$$

The strict inequality stems from the fact that:

$$V(\tfrac{\underline{U}_D}{\delta} + \tfrac{(1-\delta)c}{\delta\mu\lambda}) - V(\tfrac{U}{\delta} + \tfrac{(1-\delta)c}{\delta\mu\lambda}) < m\frac{\epsilon}{\delta}$$
$$V(\tfrac{\underline{U}_D-(1-\delta)v}{\delta}) - V(\tfrac{U-(1-\delta)v}{\delta}) = m\frac{\epsilon}{\delta}.$$

But this consists of a contradiction, since

$$V_D(U) > V(\underline{U}_D) - m\epsilon = m(\underline{U}_D - \epsilon) = V(U)$$

Therefore the inequality $\underline{U}_D \leq \underline{U}$ holds. By analogous arguments one can also show that $\bar{U}_D \geq \bar{U}$, hence the proof is complete.

A.5 Proof of Proposition 1.1

To complete the proof, one just needs to show that the value function V is strictly concave over the interval $[\underline{U}_D, \bar{U}_D]$.

It is already shown that $\underline{U}_D \leq \underline{U} < \bar{U} \leq \bar{U}_D$. Moreover, the description of the optimal

approval decisions (x_s, x_{ns}), and the continuation utilities (U_s, U_{ns}) are provided in lemma 1.2. Then, by using the fact that $V = V_D$ over $[\underline{U}_D, \bar{U}_D]$, one can get:

$$V(U) = \begin{cases} (1-\delta)(1-\mu\lambda)\frac{U}{(1-\delta)v}(v-\mu_{ns}) + \delta\mu\lambda V(U_s) & \text{if } U_D \in [\underline{U}_D, \underline{U}) \\ (1-\delta)\pi + \delta[(1-\mu\lambda)V(U_{ns}) + \mu\lambda V(U_s)] & \text{if } U_D \in [\underline{U}, \bar{U}] \\ (1-\delta)[\pi + \mu\lambda(\frac{U+\frac{(1-\delta)c}{\mu\lambda}-\delta v}{(1-\delta)v})(v-1)] + \delta(1-\mu\lambda)V(U_{ns}) & \text{if } U_D \in (\bar{U}, \bar{U}_D] \end{cases}$$

Where the continuation utilities are given by:

$$U_s = \frac{U}{\delta} + \frac{(1-\delta)c}{\delta\mu\lambda} \text{ and } U_{ns} = \frac{U}{\delta} - \frac{(1-\delta)v}{\delta}.$$

Hence the derivative of V over $[\underline{U}_D, \bar{U}_D]$ satisfies:

$$V'(U) = \begin{cases} (1-\mu\lambda)(\frac{v-\mu_{ns}}{v}) + \mu\lambda V'(U_s) & \text{if } U_D \in (\underline{U}_D, \underline{U}] \\ (1-\mu\lambda)V'(U_{ns}) + \mu\lambda V'(U_s) & \text{if } U_D \in [\underline{U}, \bar{U}] \\ (1-\mu\lambda)V'(U_{ns}) + \mu\lambda(\frac{v-1}{v}) & \text{if } U_D \in (\bar{U}, \bar{U}_D] \end{cases}$$

I want to show that, there does not exist an interval in $[\underline{U}_D, \bar{U}_D]$ over which the value function is linear. Suppose not in order to get a contradiction. Let $\mathcal{I}$ be the largest interval with a linear V.

First, note that $\frac{v-\mu_{ns}}{v} > \bar{s}$, and $(\frac{v-1}{v}) < \underline{s}$, where $\bar{s}$ and $\underline{s}$ are the upper and lower bounds for V' respectively.[3] Therefore, V' cannot stay constant in any neighborhoods of $\underline{U}$ and $\bar{U}$. As a result, there are three possible cases: i)$\mathcal{I} \subset [\underline{U}_D, \underline{U})$, ii) $\mathcal{I} \subset (\bar{U}, \bar{U}_D]$, iii) $\mathcal{I} \subset [\underline{U}, \bar{U}]$.

For the first case, due to the concavity of V, V' must be constant over $U_s|_{\mathcal{I}} = (U_s(inf(\mathcal{I})), U_s(sup(\mathcal{I})))$. Moreover, due to the lower bound on δ, $U_s(inf(\mathcal{I})) < \bar{U}$, hence $U_s|_{\mathcal{I}} \subset [\underline{U}_D, \bar{U}_D]$. This gives an immediate contradiction with the definition of $\mathcal{I}$, since the

[3] See lemma 1.1.

length of $U_s|_{\mathcal{I}}$ is larger than the length of $\mathcal{I}$. An analogous contradiction carries over the second case.

For the last case, V' must be constant along the intervals $U_s|_{\mathcal{I}}$ and $U_{ns}|_{\mathcal{I}}$, due to the concavity of V. Moreover, either $U_s(inf(\mathcal{I})) < \bar{U}$ or $U_{ns}(sup(\mathcal{I})) > \bar{U}$ must be correct, because both of them cannot be wrong at the same time due to the lower bound on δ. Therefore, either $U_s|_{\mathcal{I}}$ or $U_{ns}|_{\mathcal{I}}$ must be the subset of $[\underline{U}, \bar{U}] \subset [\underline{U}_D, \bar{U}_D]$, which contradicts with the definition of $\mathcal{I}$, because both of these intervals have a larger length than $\mathcal{I}$.

A.6 Proof of Lemma 1.5

Define:

$$\underline{b} = \inf\{U \in [0, v] \mid \tilde{V}'(U) \leq \frac{v-\mu}{v}\},$$

$$\bar{b} = \sup\{U \in [0, v] \mid \tilde{V}'(U) \geq \frac{v-\mu}{v}\}.$$

Then the optimal approval decision conditional on no self-monitoring satisfies:

$$\tilde{x}_n(U_N) = \begin{cases} 0 & \text{if } U_N \leq \delta\underline{b} \\ \in (0,1) & \text{if } U_N \in \big(\delta\underline{b}, \delta\bar{b} + (1-\delta)v\big) \\ 1 & \text{if } U_N \geq \delta\bar{b} + (1-\delta)v \end{cases} \cdot \tag{A.1}$$

Therefore:

$$\tilde{V}'_N(U_N) = \begin{cases} \tilde{V}'(\frac{U}{\delta}) & \text{if } U_N \leq \delta\underline{b} \\ \frac{v-\mu}{v} & \text{if } U_N \in (\delta\underline{b}, \delta\bar{b} + (1-\delta)v) \\ \tilde{V}'(\frac{U-(1-\delta)v}{\delta}) & \text{if } U_N \geq \delta\bar{b} + (1-\delta)v \end{cases}$$

Clearly $\tilde{V}_N(v) = \tilde{V}(v)$, since there is only one possible way to provide the maximal utility to the agent. However, unlike the previous case, the equality $\tilde{V}_N(0) = \tilde{V}(0)$ does not hold. To see this, first observe that $\tilde{V}_N(0) = 0$. Then, in order to examine the value of

$\tilde{V}_D(0)$, by using the fact that the incentive and the promise keeping constraints are binding, we get:

$$\tilde{U}_{ns}(0) = -\frac{(1-\delta)v}{\delta}\tilde{x}_{ns}(0),$$

$$\tilde{U}_s(0) = \frac{1-\delta}{\delta\mu\lambda}c - \frac{(1-\delta)v}{\delta}\tilde{x}_s(0).$$

Obviously, $\tilde{x}_{ns} = 0$ since the continuation utility $\tilde{U}_{ns}(0)$ cannot be negative. Moreover, $\tilde{U}_s(0) = \frac{1-\delta}{\delta\mu\lambda}c$, since the optimal choice of $\tilde{x}_s(0)$ is 0.[4] As a result, $\tilde{V}_D(0) = \delta\tilde{V}(\frac{(1-\delta)c}{\delta\mu\lambda})$, which is strictly positive. Therefore we must have $\tilde{V}_N < \tilde{V}(0) = \tilde{V}_D(0)$. Then by using this and fundamental theorem of calculus together with the following

$$\tilde{V}_N'(U) \leq \tilde{V}'(U) \text{ if } U \in [0, \delta\underline{b}]$$

$$\tilde{V}_N'(U) < \tilde{V}'(U) \text{ if } U \in (\delta\underline{b}, \underline{b})$$

$$\tilde{V}_N'(U) = \tilde{V}'(U) \text{ if } U \in [\underline{b}, \bar{b}]$$

$$\tilde{V}_N'(U) > \tilde{V}'(U) \text{ if } U \in (\bar{b}, \delta\bar{b} + (1-\delta)v)$$

$$\tilde{V}_N'(U) \geq \tilde{V}'(U) \text{ if } U \in [\delta\bar{b} + (1-\delta)v, v]$$

we can conclude that there exists $\tilde{U}_N$ such that $\tilde{V}(U) = \tilde{V}_N(U)$ if and only if $U \in [\tilde{\tilde{U}}_N, v]$. In addition, the facts that $\tilde{U}_N$ is strictly smaller than v, and $\tilde{V}$ is linear over $[\frac{\tilde{U}_N-(1-\delta)v}{\delta}, v]$ follow from exactly the same arguments provided in the proof of lemma 1.3. On the other hand, since $\tilde{\tilde{U}}_N \in [\delta\bar{b} + (1-\delta)v, v]$, it immediately follows that $\tilde{x}_n = 1$ over $[\tilde{\tilde{U}}_N, v]$.

A.7 Proof of Proposition 1.2

Initially note that it is already known that the equality $\tilde{V} = \tilde{V}_D$ holds at 0 and $\tilde{U}_D$, Then the same logic that is used in lemma 1.3, it immediately follows that the equality $\tilde{V} = \tilde{V}_D$ holds for each $U \in [0, \tilde{U}_D]$.

[4] Refer to earlier discussions to see why it is optimal to set $\tilde{x}_s$ equal to 0.

When it comes to the approval decisions, the case for $\tilde{x}_s$ has already been discussed. Here the focus will be on the contingency of no self-reporting, i.e the choice variables $\tilde{x}_{ns}$, and $\tilde{U}_{ns}$. Since these variables are isolated from the other contingency, their choice satisfy the following local problem:

$$\max_{\tilde{x}_{ns}} (1-\delta)\tilde{x}_{ns}(v - \tfrac{\mu(1-\lambda)}{1-\mu\lambda}) + \delta\tilde{V}(\tilde{U}_{ns})$$
$$\text{s.t. } (1-\delta)v\tilde{x}_{ns} + \delta\tilde{U}_{ns} = U$$

From the constraint, it is possible to substitute between $\tilde{x}_{ns}$ and $\tilde{U}_{ns}$ at a rate 1 to $\frac{(1-\delta)v}{\delta}$; moreover their marginal returns for the principal are $(1-\delta)(v - \frac{\mu(1-\lambda)}{1-\mu\lambda})$, and $\delta\tilde{V}'(U_{ns})$ respectively. Therefore, showing that $\tilde{V}' < 1 - \frac{\mu(1-\lambda)}{(1-\mu\lambda)v}$, would be sufficient to complete the proof. Suppose not to get a contradiction, and define:

$$\underline{d} = \inf\{U \in [0,v] \mid \tilde{V}'(U) \leq 1 - \frac{\mu(1-\lambda)}{(1-\mu\lambda)v}\},$$

$$\bar{d} = \sup\{U \in [0,v] \mid \tilde{V}'(U) \geq 1 - \frac{\mu(1-\lambda)}{(1-\mu\lambda)v}\}.$$

From the hypothesis it is known that $\bar{d} > 0$. Moreover, $\tilde{x}_{ns}(U) = 0$, $\tilde{U}_{ns}(U) = \frac{U}{\delta}$, and $\tilde{x}_s(U) = \frac{U}{\delta} + \frac{(1-\delta)c}{\delta\mu\lambda}$, $\forall U \in [0, \delta\underline{d}]$. This would require that $\tilde{V}' = (1-\mu\lambda)\tilde{V}'(\tilde{U}_{ns}) + \mu\lambda\tilde{V}'(\tilde{U}_s)$, i.e the derivative of the value function is equal to the the expectation of the derivative over the continuation values. This in turn requires $\tilde{V}'$ to be constant over $[0, \underline{d} + \frac{(1-\delta)c}{\delta\mu\lambda}]$, due to the concavity together with the fact that $\tilde{U}_s(U), \tilde{U}_{ns}(U) > U$, for every $U \in [0, \delta\underline{d}]$. Therefore $\underline{d}$ must be equal to 0.

Now it is known that, $\tilde{V}'(U) = 1 - \frac{\mu(1-\lambda)}{(1-\mu\lambda)v}$, for every $U \in [0, \bar{d}]$. Then again from the same logic, this constant slope must carry over to a larger region, and hence constitutes a contradiction with the definition of $\bar{d}$, as it requires that $V'(\tilde{U}_s(U)) = 1 - \frac{\mu(1-\lambda)}{(1-\mu\lambda)v}$, $\tilde{x}_s = 0$, $\forall U \in [0, \bar{d}]$. This stems from the fact that $\tilde{V}' = (1-\mu\lambda)\left(1 - \frac{\mu(1-\lambda)}{(1-\mu\lambda)v}\right) + \mu\lambda\tilde{V}'(\tilde{U}_s)$, as the

principal is indifferent between the marginal increase on $\tilde{x}_{ns}$ and $\tilde{U}_{ns}$ in this region. As a result we have $\tilde{V}' < 1 - \frac{\mu(1-\lambda)}{(1-\mu\lambda)v}$, and the result immediately follows.

Finally the strict concavity, and the inequality $\tilde{U}_D > \bar{U}$ follows from the identical arguments presented in the previous case, hence not repeated here.

A.8 Proof of Lemma 1.6

First of all, $U_{max} = \sup\{W \mid V_W(W) \geq 0\}$ exists since W takes its values from a bounded interval. Therefore, to complete the proof, one just needs to show that $V_{U_{max}}(U) \geq 0$ for each $U \in [0, U_{max}]$. Suppose not to get a contradiction, and assume that $V_{U_{max}}(U_{max}) < 0$. Then the definition of U_{max} implies that $\forall \epsilon > 0$, $\exists U_\epsilon \in [U_{max} - \epsilon, U_{max}]$, such that $V_{U_\epsilon}(U_\epsilon) \geq 0$. However, it is known that $V_W(U)$ is continuous on U and W, therefore $V_W(W)$ must be continuous in W. Therefore $\exists \epsilon^* > 0$ such that $\forall W \in [U_{max} - \epsilon^*, U_{max}]$, $V_W(W) < 0$. Which is a contradiction. Therefore $V_{U_{max}}(U_{max}) \geq 0$.

To show that the last inequality holds with equality, suppose that $V_{U_{max}}(U_{max}) > 0$. Then the principal could grant the agent a utility level that is slightly higher than U_{max} due to the continuity of $V_W(w)$, and this contradicts with the definition of U_{max}.

Finally, in order to show that increasing μ strictly decreases U_{max}, let μ and μ' are two values with $\mu > \mu'$. Note that, an incentive compatible policy when the prior is equal to μ is also incentive compatible when prior is μ' and vice versa. This stems from the fact that the agent does not care about the type of the project. Let U_{max} be the corresponding maximal state variable when prior is μ. Take an optimal policy when the agent is given U_{max}. It is clear that this policy would bring a strictly positive value to principal when the prior is μ', therefore the corresponding maximal state variable, U'_{max} must be strictly larger than U_{max}.

Appendix B

Appendix to Chapter 2

B.1 Proof of Lemma 2.1

The "only if part" of the statement follows directly from revealed preferences. In other words, if the mechanism is incentive compatible, then the buyer of type x does not want to mimic type x' and vice versa. More precisely take $x > x'$, then:

$$\begin{aligned} U(x) \geq \ & e^{-rt_{x'}}(\frac{x}{r} - p_{t_{x'}}) = U(x') + e^{-rt_{x'}}(\frac{x}{r} - \frac{x'}{r}) \\ U(x') \geq \ & e^{-rt_{x}}(\frac{x'}{r} - p_{t_{x}}) = U(x) - e^{-rt_{x}}(\frac{x}{r} - \frac{x'}{r}) \end{aligned}$$

Then we get,

$$\frac{e^{-rt_x}}{r} \geq \frac{U(x) - U(x')}{x - x'} \geq \frac{e^{-rt_{x'}}}{r} \tag{B.1}$$

which requires $t_x \leq t_{x'}$. Therefore we get $i)$. Now, since t_x is monotone it is differentiable and continuous almost everywhere. Therefore e^{-rt_x} is differentiable and continuous a.e. and hence, $lim_{x' \to x} \frac{e^{-rt_{x'}}}{r} = \frac{e^{-rt_x}}{r}$ a.e. We also know that $U(x)$ is continuous and differentiable

a.e so by taking the limit of the expression $(B.1)$, when $x' \to x$, we get

$$\frac{\partial U(x)}{\partial x} = \frac{1}{r} e^{-rt_x} \quad \text{a.e.}$$

Hence, $ii)$ follows immediately.

For the "if" part, suppose for a given mechanism conditions $i)$ and $ii)$ are satisfied, and we want to show that this mechanism is incentive compatible. Take any two arbitrary types x, and x' and WLOG assume $x > x'$. First, we want to show that x does not want to report his type as x'. In other words the following must be true

$$U(x) \geq \; e^{-rt_{x'}}(\frac{x}{r} - p_{t_{x'}}) = \; U(x') + e^{-rt_{x'}}(\frac{x}{r} - \frac{x'}{r}). \tag{B.2}$$

However by $ii)$ we know that

$$U(x) - U(x') = \frac{1}{r} \int_{x'}^{x} e^{-rt_{\tilde{x}}} d\tilde{x}.$$

Hence, expression $(B.2)$ boils down to:

$$\int_{x'}^{x} e^{-rt_{\tilde{x}}} d\tilde{x} \geq e^{-rt_{x'}}(x - x').$$

But this is correct given monotonicity in $i)$. Similar arguments follow for the reports of x' as well. Hence the statement is true.

B.2 Proof of Lemma 2.4

i) Monotonicity: Take arbitrarily two agents of type x, and x', where $x > x'$ without loss of generality, we want to show that $t_x \leq t_{x'}$. Showing that purchasing the good at time $t > t_{x'}$ is worse then purchasing it at $t_{x'}$ for agent x is sufficient to prove monotonicity. To this end, take an arbitrary t satisfying $t > t_{x'}$. We know by revealed preferences of agent

x' that:

$$U(x') \;\geq\; e^{-(r+\lambda)t}\,(\frac{x'}{r} - p_t) + \int_0^t e^{-(r+\lambda)T}\lambda V^T_{\beta x'}dT + \int_t^\infty e^{-(r+\lambda)T}\lambda V^T_{(\beta-1)x'}dT$$

Then we get:

$$\frac{x'}{r}(e^{-(r+\lambda)t_{x'}} - e^{-(r+\lambda)t}) - (e^{-(r+\lambda)t_{x'}}p_{t_{x'}} - e^{-(r+\lambda)t}p_t) \geq \int_{t_{x'}}^t e^{-(r+\lambda)T}\lambda(V^T_{\beta x'} - V^T_{(\beta-1)x'})dT$$

We want show that the symmetric version of the above expression holds for agent x as well. Hence we need show

$$\frac{x-x'}{r}(e^{-(r+\lambda)t_{x'}} - e^{-(r+\lambda)t}) \;\geq\; \int_{t_{x'}}^t e^{-(r+\lambda)T}\lambda\big((V^T_{\beta x} - V^T_{(\beta-1)x}) - (V^T_{\beta x'} - V^T_{(\beta-1)x'})\big)dT. \tag{B.3}$$

Now to show the inequality above is correct we need to consider two cases.

- Case 1 : $[x' < \frac{\beta-1}{\beta}x]$

Incentive compatibility in the second step is satisfied by hypothesis. Therefore, by the second condition in Lemma 2.3 we know that the highest possible value of $\big((V^T_{\beta x} - V^T_{(\beta-1)x}) - (V^T_{\beta x'} - V^T_{(\beta-1)x'})\big)$ can be attained by arranging $t^T_z = 0$, for all $z \in [(\beta-1)x, \beta x]$, and $t^T_z = \infty$, for all $z \in [0,\ (\beta-1)x)$. Therefore

$$\big((V^T_{\beta x} - V^T_{(\beta-1)x}) - (V^T_{\beta x'} - V^T_{(\beta-1)x'})\big) \;\leq\; \frac{x}{r},$$

which leads to

$$\int_{t_{x'}}^t e^{-(r+\lambda)T}\lambda\big((V^T_{\beta x} - V^T_{(\beta-1)x}) - (V^T_{\beta x'} - V^T_{(\beta-1)x'})\big)dT \;\leq\; x\frac{\lambda}{r(r+\lambda)}(e^{-(r+\lambda)t_{x'}} - e^{-(r+\lambda)t}).$$

However, since $x' < \frac{\beta-1}{\beta}x$, and $\lambda < \frac{r}{\beta-1}$, we know that

$$\frac{x-x'}{r}(e^{-(r+\lambda)t_{x'}} - e^{-(r+\lambda)t}) \;>\; x\frac{\lambda}{r(r+\lambda)}(e^{-(r+\lambda)t_{x'}} - e^{-(r+\lambda)t}).$$

Therefore equation (B.3) is satisfied and we are done for this case.

- Case 2 : $[x' \geq \frac{\beta-1}{\beta}x]$

Again by Lemma 2.3 the highest possible value of $\big(((V^T_{\beta x} - V^T_{(\beta-1)x}) - (V^T_{\beta x'} - V^T_{(\beta-1)x'})\big)$ can be attained by arranging $t^T_z = 0$, for all $z \in [\beta x', \beta x]$, and $t^T_z = \infty$, for all $z \in [0, \beta x')$. Therefore,

$$\big((V^T_{\beta x} - V^T_{(\beta-1)x}) - (V^T_{\beta x'} - V^T_{(\beta-1)x'})\big) \;\leq (x-x')\frac{\beta}{r}.$$

Hence

$$\int_{t_{x'}}^{t} e^{-(r+\lambda)T}\lambda\big((V^T_{\beta x} - V^T_{(\beta-1)x}) - (V^T_{\beta x'} - V^T_{(\beta-1)x'})\big)dT \;\leq (x-x')\frac{\beta\lambda}{r(r+\lambda)}(e^{-(r+\lambda)t_{x'}} - e^{-(r+\lambda)t}).$$

However, since $\lambda < \frac{r}{\beta-1}$, we know that

$$\frac{x-x'}{r}(e^{-(r+\lambda)t_{x'}} - e^{-(r+\lambda)t}) \;>\; (x-x')\frac{\lambda\beta}{r(r+\lambda)}(e^{-(r+\lambda)t_{x'}} - e^{-(r+\lambda)t}).$$

So equation (B.3) is valid for this case as well. Hence we are done to show monotonicity.

ii) Derivative of V_x: By truthfully reporting, agent get the utility

$$U(x) = V_x + \int_0^{t_x} e^{-(r+\lambda)T}\lambda V^T_{\beta x}dT + \int_{t_x}^{\infty} e^{-(r+\lambda)T}\lambda V^T_{(\beta-1)x}dT.$$

Now, for a given type x' with $x > x'$, what would happen if the agent x reports his type as x' at first stage reports? He would be allocated the first version of the good at time $t_{x'}$ rather than t_x where, from from part i), we know that $t_x \leq t_{x'}$. This deviation from truth-telling will affect his utility via two different channels. The first channel is a direct

effect as he now acquires the first version at a different time.[1] The second channel is an indirect effect due to the change on second stage utility.

We know that the second step of the mechanism is incentive compatible and so the agent x reports his effective type truthfully in the second stage. Because of this, misreporting in the first stage alters the reports of the second stage only if it alters the effective types at the realized arrival time. This happens only if the arrival occurs between t_x and $t_{x'}$. In particular, after truthful reporting, the effective type of agent x would be $(\beta-1)x$ inside the time interval $(t_x, t_{x'})$, and it would be βx if he deviates and misreports its type as x'. Therefore, the incentive constraint of agent x preventing him to not to mimic x' is

$$\begin{aligned} U(x) \geq\ & e^{-(r+\lambda)t_{x'}}\big(\frac{x}{r} - p_{t_{x'}}\big) + \int_0^{t_{x'}} e^{-(r+\lambda)T}\lambda V^T_{\beta x} dT + \int_{t_{x'}}^{\infty} e^{-(r+\lambda)T}\lambda V^T_{(\beta-1)x} dT \\ =\ & V_{x'} + e^{-(r+\lambda)t_{x'}}\frac{(x-x')}{r} + \int_0^{t_{x'}} e^{-(r+\lambda)T}\lambda V^T_{\beta x} dT + \int_{t_{x'}}^{\infty} e^{-(r+\lambda)T}\lambda V^T_{(\beta-1)x} dT. \end{aligned}$$

The incentive constraint of the agent x' preventing him to mimic x is a symmetric version of the above expression. Then by combining these two inequalities we get

$$\frac{e^{-(r+\lambda)t_x}}{r} + \frac{\int_{t_x}^{t_{x'}} e^{-(r+\lambda)T}\lambda(V^T_{\beta x'} - V^T_{(\beta-1)x'})dT}{x-x'} \geq \frac{V_x - V_{x'}}{x-x'} \geq \frac{e^{-(r+\lambda)t_{x'}}}{r} + \frac{\int_{t_x}^{t_{x'}} e^{-(r+\lambda)T}\lambda(V^T_{\beta x} - V^T_{(\beta-1)x})dT}{x-x'}.$$

First of all, we know t_x is monotone. Therefore it is continuous almost everywhere, and so when $x' \to x$, $\frac{e^{-(r+\lambda)t_{x'}}}{r} \to \frac{e^{-(r+\lambda)t_x}}{r}$, almost everywhere. Moreover, when $x' \to x$, by using Leibniz Rule, L'Hopital's Rule, almost everywhere continuity of t_x and incentive constraints of second step reports which we have proven in previous Lemma 2.3, we get the following :

$$\begin{aligned} lim_{x'\to x}\ \frac{\int_{t_x}^{t_{x'}} e^{-(r+\lambda)T}\lambda(V^T_{\beta x'} - V^T_{(\beta-1)x'})dT}{x-x'} &= lim_{x'\to x}\ \frac{\int_{t_x}^{t_{x'}} e^{-(r+\lambda)T}\lambda(V^T_{\beta x'} - V^T_{(\beta-1)x'})dT}{x-x'} \\ &= -\lambda\ e^{-(r+\lambda)t_x}\ \frac{\partial t_x}{\partial x}(V^{t_x}_{\beta x} - V^{t_x}_{(\beta-1)x})\ \ a.e \end{aligned}$$

[1] If $t_x = t_{x'}$ then we do not need to worry about first stage incentive constraints.

Therefore we can conclude that

$$\frac{\partial V_x}{\partial x} = \frac{e^{-(r+\lambda)t_x}}{r} - \lambda\, e^{-(r+\lambda)t_x}\, \frac{\partial t_x}{\partial x}(V_{\beta x}^{t_x} - V_{(\beta-1)x}^{t_x}) \quad a.e$$

Then by integrating it we get the result ii).

B.3 Proof of Proposition 2.1

For now, rather than the problem in (2.4), we consider an auxiliary problem in which the second term of the objective function is omitted. Precisely

$$\begin{aligned} \max_{\{t_x\}_{x\in[0,1]},\{\{t_x^T\}_{x\in[0,1]}\}_{T>0}} \quad & \frac{1}{r}\int_0^1 e^{-(r+\lambda)t_x}\Big(x - \frac{1-F(x)}{f(x)}\Big) f(x)dx \\ & + \frac{\lambda}{r}\int_0^\infty e^{-(r+\lambda)T}\Big(\int_0^1 e^{-rt_x^T}\Big(x - \frac{1-F_T(x)}{f_T(x)}\Big) f_T(x)dx\Big)dT \end{aligned}$$

$$\begin{aligned} \textit{subject to} \quad & \bullet\ t_x \textit{ is non-increasing in } x \\ & \bullet\ t_x^T \textit{ is non-increasing in } x, \quad \forall T \in [0,\infty) \end{aligned} \tag{B.4}$$

In this problem, the contingent allocation terms $\{t_x^T\}_{x\in[0,1]}$, for the second version of the good, appear only on the last term of the objective function. Then the optimal $\{t_x^T\}_{x\in[0,1]}$ for this problem will be similar to the one of the canonical model. Therefore it follows a cutoff rule, where the value of the cutoff is a function of the distribution $f_T(.)$. Hence we denote the cutoff value by $x^\star(f_T)$, and its value is exactly the same as the value of the cutoff for the static monopoly with distribution $f_T(.)$. So that for a given f_T the allocations are of the form:

$$t_x^T = \begin{cases} 0 & x \geq x^\star(f_T) \\ \infty & x < x^\star(f_T) \end{cases}$$

On the other hand, the first version allocations $\{t_x\}_{x\in[0,1]}$ are affecting both lines of the objective function as they alter the distribution functions $f_T(.)$ of the effective types. If this indirect effect did not exist, then the optimal allocation rule would be the immediate allocation for those agents having a type higher then $x^\star$ i.e the static monopoly allocation. Despite this additional effect, the optimal allocations have a similar structure to the one of the static monopoly in the sense that it also follows a cutoff rule. This is because of the stationary structure of the environment

Claim: There is a cutoff value $\hat{x}$, depending on the values of λ, α, r, such that the optimal solution of the program (B.4) satisfies:

$$t_x = \begin{cases} 0 & x \geq \hat{x} \\ \infty & x < \hat{x} \end{cases}$$

Proof of the Claim: This is due to the stationary structure resulting from the Poisson arrival process. In particular, if at $t \neq 0$ an agent is allocated the first version of the good then it must be the case that the total effect of allocating the first version to this agent on the objective function is positive. But then it must be positive t=0 as well since the environment is stationary. Therefore it is better for the monopolist to allocate the good to this agent at the beginning $t = 0$. Hence we know that the term t_x must be either 0 or ∞ for every x. Furthermore, since t_x is restricted to be monotone with respect to x, optimal solution must incorporate a structure as given above. $\diamond$

Then we have the solution of the problem (B.4), as:

$$t_x = \begin{cases} 0 & x \geq \hat{x} \\ \infty & x < \hat{x} \end{cases} \qquad \forall T, t_x^T = \begin{cases} 0 & x \geq x^\star(\hat{x}) \\ \infty & x < x^\star(\hat{x}) \end{cases}$$

Since the allocation of the first version only occurs at $t = 0$, the distribution of the

effective types is independent of the realized arrival time T, and just depending on the cutoff of the first version allocations. Moreover, the allocation of the second version is same with the static monopoly allocations corresponding to the effective type distribution.

Turning back to the original problem of the monopolist as defined in (2.4) we know that the second term, which is omitted in the relaxed problem, would be equal to zero under the allocation rule that is specified above. This is because of the fact that $\frac{\partial t_x}{\partial x} = 0$ almost everywhere. Furthermore, we also know that the highest possible value of this term is also zero, since t_x must be non-increasing and hence its derivative is never strictly positive. Therefore the solution of program (B.4), which is defined as above, is also the solution for the original problem (2.4) as it is maximizing the second term as well. Then we have $x_1 = \hat{x}$, and $x_2 = x^\star(\hat{x})$.

B.4 Proof of Lemma 2.5

To start with the first part suppose that the L-type buyers purchase both versions, hence $p_T < (\beta - 1)L$. Then the H-type buyers purchase the second version as well, since $p_T < (\beta - 1)H$. Moreover, since purchasing the first version conditional on purchasing the second version has a positive return for the L-type buyers, it must have a positive return for the H-type buyers as well hence H-type buyers also purchase the first version.

For the second part, suppose the L-type buyers purchase the first version at $t < t^\star$. Then, to prove the statement, we just need to show that the H-type buyers purchase the first version of the good, thanks to our observation given before. Suppose not to get a contradiction. Then it must be the case that the L-type buyers are only purchasing the first version of the good while H-type buyers are only purchasing the second version, because otherwise if the L-type buyers were purchasing the second version, then, from the first part of the lemma, the H-type buyers would purchase both versions of the good. Also, if a H-type buyers is not purchasing the second version, then it means that he is not purchasing any

versions of the good which would also be a contradiction. Then by the revealed preferences:

$$e^{-rt}(L - p_t) \geq e^{-rT}(\beta L - p_T)$$
$$e^{-rT}(\beta L - p_T) \geq e^{-rt}(H - p_t)$$

Which is a direct contradiction since $H > L$ and $e^{-rt} > e^{-rt^\star} = \beta e^{-rT}$.

B.5 Proof of Lemma 2.6

Showing that under the optimal price path there must be an agent of some type that is purchasing at $t = 0$ would be sufficient to prove this lemma due to the second part of Lemma 2.5.

Assume that nobody purchases at t=0 to get a contradiction. There must be a sale of the first version at some time before T, because otherwise, if there is a sale of only the second version good, we would get a contradiction immediately, as the monopolist could deviate and sell only the first version of the good at $t = 0$ to the agents that are purchasing the second version. This is better for the firm as it can get a higher discounted payment due to assumption 2.2.

Denote the earliest time period at which a sale of the first version occurs by $\underline{t}$. We want to show that it is equal to 0. Suppose $\underline{t} > 0$ to get a contradiction. Then there must be a sale of the second version of the good, because otherwise there exists an obvious profitable deviation, which is selling at $t = 0$ with the price level $p_{\underline{t}}$. If the agent purchasing the first version at $\underline{t}$ also purchases the second version, then it must be a H-type from lemma 2.5. Then the monopolist can be made better off by changing the price level at $t = 0$ so that the H-type is indifferent between purchasing at 0 and $\underline{t}$ as that does not alter the incentives of the L-type buyers. We get a similar contradiction for the other case in which the agent purchasing the first version at $\underline{t}$ is not purchasing the second version.

B.6 Proof of Proposition 2.2

To prove the proposition, we first treat the price of the second version of the good as a fixed value. We then find the corresponding optimal price path $\{p_t\}_{t\in[0,T)}$ of the first version of the good for any given value of p_T and we finally optimize p_T at the end. There are 5 cases to consider for p_T.

i) $\beta H < p_T$.

In this case there is no sale of the second version of the good. We know that in an optimal policy H-type buyers purchase the first version at t=0, and the maximum amount that they are willing to pay at t=0 is H. On the other hand, at any time $t > T$, the L-type buyers are willing to pay at most L (given that there is no sale of the second version). Therefore, if the monopolist is going to sell the first version of the good to L-type buyers at a time t, then he should arrange the price as $p_t = L$. However this will affect the incentives of the H-type buyers that are purchasing at $t = 0$. Given that $p_t = L$ for some t, the maximum amount that the H-type buyers are willing to pay at $t = 0$, which we denote by $\bar{p}$, satisfies

$$H - \bar{p} = e^{-rt}(H - L)$$
$$\bar{p} = (1 - e^{-rt})H + e^{-rt}L$$

And the corresponding profit of the monopolist, when L-type buyers purchase at time t, is

$$\Pi_t = \mu((1 - e^{-rt})H + e^{-rt}L) + (1 - \mu)e^{-rt}L. \tag{B.5}$$

Note that the expression above is linear in e^{-rt}, hence it is maximized either at $t = 0$ or at $t = \infty$[2]. If $t = 0$ is optimal, then both types purchase the good at $t = 0$ and the price level is equal to L. On the other hand, if $t = \infty$ is optimal, then only the H-type buyers

[2]It is also possible have that any $t > 0$ is a maximizer of the expression B.5. In such a case restricting t to be either 0 or ∞ is wlog.

purchase the good (at $t = 0$) at price H.

Therefore, for the first case there are two candidates of the optimal policy.

- A1: Sell the first version of the good to agents of both types at t=0 at a price level L and have no sales of the second version.
- A2: Sell the first version only to the H-types at t=0 at a price level H, and have no sales of the second version.

Actually, this is analogous to the result of Stokey (1979) and intuitively follows because if there is no sales of the second version, we turn back to canonical model.

ii) $(\beta - 1)H < p_T \leq \beta H$.

At the optimal policy, the marginal benefit from the second version of the good is $(\beta - 1)H$ for H-type buyers since they purchase the first version at $t = 0$. Therefore, there is no sale of the second version in this case as well. However, the situation is different than the previous case in the sense that now there is an additional option for H-type buyers. In other words, by purchasing only the second version of the good, they can guarantee a non-negative utility. As a result, the maximum amount that H-type buyers are willing to pay at t=0 is less than H. Denote this maximum price level by $\bar{p}$, which satisfies:

$$H - \bar{p} = e^{-rT}\beta H - e^{-rT}p_T.$$
$$\bar{p} = (1 - \beta e^{-rT})H + e^{-rT}p_T$$

Where the LHS of the first line is the utility from the purchase of the first version of the good, and the RHS is from the purchase of the second version. Similar to the previous case, a candidate optimal policy is selling the first version good at t=0, to both types of agents at a price L[3]. However, the corresponding policy would be equivalent to A1, so that we do

[3]Note that L is less than $\bar{p}$, since $p_T \leq \beta H$, $\beta e^{-rT} < 1$, and $(\beta - 1)H > \beta L$. Therefore H-type buyers are willing to buy at this price level.

not write it again here.

We can think of another candidate, which is a modified version of the policy A2, that is selling version 1 at t=0 only to the H-type buyers but now with a payment $\bar{p}$, rather than H. However, this policy is strictly dominated since the payment is less than the one of A2.

Finally, in policy A2, there is no sale of the first version to the L-types, which is due to the fact that the monopolist needs to decrease the price level at $t = 0$ (which was equal to H) to be able to sell to the L-type buyers at any time. However, in this case, by departing from the case 1, there may exists a time period earlier than T, at which selling the good to the L-type buyers at the maximum price that they are willing to pay (which is equal to L) does not hurt the incentives of the H-types. Hence it does not require to decrease the price at t=0[4], because, now H-type buyers are having a positive utility from the first version purchase at price $\bar{p}$. Nevertheless, even if such a period of time exists, doing any better than both of the policies A1 and A2 is not possible due to the fact that the resulting policy would be equivalent to one of the intermediate policies in the expression B.5. Therefore, the firm cannot do any better here in this case.

iii) $\beta L < p_T \leq (\beta - 1)H$.

In this case, H-type buyers purchase the second version of the good, since p_T is always smaller then their marginal benefit and L-type buyers do not purchase. The maximum amount that H-type buyers are willing to pay at t=0 for the first version of the good, $\bar{p}$, in this case satisfies the following:

$$H - \bar{p} + e^{-rT}((\beta - 1)H - p_T) = e^{-rT}(\beta H - p_T)$$
$$\bar{p} = (1 - e^{-rT})H$$

where the LHS of the first line is the utility of the H-type buyers from purchasing both

[4] We can find such a time period by finding a t so that the H-type buyers are indifferent between purchasing at t at price L and purchasing at time 0 with price $\bar{p}$. Time period t satisfying this indifference condition should be less than T.

versions of the good, and the RHS is the utility from purchasing only the second version of the good. Note that $\bar{p}$ is higher than L.[5] Hence the L-type buyers are not willing to purchase at t=0 with price $\bar{p}$. Then the monopolist should either set a smaller price than $\bar{p}$ at $t = 0$ to be able to sell the L-type buyers at $t = 0$, or he can sell it at a later time to them. Note that, at any time t, the highest amount that L-type buyers are willing to pay for the first version of the good is L as they do not purchase the second version of the good in this case. Like in the previous cases, charging a price level that is equal to L will affect the incentives of the H-type buyers.

One possible optimal policy here is to sell the first version of the good at t=0 to both types of buyers at a price L, and set $p_T = (\beta - 1)H$. Another possibility is to sell the first version of the good to the H-type buyers at a price $\bar{p}$ at $t = 0$, and to the L-type buyers at a later time (before T) without hurting the incentives of the H-type buyers. As we discussed in case 2, finding such a time period is possible here, because H-type buyers are having a positive utility from purchase of the first version of the good. To find this time period, let's denote p_t^H as the price level, which leaves the H-type buyers indifferent between purchasing good at t=0 with payment $\bar{p}$, and purchasing at t with payment p_t^H. In particular

$$\begin{aligned} H - \bar{p} &= e^{-rt}(H - p_t^H) \\ p_t^H &= \frac{(e^{-rt} - e^{-rT})H}{e^{-rt}} \end{aligned}$$

Then we can find the earliest possible time period $\bar{t}$ at which the firm can sell the first version of the good to the L-type buyers at price L without hurting the incentives of the H-type by using the equality $p_t^H = L$. In particular,

$$\begin{aligned} (e^{-r\bar{t}} - e^{-rT})H &= e^{-r\bar{t}}L \\ e^{-r\bar{t}} &= e^{-rT}\frac{H}{H - L} \end{aligned}$$

[5] This is because $\beta e^{-rT} < 1$, and $(\beta - 1)H > \beta L$.

Note that $\bar{t} < T$ is always satisfied due to assumption 2.2 and $\beta L < (\beta - 1)H$. Hence selling the first version of the good to the L-type buyers at $\bar{t}$ at a price L is a feasible policy. To sum up, we have the following two candidates for this case.

- A3: Sell the first version to both type of buyers at t=0 at price L, and sell the second version to only to the H-type buyers at a price $(\beta - 1)H$.[6]

- A4: Sell the first version of the good to the H-type buyers at t=0 with price $\bar{p} = (1 - e^{-rT})H$, and also sell to L-type buyers at $t = \bar{t}$ with payment $p_{\bar{t}}^H = L$. Sell the second version to H-type buyers at price $(\beta - 1)H$.

Note that, in this case there can not be any better policy then these two due to the linearity of the profit function as we have discussed in case 1. For instance, take the policy A4, if it is better to decrease the price at t=0 to sell to the L-type buyers earlier than $\bar{t}$, then the monopolist should continue to decrease price level at 0 until it reaches L at which L-type is willing to buy; and this corresponds to the policy A3.

iv) $(\beta - 1)L < p_T \leq \beta L$.

In this case, the H-type buyers always purchase the second version of the good while L-type buyers purchase the second version only if they have not purchased the first one. We can easily see that the maximum amount that the H-type buyers are willing to pay at $t = 0$ for the first version of the good is same as in case 3 and so it is equal to $\bar{p} = (1 - e^{-rT})H$.

There are three candidates for the optimal policy in this case. The first one is to sell the first version of the good to both types of buyers at t=0 at a price level that leaves the L-type buyers indifferent between purchasing only the first version and purchasing only the second version, and to sell the second version only to the H-type buyers. However, this policy is strictly dominated by A3. In particular, the maximum amount of the payment that L-type buyers are willing to pay at $t = 0$ is less than L, and the amount charged for the second version of the good at time T is strictly less than the one of A3. The second policy, is to sell

[6]Note that this policy is strictly dominating the policy A1.

the first version of the good to the H-type buyers at t=0 with price $\bar{p} = (1 - e^{-rT})H$, and to the L-type buyers at a later time, and to sell the second version only to the H-type buyers. This is dominated by the policy A4 for the same reason above. Then the final candidate is:

- A5: Sell the first version of the good only to the H-type buyers at t=0 at the price $\bar{p} = (1 - e^{-rT})H$, and sell the second version of the good to both at the price βL.

v) $p_T \leq (\beta - 1)L$.

In this case, both types of the buyers purchase the second version of the good regardless of their decision on the first version sales. From the same reasoning as above the maximum amount that the H type buyers are willing to pay at t=0 is $\bar{p} = (1 - e^{rT})H$, and he is indifferent between purchasing the first version at $t = 0$ with payment $\bar{p}$ and purchasing at t with payment $p_t^H = \frac{(e^{-rt} - e^{-rT})H}{e^{-rt}}$. Similarly, the maximum amount that the L-type buyers can are willing to pay for the first version of good at time t is $p_t^L = \frac{(e^{-rt} - e^{-rT})L}{e^{-rt}}$.

Since $p_t^L < p_t^H$, there does not exist a time period in which the monopolist can sell the first version of the good tothe L-type buyers at price p_t^L without hurting the incentives of the H-type buyers when they are purchasing at $t = 0$ with price $\bar{p}$. Therefore the monopolist must decrease the initial price to be able to sell the L-type agents at any time. Then, again due to the linearity of the monopolist profit, there are two possibilities for the optimal policy in this case. However one of them, which is selling the first version of the good only to the H-type buyers with price $\bar{p}$ and selling the second version of the good to both types with a price $(\beta - 1)L$ is strictly dominated by the policy A5. Therefore the only option that we are left with is:

- A6: Sell the first version of the good to both type of buyers at t=0 at a price $(1 - e^{-rT})L$, and sell the second version of the good to both types of buyers at a price $(\beta - 1)L$.

We have considered all of the possible optimal policies for the monopolist. First note

that the policy A1 is strictly dominated by the policy A3, (hence we omit A1). Then the corresponding profit level for each are listed below as follows:

Table B.1: Sale Policies and Corresponding Profits.

Policy	Profit
A2	μH
A3	$L + \mu e^{-rT}(\beta - 1)H$
A4	$\mu[1 + e^{-rT}(\beta - 2)]H + (1 - \mu)e^{-rT}\frac{LH}{H-L}$
A5	$\mu(1 - e^{-rT})H + e^{-rT}\beta L$
A6	$[1 + e^{-rT}(\beta - 2)]L$

Each of these 5 policies may be the optimal one depending on the values of β, μ, H, L and r. All the policies except $A4$ involves immediate allocations like in the stochastic arrival case. Hence time is not used to discriminate over people in those policies. On the contrary, in policy $A4$ the price of the first version of the good is decreasing over time. As a result, purchase times of agents for the first version of the good are different for L and H-types of buyers. More precisely, the H-type buyers purchase at the beginning, whereas the L-type buyers purchase at a later time (before T). The reason for such a pattern is based on the anonymous structure of the posted prices for the second version sales. The existence of the second version puts a restriction on the amount that a H-type buyer is willing to pay for the first version of the good since it is possible for him to give up from the purchase of the first version of the good and purchase only the second one. As a result there exists a time period so that the monopolist can sell the first version of the good to the L-type buyers without hurting the incentives of the H-type buyers. This is not possible in the canonical durable good monopoly model.

Appendix C

Appendix to Chapter 3

C.1 Proof of Proposition 3.2

One can rewrite the incentive constraint $\mathcal{IC}_{\mathcal{S}}$ as follows:

$$(1-\sigma)(p_1-p_0)[p_1(w_{11}-w_{01})+(1-p_1)(w_{10}-w_{00})] \geq c$$

Then, it is straightforward to see that $\mathbf{w}^s$ solves the problem $\mathcal{P}_S$. Moreover, when the incentive scheme is $\mathbf{w}^s$, (work, work) is a Nash equilibrium since $\mathcal{IC}_{\mathcal{S}}$ is satisfied under $\mathbf{w}^s$. Next, we argue that (work, work) is indeed the unique Nash equilibrium, hence comprises a team equilibrium as well. Suppose that the agent j chooses high level effort with probability e_j. Then, under the incentive scheme $\mathbf{w}^s$, the expected payment of agent i from working and shirking are given by:

$$\pi(1, e_j, \mathbf{w}^s) = (1-\sigma)p_1[e_j(1-p_1)+(1-e_j)(1-p_0)]w^s_{10},$$
$$\pi(0, e_j, \mathbf{w}^s) = (1-\sigma)p_0[e_j(1-p_1)+(1-e_j)(1-p_0)]w^s_{10},$$

respectively. After plugging the value of w_{10}^s, the difference between these expected payments, which we denote by $\Delta\pi(e_j, \mathbf{w}^s)$, becomes:

$$\Delta\pi(e_j, \mathbf{w}^s) = [e_j(1-p_1) + (1-e_j)(1-p_0)]\frac{c}{(1-p_1)}.$$

Note that, $\Delta\pi(e_j, \mathbf{w}^s) > c$ as long as $e_j < 1$. Therefore, (work, work) is the unique Nash equilibrium, and also the team equilibrium. Therefore $\mathbf{w}^s$ is the optimal incentive scheme in the static setting.

C.2 Proof of Lemma 3.2

We examine the problem $\mathcal{P}_{\mathcal{D}}$ by considering it as a nested optimization problem. More precisely, we divide the entire space of incentive schemes into two sub-spaces, depending on the comparison between the values of w_{11}, and w_{10}. Then we characterize the solution of $\mathcal{P}_D$ in these two sub-spaces separately. Getting the general solution is then just a matter of a simple comparison between these two local solutions.

First, we analyze the case in which $w_{11} \leq w_{10}$. The incentive constraint $\mathcal{IC}_{\mathcal{D}}$ boils down the incentive constraint of the static setting, i.e., $\mathcal{IC}_{\mathcal{S}}$. Therefore, under the restriction $w_{11} \leq w_{10}$, $\mathcal{P}_{\mathcal{D}}$ is equivalent to $\mathcal{P}_{\mathcal{S}}$, hence $\mathbf{w}^s$ is the solution.

On the other hand, when $w_{11} \geq w_{10}$, the incentive scheme is a JPE, since $w_{01} = w_{00} = 0$, and $\mathcal{IC}_{\mathcal{D}}$ is equivalent to:

$$\pi(1,1,\mathbf{w}) - c \geq (1-\delta)\pi(0,1,\mathbf{w}) + \delta\pi(0,0,\mathbf{w}). \tag{$\mathcal{IC}_{\mathcal{J}}$}$$

Therefore, one can write the conditional problem in this case as follows:

$$\min_{\mathbf{w}} \ \pi(1,1,\mathbf{w}) \ \text{ s.t. } \ \mathcal{IC}_{\mathcal{J}}, \ w_{11} \geq w_{10}. \tag{$\mathcal{P}_{\mathcal{J}}$}$$

Due to the linearity of the objective function and the constraints, the optimality requires either to increase or to decrease w_{11} as much as possible. If latter is the case, then we must have $w_{11} = w_{10}$, due to the condition $w_{11} \geq w_{10}$. Moreover we know that the incentive constraint $\mathcal{IC}_{\mathcal{J}}$ can be rewritten as:

$$(p_1 + \delta p_0)w_{11} + [1 - (p_1 + \delta p_0)]w_{10} \geq \hat{c}$$

As a result, the solution to this auxiliary problem must be either $\mathbf{w}^j = (\frac{\hat{c}}{p_1+\delta p_0}, 0, 0, 0)$, or $\mathbf{w}^I = (\hat{c}, \hat{c}, 0, 0)$, where $\hat{c} = \frac{c}{(1-\sigma)(p_1-p_0)}$.

Finally, in order to get the general solution of the problem $\mathcal{P}_{\mathcal{D}}$, one just needs to compare the expected costs of the incentive schemes $\mathbf{w}^j$, $\mathbf{w}^s$, and $\mathbf{w}^I$. The statement follows from this comparison immediately.

C.3 Proof of Lemma 3.3

The problem $\mathcal{P}_{\mathcal{R}}$ is defined as follow:

$$\min_{w_{11}, w_{10}} \quad [\sigma + (1-\sigma)p_1]w_{11} + [(1-\sigma)p_1(1-p_1)]w_{10} \quad \text{s.t.} \quad \mathcal{IC}_{\mathcal{S}},\ \widehat{\mathcal{IC}_{\mathcal{R}}},\ w_{11} \leq w_{10}.$$

In order to solve this problem, we need to have a precise comprehension of its constraints. First note that, we can rewrite $\mathcal{IC}_{\mathcal{S}}$, $\mathcal{IC}_{\mathcal{R}}$, and $\mathcal{IC}''_{\mathcal{R}}$ respectively as follows:

$$w_{10} + p_1(w_{11} - w_{10}) \geq \hat{c} \qquad (\mathcal{IC}_{\mathcal{S}})$$

$$w_{10} + (1+\delta)p_1(w_{11} - w_{10}) > \hat{c} \qquad (\mathcal{IC}_{\mathcal{R}})$$

$$w_{10} + 2p_1(w_{11} - w_{10}) > \hat{c} \qquad (\mathcal{IC}''_{\mathcal{R}})$$

The constraint $\mathcal{IC}'_{\mathcal{R}}$, on the other hand, is given by:

$$\left((\delta+\delta^2)p_1-(1-\delta^2)p_0\right)w_{11}+\left((\delta+\delta^2-1)-(\delta+\delta^2)p_1+(1-\delta^2)p_0\right)w_{10}>(\delta+\delta^2-1)\hat{c}.$$

As we can see from this expression, the structure of $\mathcal{IC}'_{\mathcal{R}}$ crucially depends on the value of δ. In particular, when $\delta+\delta^2=1$, this constraint boils down to $w_{11}-w_{10}>0$, which can never be satisfied since the problem is defined conditional on $w_{11}\leq w_{10}$. In such a circumstance, to satisfy the constraint $\widehat{\mathcal{IC}_{\mathcal{R}}}$, the incentive scheme has to satisfy either $\mathcal{IC}_{\mathcal{R}}$, or $\mathcal{IC}''_{\mathcal{R}}$. When $\delta+\delta^2\neq 1$, we can express $\mathcal{IC}'_{\mathcal{R}}$ more explicitly as follows:

$$\mathcal{IC}'_{\mathcal{R}}=\begin{cases} w_{10}+\left(\frac{1-\delta^2}{1-\delta-\delta^2}p_0-\frac{\delta+\delta^2}{1-\delta-\delta^2}p_1\right)(w_{11}-w_{10})<\hat{c} & \text{if } \delta+\delta^2<1 \\ w_{10}+\left(\frac{\delta+\delta^2}{\delta+\delta^2-1}p_1-\frac{1-\delta^2}{\delta+\delta^2-1}p_0\right)(w_{11}-w_{10})>\hat{c} & \text{if } \delta+\delta^2>1 \end{cases}$$

The focus is on the incentive schemes satisfying $w_{11}\leq w_{10}$. Thus, the optimal values of w_{11}, and w_{10} must belong to the dashed region of the figure C.1.

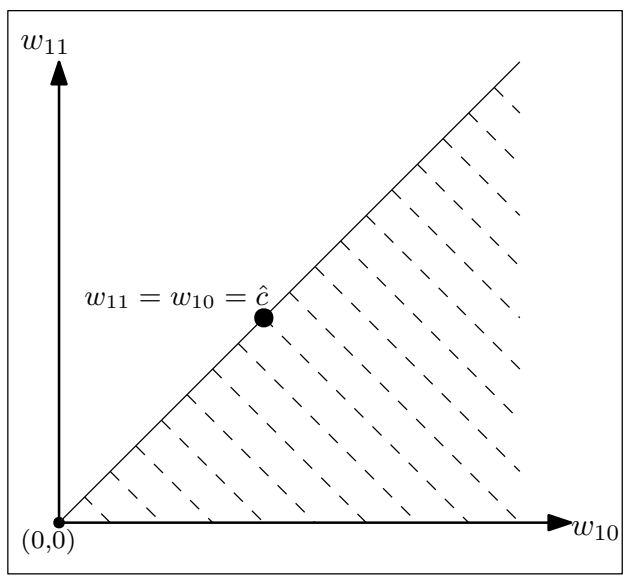

Figure C.1: The region satisfying $w_{11}<w_{10}$, together with $w_{11}\geq 0$, and $w_{10}\geq 0$.

The linear lines describing the constraints $\mathcal{IC}_{\mathcal{S}}$, $\mathcal{IC}_{\mathcal{R}}$, $\mathcal{IC}'_{\mathcal{R}}$, $\mathcal{IC}''_{\mathcal{R}}$ pass through the point $w_{11} = w_{10} = \hat{c}$. Moreover, these lines never intersect again as they have different slopes. As a result, there are two possibilities for the optimal solution: i) Increase w_{11} as much as possible and set it equal to w_{10}. ii) Decrease w_{11} as much as possible and set it equal to 0. If precedent is the case, then we have $w_{11} = w_{10} = \hat{c}$. The corresponding incentive scheme is an IPE, and the agents' payments are independent from the performance measure of each other. Such a scheme is obviously collusion-proof, and hence satisfies the constraints.[1]

The analysis is more tricky when the optimality requires to decrease $w_{11} = 0$. It turns out that, $\mathcal{IC}_{\mathcal{R}}$ is the essential constraint to satisfy for fulfilling the comprehensive constraint $\widehat{\mathcal{IC}_{\mathcal{R}}}$. In other words, $\mathcal{IC}'_{\mathcal{R}}$, and $\mathcal{IC}''_{\mathcal{R}}$ are redundant. The redundancy of $\mathcal{IC}''_{\mathcal{R}}$ is rather obvious, because it is always harder to satisfy in comparison to $\mathcal{IC}_{\mathcal{R}}$ since $2p_1 > (1-\delta)p_1$. To see the redundancy of $\mathcal{IC}'_{\mathcal{R}}$, we consider two cases separately. When $\delta + \delta^2 > 1$, $\mathcal{IC}'_{\mathcal{R}}$ is always harder to satisfy in comparison to $\mathcal{IC}_{\mathcal{R}}$, because $\left(\frac{\delta+\delta^2}{\delta+\delta^2-1}p_1 - \frac{1-\delta^2}{\delta+\delta^2-1}p_0\right) > (1+\delta)p_1$ in this case. On the other hand, when $\delta + \delta^2 < 1$, the direction of the inequality is reversed. In this case, if $\mathcal{IC}'_{\mathcal{R}}$ was the relevant constraint for fulfilling $\widehat{\mathcal{IC}_{\mathcal{R}}}$, then the incentive scheme would satisfy $\mathcal{IC}_{\mathcal{S}}$, and $\mathcal{IC}'_{\mathcal{R}}$ at the same time. But this is impossible as we have $w_{11} = 0$. This stems from the fact that the constraint $\mathcal{IC}'_{\mathcal{R}}$, which is equivalent to $\left(1 - \frac{1-\delta^2}{1-\delta-\delta^2}p_0 + \frac{\delta+\delta^2}{1-\delta-\delta^2}p_1\right) w_{10} < \hat{c}$ in this case, is always violated, even for the smallest value of w_{10} satisfying the constraint $\mathcal{IC}_{\mathcal{S}}$, i.e., $\frac{c}{1-p_1}$, regardless the value of δ.

To sum up, when it is optimal to set $w_{11} = 0$, the value of w_{10} must be the minimum amount fulfilling $\mathcal{IC}_{\mathcal{R}}$, and $\mathcal{IC}_{\mathcal{S}}$. Therefore, w_{10} is equal to $\frac{\hat{c}}{1-(1+\delta)p_1}$.[2]

Now we would like to show that, when the incentive scheme is restricted to satisfy $w_{11} \leq w_{10}$, the solution of the problem $\mathcal{P}_{\mathcal{R}}$ coincides with the solution of the principal's problem. We already argued that the incentive scheme $\mathbf{w}^I$ is free from collusion, and induces (work,work)$^\infty$ as a team equilibrium. This suggests that whenever $\mathbf{w}^I$ solves $\mathcal{P}_{\mathcal{R}}$, it also

[1] The constraint $\widehat{\mathcal{IC}_{\mathcal{R}}}$ is satisfied as $\mathcal{IC}''_{\mathcal{R}}$ is satisfied.

[2] Since the constraint $\mathcal{IC}_{\mathcal{R}}$ is based on a strict inequality, the actual optimal value of w_{10} is infinitesimally larger than $\frac{\hat{c}}{1-(1+\delta)p_1}$.

solves the principal's problem conditional on $w_{11} \leq w_{10}$.

In order to complete the proof, one needs to show that, $\mathbf{w}^r$ is also free from collusion.[3] To this end, we argue that, (work,work)$^\infty$ is the unique sub-game perfect equilibrium when the incentive scheme is $\mathbf{w}^r$; hence it is also the unique team equilibrium.

First step: There is no loss of generality by restricting attention into the strategy profiles in which the agents condition their effort choices only to the realized effort decisions in the previous periods. Generally speaking, the agents can condition their effort choices on the realized performance signals as well. However, for a given sub-game perfect equilibrium of this sort, one can find another sub-game perfect equilibrium in which the strategy profile is just a function of the effort decisions of the earlier periods. This stems from the fact that, the agents can use a public randomization device to coordinate over all the possible probability distributions over the action profiles that they can reach by conditioning their efforts on the realized signals. Moreover, modifying a given strategy profile by using such public randomization would not alter the incentives as the agents only care about their continuation utilities while choosing their actions.

Second step: When the incentive scheme is $\mathbf{w}^r$, the agents never shirk together in a sub-game perfect equilibrium. Note that, in order to convince an agent to stick with (shirk, shirk) at some period, his continuation utility must be at least equal to U_{00}, which is defined by:

$$(1-\delta)\pi(0,0,\mathbf{w}^r) + \delta U_{00} = (1-\delta)\pi(1,0,\mathbf{w}^r) + \delta\pi(1,1,\mathbf{w}^r).$$

However, when the incentive scheme is $\mathbf{w}^r$, there does not exist a strategy profile delivering an expected utility that is higher than or equal to U_{00} to both agents at the same time. To see this, consider a strategy profile, and let $\lambda_{e_1 e_2}$ be the discounted probability that the effort pair (e_1, e_2) being chosen following this initial period at which the agents play (shirk, shirk). For instance, if the agents alternate between (work, shirk), and (shirk, work)

[3] Throughout the entire proof, we assume that $[1-(1+\delta)p_1] > 0$, so that w_{10}^r is well defined. Otherwise, $\mathbf{w}^I$ would always be the optimal solution of $\mathcal{P}_{\mathcal{R}}$, and the statement would be trivially correct.

afterwards, then we have $\lambda_{11} = 0$, $\lambda_{10} = \frac{1}{1+\delta}$, $\lambda_{01} = \frac{\delta}{1+\delta}$, and $\lambda_{00} = 0$. It is crucial to note that, $\lambda_{11} + \lambda_{10} + \lambda_{01} + \lambda_{00} = 1$, since all the terms are multiplied with $1 - \delta$. Then the expected continuation utilities of the agents satisfy:

$$U_1 = \lambda_{11}[\pi(1,1,\mathbf{w}^r) - c] + \lambda_{10}[\pi(1,0,\mathbf{w}^r) - c] + \lambda_{01}\pi(0,1,\mathbf{w}^r) + \lambda_{00}\pi(0,0,\mathbf{w}^r)$$
$$U_2 = \lambda_{11}[\pi(1,1,\mathbf{w}^r) - c] + \lambda_{01}[\pi(1,0,\mathbf{w}^r) - c] + \lambda_{10}\pi(0,1,\mathbf{w}^r) + \lambda_{00}\pi(0,0,\mathbf{w}^r)$$

By using the fact that, $\lambda_{00} = 1 - \lambda_{11} - \lambda_{10} - \lambda_{01}$, and plugging in the values, one can get:

$$U_1+U_2-2U_{00} = \frac{2c}{(p_1-p_0)(1-(1+\delta)p_1)}\Big[\lambda_{11}(\delta p_1-p_0)+(\lambda_{10}+\lambda_{01})[(1+\delta)p_1-2p_0]-\frac{2(p_1-p_0)}{\delta}\Big].$$

The value of $U_1 + U_2 - 2U_{00}$ is strictly negative regardless of the choices of λ_{11}, λ_{10}, and λ_{01}. Therefore, there does not exist a strategy profile delivering a continuation utility that is higher than U_s to both agents at the same time, hence they never shirk together when the incentive scheme is $\mathbf{w}^r$.

Third step: Unilateral shirking can not be supported in a sub-game perfect equilibrium either, when the incentive scheme is $\mathbf{w}^r$. To prove this, one needs to show that the action pair (work, shirk) never appears in an equilibrium. From the incentive constraint $\mathcal{IC}_{\mathcal{R}}$, we know that in order to convince the first agent to stick with (shirk, work) when the incentive scheme is $\mathbf{w}^r$, his continuation utility must be set larger than U_{01}, where:

$$U_{01} = \frac{1}{1+\delta}[\pi(1,0,\mathbf{w}^r) - c] + \frac{\delta}{1+\delta}[\pi(0,1,\mathbf{w}^r)].$$

To complete this step of the proof, we will mainly argue that, while providing an expected utility that is larger than U_{01} to an agent, it is impossible to not to violate the other agent's incentive constraints. First notice that, by the the definition of $\mathbf{w}^r$, we know that:

$$\pi(1,1,\mathbf{w}^r) - c = \frac{\delta}{1+\delta}[\pi(1,0,\mathbf{w}^r) - c] + \frac{1}{1+\delta}[\pi(0,1,\mathbf{w}^r)].^4$$

[4]Since there is an ϵ term in the value of $\mathbf{w}^r$, there is an arbitrarily small difference between the values of

Then by using the facts $\lambda_{00} = 0$, and $\lambda_{11} = 1 - \lambda_{10} - \lambda_{01}$, we can get:

$$U_1 = \frac{\delta + \lambda_{10} - \delta\lambda_{01}}{1+\delta}[\pi(1,0,\mathbf{w}^r) - c] + \frac{1 - \lambda_{10} + \delta\lambda_{01}}{1+\delta}\pi(0,1,\mathbf{w}^r)$$
$$U_2 = \frac{\delta + \lambda_{01} - \delta\lambda_{10}}{1+\delta}[\pi(1,0,\mathbf{w}^r) - c] + \frac{1 - \lambda_{01} + \delta\lambda_{10}}{1+\delta}\pi(0,1,\mathbf{w}^r)$$

Thus, in order to keep the first agent's expected utility higher than U_{01}, we need to have $\lambda_{10} - \delta\lambda_{01} > 1 - \delta$. As a result, if an agent's utility is larger than U_{01}, then his teammate's payoff must be smaller than $\pi(1,1,\mathbf{w}^r) - c$. This stems from the fact that the maximum value of U_2, conditional on $\lambda_{10} - \delta\lambda_{01} \geq 1 - \delta$, can be achieved when $\lambda_{10} = \frac{1}{1+\delta}$, $\lambda(0,1) = \frac{\delta}{1+\delta}$, and $\lambda_{11} = 0$.

Suppose that the agents play (shirk, work) at period 1 without loss of any generality. Let $\lambda^2_{e_1,e_2}$ be the corresponding probability of agents playing (e_1, e_2) at $t = 2$. Similarly, let $\lambda^t_{e_1,e_2}$, be the probability of agents playing (e_1, e_2) at time t conditional on having played (shirk, work) in all the previous periods. If probability of always playing (shirk, work) at periods $1, 2, ..., t-1$ is 0, then $\lambda^t_{e_1,e_2}$ is equal to 0 as well.

Now, we will show that there does not exist a continuation play supporting (shirk, work) as a sub-game perfect equilibrium. First, there is no loss of generality to assume that, once the agents play (work, work) in a period, they continue to work together afterwards. By modifying a given strategy profile, without altering the incentives, one can reach to another strategy profile that satisfies this condition. Moreover, from the previous paragraph, we know that, once (work, shirk) is played, the first agent's continuation utility must be lower than $\pi(1,1,\mathbf{w}^r) - c$. Then by putting all these together, we know that the first agent's

right and left sides of the equlity. However, this does not effect our arguments.

continuation utility starting from $t = 2$ satisfies:

$$U_1 < \sum_{t=2}^{\infty} \Lambda_{01}^{t-1}\delta^{t-2}\lambda_{11}^{t}[\pi(1,1,\mathbf{w}^r) - c] + \sum_{t=2}^{\infty} \Lambda_{01}^{t-1}\delta^{t-2}\lambda_{01}^{t}(1-\delta)\pi(0,1,\mathbf{w}^r)$$
$$+ \sum_{t=2}^{\infty} \Lambda_{01}^{t-1}\delta^{t-2}\lambda_{10}^{t}[(1-\delta)(\pi(1,0,\mathbf{w}^r) - c) + \delta(\pi(1,1,\mathbf{w}^r) - c)],$$

where $\Lambda_{01}^{t} = \prod_{k=1}^{t} \lambda_{01}^{k}$ is the probability of having $(0,1)^t$ on the path of play starting from the first period.[5] Then by using the facts, $\lambda_{10}^{t} = 1 - \lambda_{11}^{t} - \lambda_{01}^{t}$, and $\pi(1,1,\mathbf{w}^r) - c = \frac{\delta}{1+\delta}[\pi(1,0,\mathbf{w}^r) - c] + \frac{1}{1+\delta}[\pi(0,1,\mathbf{w}^r)]$, one can rewrite the above inequality as follows:

$$U_1 < \sum_{t=2}^{\infty} \Lambda_{01}^{t-1}\delta^{t-2}\left(\frac{1}{1+\delta} - \frac{1-\delta}{1+\delta}\lambda_{11}^{t} - \frac{1}{1+\delta}\lambda_{01}^{t}\right)[\pi(1,0,\mathbf{w}^r) - c]$$
$$\sum_{t=2}^{\infty} \Lambda_{01}^{t-1}\delta^{t-2}\left(\frac{\delta}{1+\delta} + \frac{1-\delta}{1+\delta}\lambda_{11}^{t} + \frac{1-\delta^2-\delta}{1+\delta}\lambda_{01}^{t}\right)\pi(0,1,\mathbf{w}^r)$$

However, we know that $\Lambda_{01}^{t} = \Lambda_{01}^{t-1}\lambda_{01}^{t}$, therefore:

$$U_1 < \left(\frac{1}{1+\delta} - \frac{1-\delta}{1+\delta}\lambda_{11}^{2} - \sum_{t=3}^{\infty} \Lambda_{01}^{t-1}\delta^{t-2}\left(\frac{1-\delta}{\delta(1+\delta)} + \frac{1-\delta}{1+\delta}\lambda_{11}^{t}\right)\right)[\pi(1,0,\mathbf{w}^r) - c]$$
$$+ \left(\frac{\delta}{1+\delta} + \frac{1-\delta}{1+\delta}\lambda_{11}^{2} + \sum_{t=3}^{\infty} \Lambda_{01}^{t-1}\delta^{t-2}\left(\frac{1-\delta}{\delta(1+\delta)} + \frac{1-\delta}{1+\delta}\lambda_{11}^{t}\right)\right)\pi(0,1,\mathbf{w}^r)$$

The right side of the inequality is maximized by setting $\lambda_{11}^{2} = \lambda_{01}^{2} = 0$, and is equal to U_{01}.[6] Therefore, it is not possible to convince the first agent to stick with (shirk, work) at $t = 1$, while respecting the other agent's incentives for the future periods. To sum up, (work,work)$^{\infty}$ is the unique sub-game perfect equilibrium, hence the team equilibrium, when the incentive scheme is $\mathbf{w}^r$.

[5] Note that $\Lambda_{01}^{1} = 1$.

[6] Since $\lambda_{01}^{2} = 0$, $\Lambda_{01}^{t} = 0$ for each $t \geq 2$.

C.4 Proof of Proposition 3.4

This immediately follows from the comparison between the optimal production plans for human-human and human-machine production teams.

C.5 Proof of Lemma 3.1

So far, all the proofs are based on the assumption that $w_{01} = w_{00} = 0$.[7] Here, we will show that this assumption is correct. The proof is based on showing that all the "corresponding problems" that one can define without $w_{00} = w_{01} = 0$ boil down to their counterparts assuming $w_{00} = w_{01} = 0$. While defining these "corresponding problems", we take all the consequences of not having $w_{00} = w_{01} = 0$ into account.

We examine the principal's problem by means of two conditional problems as before. More precisely, we divide the entire space of incentive schemes into two, the ones satisfying $w_{11} \geq w_{10}$, and $w_{11} \leq w_{10}$ respectively. We focus on characterizing the optimal scheme in these two sub-spaces separately. We then show that it is optimal to set $w_{01} = w_{00} = 0$ in both cases.

First step: We, first, limit our attention into the incentive schemes satisfying $w_{11} \geq w_{10}$. The arguments here basically follow the same steps with the proof of lemma 3.2 without having any restriction on the values of w_{01}, and w_{00}. We have the following problem:

$$\min_{\mathbf{w}} \ (\sigma + (1-\sigma)p_1^2)w_{11} + (1-\sigma)p_1(1-p_1)(w_{10} + w_{01}) + (1-\sigma)(1-p_1)^2 w_{00}$$
$$\text{s.t.} \quad \mathcal{IC}_{\mathcal{D}}, \ w_{11} \geq w_{10}.$$

In the absence of the restrictions imposed on the values of w_{01}, and w_{00}, the condition

[7] Note that the proof of this lemma is presented after the proofs of lemma 3.2, and lemma 3.3. The reason to follow this expositional order is due to the fact that this proof uses the notation and the definitions (of problems and constraints) that are described after lemma 3.1 in the main text.

$w_{11} \geq w_{10}$ does not directly imply that the incentive scheme is a JPE. In this regard, $\mathcal{IC}_{\mathcal{D}}$ may be equivalent to $\mathcal{IC}_{\mathcal{S}}$, or $\mathcal{IC}_{\mathcal{J}}$.

$$p_1 w_{11} + (1-p_1)w_{10} - p_1 w_{01} - (1-p_1)w_{00} \geq \hat{c}, \tag{$\mathcal{IC}_{\mathcal{S}}$}$$

$$(p_1 + \delta p_0)w_{11} + (1 - p_1 - \delta p_0)w_{10} + (\delta - p_1 - \delta p_0)w_{01} - (1 + \delta - p_1 - \delta p_0)w_{00} \geq \hat{c}. \tag{$\mathcal{IC}_{\mathcal{J}}$}$$

Obviously, the solution to above problem requires to set $w_{00} = 0$, since the coefficients of w_{00} are negative in the constraint functions, and positive in the objective function. On the other hand, the variables w_{10}, and w_{01} have the same coefficients in the objective function, while the latter has a smaller coefficient in the constraints. Thus, it is also optimal to set $w_{00} = 0$. As a result, conditional on the incentive scheme satisfying $w_{11} \geq w_{10}$, it is always optimal to set $w_{01} = w_{00} = 0$.

<u>Second step:</u> Now, we limit our attention into the set of incentive schemes satisfying $w_{11} \leq w_{10}$. We focus on a variant of the problem $\mathcal{P}_{\mathcal{R}}$ which arises in the absence of the assumption $w_{01} = w_{00} = 0$. This problem, which we denote by $\mathcal{P}'_{\mathcal{R}}$, is defined as follows:

$$\min_{\mathbf{w}} \ (\sigma + (1-\sigma)p_1^2)w_{11} + (1-\sigma)p_1(1-p_1)(w_{10} + w_{01}) + (1-\sigma)(1-p_1)^2 w_{00} \tag{$\mathcal{P}'_{\mathcal{R}}$}$$
$$\text{s.t. } \ \mathcal{IC}_{\mathcal{D}}, \ \widehat{\mathcal{IC}_{\mathcal{R}}}, \ w_{11} \leq w_{10}.$$

In the absence of $w_{01} = w_{00} = 0$, the condition $w_{11} \leq w_{10}$ is not sufficient to conclude that the incentive scheme is RPE. To this respect, the problem $\mathcal{P}'_{\mathcal{R}}$ accounts for the general constraint $\mathcal{IC}_{\mathcal{D}}$, which may be equivalent to $\mathcal{IC}_{\mathcal{S}}$, or $\mathcal{IC}_{\mathcal{J}}$.

In order to figure out how things are different here in comparison to proof of lemma 3.3, which characterizes the optimal solution of $\mathcal{P}_{\mathcal{R}}$, we need to have a better understanding of the constraints. The corresponding expressions for $\mathcal{IC}_{\mathcal{S}}$, $\mathcal{IC}_{\mathcal{J}}$ are already given in the first

step. The constraints $\mathcal{IC}_{\mathcal{R}}$, and $\mathcal{IC}''_{\mathcal{R}}$ are given by:

$$w_{10} + (1+\delta)p_1(w_{11} - w_{10}) - w_{00} + (\delta - (1+\delta)p_1)(w_{01} - w_{00}) > \hat{c}, \qquad (\mathcal{IC}_{\mathcal{R}})$$

$$w_{10} + 2p_1(w_{11} - w_{10}) - w_{00} + (1 - 2p_1)(w_{01} - w_{00}) \geq \hat{c}. \qquad (\mathcal{IC}''_{\mathcal{R}})$$

Furthermore, we can write $\mathcal{IC}'_{\mathcal{R}}$, depending on the value of δ, as follows:

$$\mathcal{IC}'_{\mathcal{R}} = \begin{cases} w_{10} + \frac{(1-\delta^2)p_0 - (\delta+\delta^2)p_1}{1-\delta-\delta^2}(w_{11} - w_{10}) - w_{00} + \frac{(\delta+\delta^2)p_1 - (1-\delta^2)p_0 - \delta^2}{1-\delta-\delta^2}(w_{01} - w_{00}) < \hat{c}, & \text{if } \delta + \delta^2 < 1, \\ (p_1 - \delta p_0)(w_{11} - w_{10}) + (\delta^2 + \delta p_0 - p_1)(w_{01} - w_{00}) > 0, & \text{if } \delta + \delta^2 = 1, \\ w_{10} + \frac{(\delta+\delta^2)p_1 - (1-\delta^2)p_0}{\delta+\delta^2-1}(w_{11} - w_{10}) - w_{00} + \frac{\delta^2 + (1-\delta^2)p_0 - (\delta+\delta^2)p_1}{\delta+\delta^2-1}(w_{10} - w_{00}) > \hat{c}, & \text{if } \delta + \delta^2 > 1. \end{cases}$$

In what follows, we will argue that the solution of $\mathcal{P}'_{\mathcal{R}}$ satisfies $w_{01} = w_{00} = 0$, and conclude that $\mathcal{P}'_{\mathcal{R}}$ is equivalent to $\mathcal{P}_{\mathcal{R}}$. Suppose not to get a contradiction, i.e., we have either $w_{01} \neq 0$, or $w_{00} \neq 0$. Then, the comprehensive constraint $\widehat{\mathcal{IC}_{\mathcal{R}}}$ must have been fulfilled by means of satisfying $\mathcal{IC}'_{\mathcal{R}}$. Otherwise, if $\widehat{\mathcal{IC}_{\mathcal{R}}}$ was carried through $\mathcal{IC}_{\mathcal{R}}$, or $\mathcal{IC}''_{\mathcal{R}}$, then we would have $w_{01} = w_{00} = 0$. The reasons for this are twofold. First, the constraints $\mathcal{IC}_{\mathcal{S}}$, $\mathcal{IC}_{\mathcal{J}}$, $\mathcal{IC}_{\mathcal{R}}$, and $\mathcal{IC}''_{\mathcal{R}}$ increase with the variable w_{00}, yet the objective function decreases. Second, in all these constraints, the coefficients of w_{10} are weakly larger than the ones of w_{01}; while their coefficients are same in the objective function. Therefore, if $\mathcal{IC}'_{\mathcal{R}}$ was redundant, then the solution of the problem $\mathcal{P}'_{\mathcal{R}}$ would satisfy $w_{01} = w_{00} = 0$. For this reason, conditional on the hypothesis, $\mathcal{IC}'_{\mathcal{R}}$ and $\mathcal{IC}_{\mathcal{D}}$ are the relevant constraints. Next, we separately examine the three cases that define $\mathcal{IC}'_{\mathcal{R}}$.

<u>The case: $\delta + \delta^2 > 1$.</u> In this case, all the constraints $\mathcal{IC}'_{\mathcal{R}}$, $\mathcal{IC}_{\mathcal{S}}$, and $\mathcal{IC}_{\mathcal{J}}$ increase with the variable w_{00}, yet the objective function decreases. Moreover, in all these constraints, the coefficients of the variable w_{10} are weakly larger than the ones of w_{01}; while they have the same coefficients in the objective function. Therefore it is optimal to set $w_{01} = w_{00} = 0$, which gives us a contradiction.

The case: $\delta + \delta^2 = 1$. In this case, the sign of the term $(\delta^2 + \delta p_0 - p_1)$ is crucial for the analysis, hence we further separate the discussion into three sub-cases.

If $\delta^2 + \delta p_0 - p_1 < 0$: $\mathcal{IC}'_{\mathcal{R}}$ requires $w_{00} > w_{01}$ as we already have $w_{11} \leq w_{10}$. Then the incentive scheme is an RPE, and $\mathcal{IC}_{\mathcal{D}}$ is equivalent to $\mathcal{IC}_{\mathcal{S}}$. Increasing w_{01}, or w_{00} always hurt the objective function as well as the constraint $\mathcal{IC}_{\mathcal{S}}$. Therefore, the only reason to assign a positive value for w_{01}, or w_{00} is to fulfill $\mathcal{IC}'_{\mathcal{R}}$. Then one can easily see that w_{00} must be set infinitesimally larger than $\frac{p_1 - \delta p_0}{p_1 - \delta^2 - \delta p_0}(w_{10} - w_{11})$, while w_{01} is equal to 0. Consequently, the values of w_{11}, and w_{10} must be arranged in a way to minimize the objective function, conditional on the corresponding value of w_{00} and the constraint $\mathcal{IC}_{\mathcal{S}}$. This choice can be seen in the following problem more explicitly.

$$\min_{w_{11} \leq w_{10}} (\sigma + (1-\sigma)p_1^2)w_{11} + (1-\sigma)p_1(1-p_1)w_{10} + (1-\sigma)(1-p_1)^2 \frac{p_1 - \delta p_0}{p_1 - \delta^2 - \delta p_0}(w_{10} - w_{11})$$
$$\text{s.t.} \quad p_1 w_{11} + (1-p_1)w_{10} - (1-p_1)\frac{p_1 - \delta p_0}{p_1 - \delta^2 - \delta p_0}(w_{10} - w_{11}) \geq \hat{c}$$

Due to the linearity of this problem, w_{11} is either equal to w_{10}, or equal to 0. The latter cannot be the case, because the constraint would be violated as it features a negative coefficient for w_{10}. If precedent is the case, then we have $w_{11} = w_{10} = \hat{c}$, which also brings us a contradiction as it requires $w_{01} = w_{00} = 0$.[8]

If $\delta^2 + \delta p_0 - p_1 > 0$: The contradiction follows here as well from similar arguments to those provided in the previous sub-case.

If $\delta^2 + \delta p_0 - p_1 = 0$: We have an immediate contradiction here. Because, it is impossible to satisfy $\mathcal{IC}'_{\mathcal{R}}$ as it boils down to $w_{11} - w_{10} > 0$, in this case.

The case: $\delta + \delta^2 < 1$. In this case, the direction of the constraint $\mathcal{IC}'_{\mathcal{R}}$ is reversed. In contrast to the previous cases, it now introduces a positive upper bound for an affine function of the choice variables. From the earlier discussions, we know that fulfilling the constraint $\mathcal{IC}'_{\mathcal{R}}$ is the only reason to assign a positive value for w_{01}, and w_{00}. Due to the linearity of the

[8] This choice satisfies $\mathcal{IC}''_{\mathcal{R}}$, therefore $\mathcal{IC}'_{\mathcal{R}}$ becomes redundant.

problem, w_{11} is either equal to w_{10}, or equal to w_{00}. If the precedent is the case, then the best we can do is to set $w_{11} = w_{10} = \hat{c}$. This gives rise to a contradiction as it requires $w_{01} = w_{00} = 0$. If, on the other hand, it is optimal to set $w_{11} = 0$, then the value of w_{10} must be sufficiently large so that we need to assign a positive value for w_{01}, or w_{00} in order to fulfill the constraint $\mathcal{IC}'_{\mathcal{R}}$. More precisely, we need $w_{10} = \frac{1-\delta-\delta^2}{1-\delta-\delta^2-(1-\delta^2)p_0+(\delta+\delta^2)p_1}\hat{c} + A$, for some positive A. Moreover, from linearity, it is without loss of generality to assume that only one of w_{01}, and w_{00} is positive. Here we will discuss the case where $w_{00} > 0$, and the other case follows from similar arguments. The value of w_{00} will be just enough to satisfy $\mathcal{IC}'_{\mathcal{R}}$, i.e., $w_{00} = A\frac{1-\delta-\delta^2}{1-\delta-2\delta^2-(1-\delta^2)p_0+(\delta+\delta^2)p_1} + \epsilon$ for some infinitesimally small $\epsilon > 0$.[9] Note that the incentive scheme is RPE in this case, hence $\mathcal{IC}_{\mathcal{D}}$ is equivalent to $\mathcal{IC}_{\mathcal{S}}$. Then, to satisfy $\mathcal{IC}_{\mathcal{S}}$, we must have $w_{00} \leq A - \frac{\frac{p_1-(1-\delta^2)p_0}{1-p_1}}{1-\delta-\delta^2-(1-\delta^2)p_0+(\delta+\delta^2)p_1}\hat{c}$. Summing up these, we get $A\frac{(\delta+\delta^2)p_1-(1-\delta^2)p_0-\delta^2}{1-\delta-2\delta^2-(1-\delta^2)p_0+(\delta+\delta^2)p_1} > \hat{c}\frac{\frac{p_1-(1-\delta^2)p_0}{1-p_1}}{1-\delta-\delta^2-(1-\delta^2)p_0+(\delta+\delta^2)p_1}$.[10] This condition basically defines a lower bound for the value of A. Then by using this lower bound, we get that $w_{10} > \frac{1-\delta^2+(\delta+\delta^2)p_1-(1-\delta^2)p_0}{(1-p_1)((\delta+\delta^2)p_1-(1-\delta^2)p_0-\delta^2)}$. But then there is no need to have a positive value for w_{00}, because the value of w_{10} is sufficiently large so that $\mathcal{IC}_{\mathcal{R}}$ is satisfied at this value when $w_{00} = 0$. Therefore, we can decrease w_{00} to 0 and improve the objective function without hurting $\mathcal{IC}_{\mathcal{S}}$. As a result $\mathcal{IC}'_{\mathcal{R}}$ become redundant, and we get a contradiction.

To sum up, the solution to problem $\mathcal{P}'_{\mathcal{R}}$ coincides with the solution of $\mathcal{P}_{\mathcal{R}}$, which induces repeated joint work as a team equilibrium outcome. As a result, conditional on the incentive scheme satisfying $w_{11} \leq w_{10}$, it is always optimal to set $w_{01} = w_{00} = 0$. And this completes our proof.

[9] We must have $1 - \delta - 2\delta^2 - (1-\delta^2)p_0 + (\delta+\delta^2)p_1 > 0$, otherwise a contradiction follows immediately since increasing w_{00} would not help to meet $\mathcal{IC}'_{\mathcal{R}}$.

[10] We must have $(\delta+\delta^2)p_1 - (1-\delta^2)p_0 - \delta^2 > 0$; otherwise A would be negative, which contradicts with our assumption.

C.6 Proof of Proposition 3.5

For given w and e_m, the agent has the following problem:

$$\max_{e} \; pwee_m - \frac{c}{2}e^2.$$

The solution to this problem gives us his best response function:

$$\mathbf{e}(w, e_m) = min\{1, \kappa w e_m\}.$$

Then we can write the principal's problem as follows:

$$\max_{w, e_m} \; (1-w)p\mathbf{e}(w, e_m)e_m - \frac{\alpha c}{2}e_m^2.$$

The net value of the success is now equal to $1 - w$, since the principal pays w to agent in case of a success. This term is multiplied with the probability of success, $p\mathbf{e}(w, e_m)e_m$. Finally, since the principal assumes the cost of operating machinery, we subtract it from her revenue.

We first figure out the optimal plan to induce the maximal effort, $e = 1$, from the agent. The wage must be set $w = \frac{1}{\kappa e_m}$, and hence the principal's problem boils down to:

$$\max_{e_m} \; (1 - \frac{1}{\kappa e_m})pe_m - \frac{\alpha c}{2}e_m^2.$$

The first order condition with respect to e_m is given by:

$$\frac{\partial}{\partial e_m} = p - \alpha c e_m.$$

which is always positive, hence $e_m = 1$, and $w = \frac{1}{\kappa}$.

If principal induces agent to choose an effort level $e < 1$, from his best response function,

we know that $e = \kappa w e_m$, and hence the wage must be set $w = \frac{e}{\kappa e_m}$. Then, we can rewrite the principal's problem as follows:

$$\max_{w,e_m} (1-w)wp\kappa e_m^2 - \frac{\alpha c}{2} e_m^2.$$

The first order condition with respect to e_m is given by

$$\frac{\partial}{\partial e_m} = 2(1-w)wp\kappa e_m - \alpha c e_m,$$

which is linear in e_m. Therefore it is optimal to set e_m at its maximum capacity unless the principal would like to shut down the production. Then by plugging in $e_m = 1$, and taking the first order condition with respect to w, one can see that it is optimal to set $w = \frac{1}{2}$. This in turn induces agent to set his effort level to $e = \frac{\kappa}{2}$.

Putting these all together, there are three possibilities for the principal: i) Shutting down the production. ii) Inducing agent to choose $e = \frac{\kappa}{2} < 1$ with $e_m = 1$, and $w = \frac{1}{2}$. iii) Inducing agent to exert maximal effort $e = 1$ with $e_m = 1$, and $w = \frac{1}{\kappa}$. The corresponding profit levels of these are 0, $\frac{p\kappa}{4} - \frac{\alpha c}{2}$, and $p - c - \frac{\alpha c}{2}$ respectively. Then the statement of the proposition immediately follows from a simple comparison between these profit levels.

C.7 Proof of Lemma 3.4

Choosing $(0, 0)$ in every period is a sub-game perfect equilibrium. Therefore, it is natural to expect to see this to be a team equilibrium when w is not sufficiently large to provide incentives.

Now assume that for some w, there is a corresponding team equilibrium that induces positive effort at least for some periods. We want to show that, in this team equilibrium, the agents choose their maximal effort levels in every period. First note that, there is at least a period in which the agents receive a positive payoff. This, however, requires that

the effort pair $(e_1, e_2) = (1, 1)$ maximizes the agents' total payoffs in the stage game due to the increasing marginal returns. Therefore, showing that the repetition of this effort pair in every period can be sustained as a sub-game perfect equilibrium would be sufficient to complete our proof. This stems from the fact that there cannot be any other SPE bringing a higher total payoff to agents. To get a contradiction, suppose that choosing $(e_1, e_2) = (1, 1)$ in every period cannot be sustained even with the most sever punishment, i.e., exerting minimal effort in all the remaining periods. Thus, deviating to the best response in a period is a profitable deviation for agent i. In other words:

$$wp - \frac{c}{2} < (1-\delta)[wp\mathbf{e}(w,1) - \frac{c}{2}(\mathbf{e}(w,1))^2].$$

It must be $\mathbf{e}(\mathbf{w}, 1) = min\{1, w\kappa\} = w\kappa$, because otherwise the agents are already best responding to each other and do not want to deviate. Therefore, we must have

$$wp < \frac{c}{2} + (1-\delta)\frac{w^2p^2}{2c}.$$

However, we already know that the team equilibrium, that we have started with, induces an effort profile $(e_1, e_2) >> (0, 0)$ at least for some periods. Therefore this profile must also be sustained as a SPE with the most sever punishment. In other words, the following incentive constraints for agent 1 and 2 must be satisfied respectively.[11]

$$wpe_1e_2 - \frac{c}{2}e_1^2 \geq (1-\delta)[\frac{w^2p^2}{2c}e_2^2]$$
$$wpe_1e_2 - \frac{c}{2}e_2^2 \geq (1-\delta)[\frac{w^2p^2}{2c}e_1^2]$$

By reorganizing these two inequalities one can get get:

$$wp \geq \frac{c}{2}\left(\frac{e_1^2+e_2^2}{2e_1e_2}\right) + (1-\delta)\frac{w^2p^2}{2c}\left(\frac{e_1^2+e_2^2}{2e_1e_2}\right)$$

[11] These expressions are based on the fact that $\mathbf{e}_i(\mathbf{w}, e_1) = min\{1, w\kappa e_j\} = w\kappa e_j$. The second equality follows from the fact that $\mathbf{e}(\mathbf{w}, 1) = min\{1, w\kappa\} = w\kappa$.

But this contradicts with $wp < \frac{c}{2} + (1-\delta)\frac{w^2p^2}{2c}$, because $e_1, e_2 \leq 1$, and hence $\frac{e_1^2+e_2^2}{2e_1e_2} \geq 1$.

C.8 Proof of Proposition 3.6

The optimal plan either induces the maximal effort in every period, or shuts down the production. The incentive constraint will be relevant only if it induces effort. Denoting $x = wp$, one can rewrite the incentive constraint as follows:

$$-(1-\delta)x^2 + 2cx - c^2 \geq 0$$

The principal will choose the minimum possible value of w satisfying the above condition. We know that $x = \frac{c}{1+\sqrt{\delta}}$ is the minimum value of x at which the above inequality is satisfied. Therefore it is optimal to set $w = \frac{c}{p(1+\sqrt{\delta})}$. Thus, the principal has a positive profit only if $\kappa \geq \frac{2}{1+\sqrt{\delta}}$, where $\kappa = \frac{p}{c}$. Otherwise it is optimal to shut down the production.

BIBLIOGRAPHY

A. Abdulkadiroğlu and K. Bagwell. Trust, reciprocity, and favors in cooperative relationships. *American Economic Journal: Microeconomics*, 5(2):213–259, 2013.

D. Acemoglu and P. Restrepo. The race between machine and man: Implications of technology for growth, factor shares and employment. *MIT working paper*, 2016.

D. Acemoglu and P. Restrepo. Robots and jobs: Evidence from us labor markets. *MIT working paper*, 2017.

A. A. Alchian and H. Demsetz. Production, information costs, and economic organization. *The American economic review*, 62(5):777–795, 1972.

D. H. Autor, F. Levy, and R. J. Murnane. Upstairs, downstairs: Computers and skills on two floors of a large bank. *ILR Review*, 55(3):432–447, 2002.

M. Battaglini. Long-term contracting with markovian consumers. *The American economic review*, 95(3):637–658, 2005.

G. S. Becker. Crime and punishment: An economic approach. In *The Economic Dimensions of Crime*, pages 13–68. Springer, 1968.

J. Benhabib and A. Rustichini. Vintage capital, investment, and growth. *Journal of Economic Theory*, 55(2):323–339, 1991.

L. Benveniste and J. Scheinkman. On the differentiability of the value function in dynamic models of economics. *Econometrica*, 47(3):727–732, 1979.

D. Bergemann and M. Said. Dynamic auctions. *Wiley Encyclopedia of Operations Research and Management Science*, 2011.

J. E. Bessen. How computer automation affects occupations: Technology, jobs, and skills. *Bessen, James E., How Computer Automation Affects Occupations: Technology, Jobs, and Skills (October 3, 2016). Boston Univ. School of Law, Law and Economics Research Paper No. 15-49.*, 2016.

S. Board. Durable-goods monopoly with varying demand. *The Review of Economic Studies*, 75(2):391–413, 2008.

S. Board and A. Skrzypacz. Revenue management with forward-looking buyers. *Unpublished manuscript, Stanford University*, 2010.

E. Brynjolfsson and A. McAfee. *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy.* Brynjolfsson and McAfee, 2012.

J. Bulow. An economic theory of planned obsolescence. *The Quarterly Journal of Economics*, 101(4):729–749, 1986.

L. Carmichael. Firm-specific human capital and promotion ladders. *The Bell Journal of Economics*, pages 251–258, 1983.

Y.-K. Che and S.-W. Yoo. Optimal incentives for teams. *American Economic Review*, pages 525–541, 2001.

Z. Cohn. A note on linked bargaining. *Journal of Mathematical Economics*, 46(2):238–247, 2010.

E&E Publishing, LLC. Puny fines, scant enforcement leave drilling violators with little to fear. `http://www.eenews.net/stories/1059956366`, November 14, 2011. [Online; accessed 4-November-2016].

Environmental Protection Agency. Incentives for self-policing: Discovery, disclosure, correction and prevention of violations; notice. *Federal Register*, 65(70):19617–19627, 11 April 2000.

H. Fang and P. Norman. To bundle or not to bundle. *The RAND Journal of Economics*, 37(4):946–963, 2006.

J. Farrell. Credible repeated game equilibrium. *Unpublished manuscript*, 1983.

J. Farrell and E. Maskin. Renegotiation in repeated games. *Games and economic behavior*, 1(4):327–360, 1989.

A. Fishman and R. Rob. Product innovation by a durable-good monopoly. *The RAND journal of Economics*, pages 237–252, 2000.

C. B. Frey and M. A. Osborne. The future of employment: how susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114:254–280, 2017.

D. Garrett. Incoming demand with private uncertainty. *Unpublished manuscript*, 2012.

C. Hauser and H. Hopenhayn. Trading favors: Optimal exchange and forgiveness. Technical report, Collegio Carlo Alberto, 2008.

B. Holmstrom. Moral hazard in teams. *Bell Journal of Economics*, 13(2):324–340, 1982.

J. Horner and Y. Guo. Dynamic mechanisms without money. 2015.

R. Hortala-Vallve. Inefficiencies on linking decisions. *Social Choice and Welfare*, 34(3): 471–486, 2010.

R. Innes. Remediation and self-reporting in optimal law enforcement. *Journal of Public Economics*, 72(3):379–393, 1999a.

R. Innes. Self-policing and optimal law enforcement when violator remediation is valuable. *Journal of Political Economy*, 107(6):1305–1325, 1999b.

M. O. Jackson and H. F. Sonnenschein. Overcoming incentive constraints by linking decisions1. *Econometrica*, 75(1):241–257, 2007.

S. Jayachandran, S. Sharma, P. Kaufman, and P. Raman. The role of relational information processes and technology use in customer relationship management. *Journal of marketing*, 69(4):177–192, 2005.

L. Kaplow and S. Shavell. Optimal law enforcement with self-reporting of behavior. *Journal of Political Economy*, 102(3):583–606, 1994.

E. P. Lazear. Pay equality and industrial politics. *Journal of political economy*, 97(3): 561–580, 1989.

E. P. Lazear and S. Rosen. Rank-order tournaments as optimum labor contracts. *Journal of political Economy*, 89(5):841–864, 1981.

J. Li, N. Matouschek, and M. Powell. Power dynamics in organizations. *American Economic Journal: Microeconomics*, 2015.

E. Lipnowski and J. Ramos. Repeated delegation. *Available at SSRN 2552926*, 2015.

M. M. Möbius. Trading favors. 2001.

D. Mookherjee. Optimal incentive schemes with many agents. *The Review of Economic Studies*, 51(3):433–446, 1984.

R. T. Moriarty and G. S. Swartz. *Automation to boost sales and marketing*. Harvard Business Review, Reprint Service, 1989.

B. J. Nalebuff and J. E. Stiglitz. Prizes and incentives: towards a general theory of compensation and competition. *The Bell Journal of Economics*, pages 21–43, 1983.

M. Pai and R. V. Vohra. Optimal auctions with financially constrained bidders. Technical report, Discussion paper//Center for Mathematical Studies in Economics and Management Science, 2008.

A. S. Pfaff and C. W. Sanchirico. Environmental self-auditing: Setting the proper incentives for discovery and correction of environmental harm. *Journal of Law, Economics, and Organization*, 16(1):189–208, 2000.

J. D. Sachs, S. G. Benzell, and G. LaGarda. Robots: Curse or blessing? a basic framework. Technical report, National Bureau of Economic Research, 2015.

J. L. Short and M. W. Toffel. Coerced confessions: Self-policing in the shadow of the regulator. *Journal of Law, Economics, and Organization*, 24(1):45–71, 2008.

S. E. Spear and S. Srivastava. On repeated moral hazard with discounting. *The Review of Economic Studies*, 54(4):599–617, 1987.

C. Speier and V. Venkatesh. The hidden minefields in the adoption of sales force automation technologies. *Journal of Marketing*, 66(3):98–111, 2002.

N. L. Stokey. Intertemporal price discrimination. *The Quarterly Journal of Economics*, pages 355–371, 1979.

P. L. Swan. Market structure and technological progress: The influence of monopoly on product innovation. *The Quarterly Journal of Economics*, 84(4):627–638, 1970.

J. Thomas and T. Worrall. Income fluctuation and asymmetric information: An example of a repeated principal-agent problem. *Journal of Economic Theory*, 51(2):367–390, 1990.

E. Van Damme. Renegotiation-proof equilibria in repeated prisoners' dilemma. *Journal of Economic theory*, 47(1):206–217, 1989.

N. Venkatraman. It-enabled business transformation: from automation to business scope redefinition. *Sloan management review*, 35(2):73, 1994.

S. Wang, P. Sun, and F. de Véricourt. Inducing environmental disclosures: A dynamic mechanism design approach. *Operations Research*, 64(2):371–389, 2016.

Washington Post. New Weapon to Force Compliance. `https://www.washingtonpost.com/archive/politics/1977/02/05/new-weapon-to-force-compliance/34549987-1943-47ad-882c-dcde9c553fae/`, February 5, 1977. [Online; accessed 4-November-2016].