UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays in Experimentation, Voting with Sabotage, and Dynamic Inconsistency

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

in

Economics

by

Danil Dmitriev

Committee in charge:

Professor T. Renee Bowen, Chair
Professor Simone Galperti
Professor Seth Hill
Professor Isabel Trevino
Professor Joel Watson

2023

The dissertation of Danil Dmitriev is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

To my family and my friends.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

From the bottom of my heart, I want to thank my advisor Renee Bowen for the vast support she has given me for almost five years. Her advice has always been thoughtful and compassionate, and I would not have been able to reach this point without her guidance.

I am immensely grateful to my mother for always believing in me and making every effort and every sacrifice to help me succeed.

There are many wonderful friends I met both before and during my graduate studies—Adrian, Ellen, Freddie, Sasha. Thank you for being there for me during these six years.

I would also like to thank Renee Bowen, Simone Galperti, Freddie Papazyan, and Adrian Wolanski for being amazing co-authors. Working with you has been a delightful experience that has made me a better researcher and a better collaborator.

Finally, I am grateful to my committee members Simone Galperti, Joel Watson, Isabel Trevino, and Seth Hill for their feedback, guidance, and encouragement.

Chapter 1, “Motivating Creativity," in part is currently being prepared for submission for publication of the material. The dissertation author, Danil Dmitriev, was the sole author of this chapter.

Chapter 2, “Sabotage-Proof Mechanism Design," is currently planned for submission for publication of the material therein. Frederick Papazyan and the dissertation author, Danil Dmitriev, were the principal co-authors of this chapter.

Chapter 3, “Dynamic Inconsistency and Convex Commitment Devices," is currently planned for submission for publication of the material therein. Adrian Wolanski and the dissertation author, Danil Dmitriev, were the principal co-authors of this chapter.

VITA

2017	Bachelor of Science, NRU Higher School of Economics
2021-2022	Research Assistant, University of California San Diego
2018-2023	Teaching Assistant, University of California San Diego
2023	Doctor of Philosophy, University of California San Diego

PUBLICATIONS

Bowen, Renee, Dmitriev, Danil and Galperti, Simone. “Learning From Shared News: When Abundant Information Leads to Belief Polarization”, *Quarterly Journal of Economics*, Volume 138, Issue 2, May 2023, pp. 955-1000

ABSTRACT OF THE DISSERTATION

Essays in Experimentation, Voting with Sabotage, and Dynamic Inconsistency

by

Danil Dmitriev

Doctor of Philosophy in Economics

University of California San Diego, 2023

Professor T. Renee Bowen, Chair

This dissertation examines the optimal design of incentives or mechanisms in various strategic and non-strategic settings.

Chapter 1 studies how to provide incentives for creativity by analyzing a stylized model of delegated creative experimentation. A principal desires an agent to frequently switch to new uncertain projects to maximize the chance of success, while the agent faces a fixed cost of switching. We show that the principal's optimal reward scheme is maximally uncertain—the agent receives transfers for success, but their distribution has extreme variance. Despite being stationary, the optimal reward scheme achieves the principal's first-best outcome provided that the agent's outside option is sufficiently

valuable. These results shed light on the value of randomized incentives for motivating creativity and provide guidance on how to design optimal bonus schemes in online platforms and other applications.

Chapter 2 studies the design of robust voting mechanisms in the presence of sabotage. We consider a preference aggregation problem in which the designer faces both genuine agents and outside saboteurs. We show that plurality voting and other standard mechanisms are typically not robust to sabotage. The optimal voting mechanism must make saboteurs indifferent between each alternative they can vote for. Based on the property, we suggest simple ways of improving standard voting mechanisms to make them more robust to outside sabotage.

Chapter 3 studies perceptions of dynamic inconsistency in labor provision over time. We present a novel laboratory experimental design intended to measure both actual and perceived dynamic inconsistency by using a convex commitment device. In a pilot experiment, we find that participants demand significant amounts of commitment despite showing little dynamic inconsistency in their labor choices. This implies that they believe themselves to be more dynamically inconsistent than they are. The results suggest caution when employing commitment devices, as their usage may be unrelated to the inconsistency they attempt to solve.

Chapter 1

Motivating Creativity

1.1 Introduction

Frequently trying new ideas (i.e., being creative) is important in many work environments. Online platforms, such as YouTube, TikTok, or Facebook, desire creators to try new content variations to find where they can capture the attention of the platform's users and generate engaging content. A recording label would like its artists to constantly try new variations of their music to maximize the chance of a "hit". These situations share two key features: (i) a profit-maximizing principal that benefits from the creator's success; (ii) the creator can recycle old variations in the hope of eventual success or develop a new variation at some cost. Since the principal is not directly involved with experimentation, she is typically free of this switching cost, which leads to a misalignment of incentives. The creator prefers to stick to each content variation for some time even in the absence of success. The principal does not internalize the creator's switching cost and would like him to experiment with new variations constantly until success arrives. How should the principal motivate the creator to be more creative?[1]

[1]While this dissertation chapter is focused on private benefits of creativity, discovery of new ideas is also a crucial factor in explaining economic growth, e.g. see Jones (2002, 2022). Studying how to properly

The contribution of this dissertation chapter is two-fold. First, we reveal a novel role that *extreme* uncertainty of rewards can play in incentivizing creative experimentation. When the agent draws a low reward on a new project, he is incentivized to switch early to another project in expectation of getting a better reward. The principal can exploit this incentive by making the high reward arbitrarily large and concentrating almost all probability on the low reward. Our second contribution is more technical. Existing literature on delegated experimentation with moral hazard finds that it is generally optimal for the principal to use a time-contingent reward scheme, i.e., a transfer rule with deadlines (e.g., see Halac et al., 2016; Guo, 2016). This solution is a natural way to affect the agent's dynamic incentives but can be costly to implement due to monitoring or enforcement costs. Stationary reward schemes (e.g., a fixed payment for achieving success) are easier to implement in practice, especially when dealing with many agents. We show that (stationary) uncertainty of rewards can provide the necessary dynamic incentives for the agent to experiment efficiently. We also derive a condition under which the restriction to stationary rewards does not prevent the principal from achieving her ideal outcome.

We analyze a stylized principal-agent model of delegated experimentation with switching costs. There is a principal and an agent. Agent engages in continuous-time experimentation, facing an unlimited number of bandit arms (projects). He can only work on one arm at a time. If an arm is good, it delivers success at a positive rate while being worked on; however, a bad arm never succeeds. Hence, good news is conclusive. There is a common prior that an unexplored arm is good. When a success happens, both the agent and the principal get payoffs. At any point, the agent can stop working on the current project, pay a fixed switching cost and start working on a new one. The principal faces no switching cost. This model captures important features of existing

motivate creative experimentation is important for both private companies and the society as a whole.

environments, such as YouTube. We can think of the agent as a new content creator. He must experiment with new content variations in order to discover the one where he can achieve success. Switching to a different type of content involves buying new software or equipment and researching the existing content of this type, which can be viewed as a switching cost for the creator. The platform is the principal and does not pay this cost because it is not directly involved in the creator's experimentation.

Without additional incentives from the principal, the agent's optimal policy follows a standard index rule. If there was no switching cost, this would result in instantaneous switching to a new project when there is no immediate success on the current one. With the switching cost, the agent spends positive time on each project before switching. Since the principal faces no such switching cost, her preferred policy is instantaneous switching. Hence, there is a conflict of incentives—the principal would like the agent to switch quicker than he does. For example, YouTube likely views a content creator that has worked on the same kind of content for a couple years without achieving success as being stuck too long.

To address this conflict, the principal can design (and commit to) a bonus scheme. We assume that the bonus scheme takes the following form: Each project is assigned a fixed payment that is paid if and when that project succeeds.[2] The agent does not know the reward on a given arm until he starts working on it. This assumption rules out a few natural mechanisms, but we believe it is practical. The main type of mechanism we rule out is any mechanism that contains time-contingent payments or deadlines (Halac et al., 2016; Guo, 2016). One could imagine promising the agent a bonus payment, but also fixing a deadline at which they must switch to a new arm. Maintaining such time-contingent schemes is costly in many situations. For example, YouTube deals with hundreds of thousands of creators at once, which would make maintaining

[2] This is reminiscent of YouTube paying content creators for the number of views on their videos.

individual time-contingent contracts with each creator quite costly. In general, implementing a mechanism relying on memory is costlier than a mechanism that does not keep track of it. We wish to investigate how well the principal can perform when she is restricted to simple, stationary reward schemes.[3]

What is the optimal bonus scheme for the principal? To motivate the discussion, let us return to YouTube. The platform incentivizes creators via a recommendation algorithm and a pay-for-ads scheme. Perhaps surprisingly, both are quite non-transparent for creators. As The Verge puts it, "YouTubers are at the mercy of a platform they do not fully understand". They must experiment with different content variations to learn how much the platform incentivizes each variation, and sometimes see that their new content is not rewarded that much. While this non-transparency is discomforting for creators and contributes to burnout (New York Times), it may be benefiting the platform in some way. Our model can shed light on the nature of this benefit.

The non-transparency of YouTube's incentive schemes can be expressed in our model as the uncertainty of the principal's bonus scheme. We will focus on two approaches that the principal can take — *transparent* and *opaque*. A transparent bonus scheme sets a fixed bonus payment for all arms, so the agent knows what he will get if he switches. In contrast, an opaque bonus scheme sets a distribution of rewards for unexplored arms but does not initially reveal rewards to the agent. When the agent switches to a new arm, its bonus is drawn randomly and revealed. This design creates an asymmetry between the agent's current arm, which has a known bonus, and an unexplored arm, whose bonus is uncertain. In this way, *stationary* uncertainty of the bonus scheme can change the agent's *dynamic* incentives for switching to a new arm.

Our first key result (Theorem 1) states that opaque bonus schemes perform strictly better for the principal than transparent ones. Specifically, we show that given

[3]We also assume non-negative transfers in the main model, which can be viewed as a liquidity constraint. We relax this assumption in Section 1.5 and show that this only strengthens our results.

any transparent bonus scheme, there exists an opaque (opaque) bonus scheme that keeps the agent just as happy and offers the principal a strictly higher expected value. To understand how opaque rewards work, suppose that the principal offers a distribution of one low reward and one high reward. If the agent draws the low reward, this provides extra endogenous pressure to switch early, since a new arm's expected reward is better than the current one. On the other hand, drawing a high reward suppresses the switching incentive for a symmetric argument. The principal thus benefits from the low-reward cases, i.e., when the agent gets unlucky. We show that the principal can design the distribution of rewards in such a way that the gain from the low-reward cases outweighs the loss from the high-reward cases.

Our second result (Theorem 2) concerns the structure of the optimal bonus scheme. We know it is opaque, but what properties does its reward distribution have? It turns out that an opaque bonus scheme performs best when it leverages an extremely large but unlikely reward to maximize the probability of a low reward that induces efficient experimentation. That is, if we fix any opaque reward distribution with a given expected value to the agent, the principal can do strictly better by increasing the value of the high reward and shifting the probability to the low reward. Doing so exploits the fact that the agent's continuation value conditional on the current arm's reward is increasing and unbounded in the reward. The principal chooses the low reward so that the agent's experimentation policy, conditional on the reward and the overall agent's value, is socially efficient. Intuitively, the principal leverages the size of the high reward (under which experimentation is suboptimal) to make the low reward (under which experimentation is optimal) almost certain. Since the goal is to make the probability of the low reward as high as possible, the principal wants to make the high reward arbitrarily large.

Our final result (Theorem 3) derives a condition under which opaque bonus schemes achieve the principal's ideal outcome.[4] That outcome is characterized by a socially efficient experimentation policy and full surplus extraction from the agent. When the principal cannot impose a negative bonus on the agent, opaque bonus schemes achieve the principal's ideal outcome if and only if the agent's outside option is sufficiently valuable. This result hinges on the combination of the agent's participation constraint and the liquidity constraint that prevents the principal from taxing the agent. The outside option sets a value that the principal must assign to an explored arm to induce the agent to begin experimentation. Dropping the low reward as low as the liquidity constraint allows, the principal induces the quickest possible switching time. This is a lower bound on the agent's switching time under the liquidity and participation constraints that the principal faces. The principal can achieve her ideal outcome if and only if this bound is below the socially efficient switching time. This is equivalent to the agent's outside option having a sufficiently high value. If the liquidity constraint is relaxed by allowing the principal to tax the agent's success payoff, the principal's ideal outcome can always be achieved (Theorem 4).

The extreme spread of rewards described above is reminiscent of the so-called "Mirrlees problem" in the classical hidden-action model. The principal observes a signal of the agent's action in the form of output, and the signal's support is infinite. When the tails of the signal distribution are sufficiently informative, it is possible to incentivize the first-best action by setting an extreme punishment for sufficiently extreme signal realizations (Bolton and Dewatripont, 2005, pp. 140-141). The mechanism in this dissertation chapter works in a similar way but is more subtle because it relies on dynamic revelation of information. The principal uses a very large reward to provide the necessary incentive for the agent to switch early conditional on drawing a bandit arm with a low reward. As opposed to the hidden-action model above, this incentive is dynamic

[4] Technically, opaque bonus schemes can get arbitrarily close to the principal's ideal value.

and comes at a potential cost for the principal. That cost vanishes as the high bonus gets arbitrarily large due to the probability being fully concentrated on the low bonus.

The role of uncertainty in incentivizing creativity sheds new light on the incentive schemes we see in practice. Consider YouTube once again. The platform has two main mechanisms that provide incentives for novice content creators—a recommendation algorithm and a share of advertising revenue (pay-for-ads). Both systems have a lot of uncertainty and non-transparency for content creators.[5] Consider the pay-for-ads scheme, which pays creators based on how much ad revenue their videos generate. The specific pay rate is sensitive to a video's content niche and the region where it is watched and can vary by an order of magnitude (as discussed in this Thinkific blog). When a creator switches to a new content variation, they cannot predict with certainty what pay rate their new videos will receive. Our results suggest that this uncertainty, even if not intended by YouTube, can induce creators to experiment more frequently and achieve success faster. If YouTube wishes to leverage this uncertainty to incentivize creative experimentation, we suggest that the distribution of rewards should be extreme—a high pay rate with low probability or a low pay rate with high probability. We discuss in detail how our model relates to YouTube and similar platforms in Section 1.4.

The rest of the chapter is structured as follows. Section 1.2 presents the main model. Section 1.3 describes the key results. Section 1.5 discusses a few important extensions, such as allowing negative transfers or introducing moral hazard into the problem. Section 1.4 discusses how our results apply to YouTube and other online platforms that rely on content creation.

[5] Discussion of YouTube algorithm transparency can be found in the following articles at the Schwartz Reisman Institute for Technology and Society, Fast Company, and Forbes.

Related literature

Broadly, this dissertation chapter belongs to the literature on experimentation with learning.[6] While we model experimentation using exponential bandit arms[7], we consider a relatively new variation where there is an unlimited supply of risky arms that the agent can explore. Our model is closely connected to that of Sadler (2021), who considers a single-agent setting with a known distribution of rewards on the unexplored arms. He analyzes how the structure of the reward distribution affects the agent's experimentation policy and a social planner's preference for using subsidies or taxes. We differ from his analysis by focusing on a principal-agent conflict of incentives that arises due to the principal facing no switching cost. We show that the value of the agent's outside option has an important effect on the efficiency of the optimal bonus scheme. Additionally, we explicitly derive the agent's optimal index policy. To do that, we rely on the applicability of the well-known Gittins index to stationary bandit problems with switching cost, as shown in Bergemann and Välimäki (2001) for a discrete-time model.

This dissertation chapter primarily contributes to the literature on delegated experimentation, which is typically focused on solving the issues of informational asymmetry[8], moral hazard[9], and adverse selection[10]. Our analysis is closest to the moral hazard strand of the literature, where we wish to highlight Halac et al. (2016). They consider a dynamic experimentation setting where the principal cannot contract over which arm the agent experiments with. However, arbitrary time-contingent transfers are allowed. Our model can be construed as a setting where the principal cannot contract over how long the agent has been experimenting, but can assign different rewards

[6]For a wide overview, see Hörner and Skrzypacz (2016).

[7]See Kaspi and Mandelbaum (1998); Keller et al. (2005).

[8]See Gomes et al. (2016); Halac et al. (2016); Guo (2016).

[9]See Bergemann and Hege (2005); Manso (2011); Bonatti and Hörner (2011); Halac et al. (2016); Bonatti and Hörner (2017); Thomas (2021).

[10]See Halac et al. (2016).

to different arms. Hence, we contribute to the literature by deriving the optimal bonus scheme when it is impractical or costly to use the standard solution of time-dependent transfers. We also derive conditions under which this restriction does not prevent the principal from achieving her first-best outcome.

Another paper we wish to highlight is Manso (2011), which studies optimal incentives for innovation using a two-armed bandit problem with two periods. It finds that incentivizing innovation sometimes requires tolerance for failure (captured by transfers that occur conditionally on early failure). In particular, pay-for-performance metrics are shown to be insufficient for achieving the principal's goal. We show that this is not the case when there are many risky projects for the agent to experiment with. The principal can use an uncertain performance-based bonus scheme to incentivize optimal exploration of new projects.

Owing to the main application, this dissertation chapter is connected to recent literature studying the design of online platforms which rely on content creation. Jain and Qian (2021) study how ad revenue sharing incentives of a platform are affected by the competition between producers and the properties of the customer base. In related work, Bhargava (2022) analyzes a more general model of a digital platform that monetizes the consumption of goods through third-party advertisers. He focuses on how the distribution of creators' talents is related to market concentration among creators and platform design. We contribute to this literature by studying how to optimally incentivize novice producers to experiment with new content and how the uncertainty of the platform's incentive scheme can accelerate their success.

This dissertation chapter is also connected to the literature on experimentation in two-sided markets. Gomes and Pavan (2016) focus on price discrimination in matching markets where agents have private characteristics that affect the match values. Peitz et al. (2017) analyze optimal experimentation by a platform in a two-sided market where the externality value of matches between consumers and producers is initially uncertain. Jullien and Pavan (2018) study platform markets where information on users' preferences is dispersed, focusing on participation decisions when a new, uncertain platform market is launched. We focus on a different aspect of experimentation in two-sided markets—new producers experimenting with products to find one that they can provide successfully.

1.2 Model

There is a principal (she) and an agent (he). Time is continuous, $t \in [0, \infty)$. The agent can experiment by choosing one of infinitely many exponential bandit arms to explore or take a safe outside option. The principal designs a bonus scheme to incentivize optimal experimentation by the agent. Both players are risk-neutral. We can think of the principal as a platform, for example YouTube, and the agent as a content creator that experiments with different content variations. We will now explain each element of the model in more detail.

Experimentation. We model each variation of content as an exponential bandit arm. Each arm can be good or bad, and there are infinitely many of them. An unexplored arm is good with probability π_0 and bad otherwise. A good arm delivers success at rate λ while the agent is exploring it. A bad arm never succeeds, and hence good news is conclusive. When an arm succeeds, the principal and the agent get payoffs y_P and y_A, respectively. Since good news is conclusive, the agent stops switching arms as soon as a success occurs. Returning to YouTube, we can think of "success" as the creator finding a content variation that attracts a lot of viewership. If the agent takes the outside option, it generates flow payoff $s > 0$ to the agent and 0 to the principal.

Learning. As the agent is working on an arm, his belief that it is good will be gradually drifting down as long as success is not occurring. After working on an arm for time t without success, the posterior belief that the arm is good equals

$$\pi_t = \frac{\pi_0 e^{-\lambda t}}{1 - \pi_0 + \pi_0 e^{-\lambda t}}.$$

Since good news is conclusive, observing a success will cause π_t to jump to 1, and cause the agent to never switch from that arm.

Switching costs. At any moment, the agent can choose to continue the current action (experimenting with a bandit arm or taking the safe payoff), or begin experimenting on a new arm. When the agent begins to experiment on a new arm, he incurs a one-time switching cost $c > 0$.

Absent additional incentives from the principal and given the switching cost, the agent's optimal policy is to experiment on each arm until its posterior belief drifts below a certain level. This is a well-known index policy that is summarized in the following lemma.

Lemma 1. *The agent switches to a new arm if and only if the current belief π satisfies $\pi < \pi^A$, where*

$$\pi^A = \frac{\hat{m}}{\lambda y_A}$$

and

$$\hat{m} = \sup_{t \geq 0} \frac{-c + \frac{\pi_0 \lambda}{r+\lambda}(1 - e^{-(r+\lambda)t}) y_A}{\frac{1}{r}\left(\frac{\pi_0 r}{r+\lambda}(1 - e^{-(r+\lambda)t}) + (1 - \pi_0)(1 - e^{-rt})\right)}.$$

The proof of this and all following results are in the Appendix.[11] Note that $\hat{m}$ is a modified version of the standard Gittins index that accounts for the switching cost. Observe that, if $c = 0$, the index simplifies to $\hat{m} = \pi_0 \lambda y_A$. This implies that without a

[11]The proof relies on the stationarity of the problem, which has already been used in Bergemann and Välimäki (2001) to prove the optimality of index policies in a discrete-time version of our model.

switching cost the cutoff belief would be $\pi^A = \pi_0$ (i.e., instantaneous switching).

Feasible bonus schemes. The agent's switching cost creates a conflict of interest between the principal- and the agent-optimal experimentation policy. With an infinite number of identical arms and no switching cost, the principal prefers instantaneous switching between arms conditional on absence of success, i.e. $\pi^P = \pi_0 > \pi^A$. To address this conflict, the principal can commit to a bonus scheme (i.e., a transfer rule) at $t = 0$.

We assume that the principal cannot contract over the time spent on each arm, but can observe the success of an arm. We also assume that the principal cannot pay directly for switching to a new arm.[12] Formally, this means that the set of feasible contracts consists of all possible transfer distributions on the available arms, where the transfers are paid upon success. We make this restriction for a few reasons. We wish to determine whether the principal can achieve her first-best outcome using a simple, stationary reward scheme. Stationary reward schemes are also less costly to implement than a history-contingent mechanism. This makes them more practical in applications, especially if they still allow the principal to achieve her first-best outcome.

Transparency of bonus schemes. There are two types of stationary bonus schemes that we will focus on—*transparent* and *opaque*. A transparent bonus scheme assigns a positive bonus payment b to each unexplored arm that the principal pays to the agent if that arm succeeds. This bonus payment is revealed to the agent at time $t = 0$. In contrast, an opaque bonus scheme assigns a distribution F of positive bonus payment b to each unexplored arm. We allow any distribution with finite support. The distribution is revealed to the agent at time $t = 0$, but the realized reward on a given arm is not revealed until the agent starts experimenting on it.[13] Hence, the continuation value of

[12]A switching bonus is ruled out because the principal would still have to keep track of the agent's history. If she did not, the agent could abuse the mechanism by constantly switching back and forth between two arms.

[13]We assume $b \geq 0$ in the main model because it is more consistent with observed mechanisms, in which the principal does not tax the agent's payoff. This assumption is relaxed in Section 1.5, and this

a new arm is uncertain until the agent pays the switching cost and starts working on it.

Equilibrium. We use Markov perfect equilibrium as our solution concept. The payoff-relevant state variables are the belief π on the current arm and the bonus payment b on the current arm. The principal sets a bonus scheme for all arms at $t = 0$, and the agent chooses an action (continue, switch, take outside option) for every belief π and reward w. In equilibrium, the two must be best-responding to each other. Formally, the principal chooses a finite-support distribution F over $b \geq 0$. The agent chooses a function $a : \mathbb{R} \times [0, 1] \to \{\text{outside option}, \text{ current arm}, \text{ new arm}\}$ that maps the bonus on the current arm and current belief into actions.

1.3 Analysis

Lemma 1 above characterizes the agent's switching policy in the absence of a principal's incentive scheme. We will begin with a characterization of the agent's optimal switching policy given a bonus scheme that offers him value V_A. This result will be useful for characterizing transparent and opaque bonus schemes in subsequent sections.

1.3.1 Agent's switching policy

Suppose the agent is currently experimenting on an arm with promised reward b and current belief π. Further suppose that switching to a new arm has value V_A to the agent. As is standard, the agent's optimal experimentation strategy consists of a cutoff switching belief.

---

only strengthens our results.

Lemma 2. *Let V_A be the agent's value of a new arm, and suppose that the current arm has bonus payment b and posterior belief π. The agent switches to a new arm if and only if*

$$\pi \leq \frac{rV_A}{\lambda(y_A + b - V_A)}.$$

Note that V_A in this lemma includes the switching cost, i.e. $V_A = -c + u_A(\pi_0, V_A)$, where $u_A(\pi_0, V_A)$ is the agent's continuation value of the new arm, given optimal policy.

The policy from Lemma 2 can also be expressed in terms of a switching time. The agent works on each arm for time $t = t^A(b, V_A)$ that satisfies

$$t^A(b, V_A) = -\frac{1}{\lambda} \ln\left(\frac{(1-\pi_0)rV_A}{\pi_0(\lambda(y_A + b - V_A) - rV_A)}\right).$$

1.3.2 Transparent bonus schemes

We will begin the analysis with a simpler case—transparent bonus schemes. The principal chooses a deterministic reward b that is offered on all arms. Whenever the agent draws a new arm, he is promised the same payment b upon success. Note that since all arms receive the same bonus b, the value of a new arm V_A depends on b and the agent's switching rule. Since Lemma 2 relies on the value V_A to characterize the switching time, we will not explicitly solve for the optimal transparent bonus scheme. We instead provide a characterization of transparent bonus schemes that will be useful for showing that they can be improved upon by opaque (stochastic) bonus schemes.

Consider a transparent bonus scheme that offers the agent some "target" value $\bar{V}_A \geq \bar{u}$. Note that this completely pins down bonus payment b. To see why, let τ be the agent's switching time. Then his overall value V_A can be written as

$$V_A = -c + \int_0^\tau \pi_0 \lambda e^{-(r+\lambda)t}(y_A + b)dt + \left(1 - \pi_0 + \pi_0 e^{-\lambda\tau}\right) e^{-r\tau} V_A$$

The first term is the switching cost that is paid immediately. The second term is the expected payoff conditional on getting a success before switching at $t = \tau$, and the third term is the discounted value of a new arm. Since the problem is recursive, the current arm's initial value and a new arm's value in the future both equal V_A. Plugging in $V_A = \bar{V}_A$ and $\tau = t^A(b, \bar{V}_A)$, we get an implicit equation that relates $\bar{V}_A$ and b. Clearly, higher b induces a higher V_A, so there is a unique payment b that satisfies this equation for a given "target" value $\bar{V}_A$. Denote that bonus payment by $b^A(\bar{V}_A)$.

We can similarly express the principal's overall value V_P as a function of b and τ:

$$V_P = \int_0^{\tau} \pi_0 \lambda e^{-(r+\lambda)t}(y_P - b)dt + \left(1 - \pi_0 + \pi_0 e^{-\lambda\tau}\right) e^{-r\tau} V_P$$

We can then plug in $\tau = t^A(b^A(\bar{V}_A), \bar{V}_A)$ and $b = b^A(\bar{V}_A)$, which gives us the principal's overall value $V_P(\bar{V}_A)$ from offering a deterministic bonus scheme with value $\bar{V}_A$ to the agent. The choice of the optimal transparent bonus scheme boils down to optimizing $V_P(\bar{V}_A)$ over $\bar{V}_A$ subject to the agent's participation constraint:

$$\max_{\bar{V}_A \geq \bar{u}} \frac{\frac{\pi_0 \lambda}{\lambda + r}\left(1 - e^{-(r+\lambda)t^A(b^A(\bar{V}_A), \bar{V}_A)}\right)(y_P - b^A(\bar{V}_A))}{1 - \left(1 - \pi_0 + \pi_0 e^{-\lambda t^A(b^A(\bar{V}), \bar{V})}\right) e^{-rt^A(b^A(\bar{V}_A), \bar{V}_A)}}.$$

However, identifying the optimal transparent bonus scheme is irrelevant, as we will next show that you can improve upon a transparent scheme by an opaque bonus scheme. The key reason why this is possible is that transparent bonus schemes are generally inefficient. The principal will only consider bonus payments $b < y_P$ in order to get a positive profit in expectation. For any bonus scheme with $b \in (0, y_P)$, the agent's switching policy leads to inefficiently high switching time,. This is because he does not fully internalize the joint benefit of success, $y_A + y_P$, but faces the full switching cost. As a result, the joint value of experimentation is not maximized for any transparent bonus scheme that offers the principal a positive value. However, opaque bonus schemes can fix this issue.

1.3.3 Opaque bonus schemes

Opaque bonus schemes are characterized by a reward distribution that a new arm's bonus payment is drawn from when the agent switches to it for the first time. Once a reward is drawn, it is fixed for that arm. This introduces endogenous disparity between the current arm and a new arm that affects the agent's dynamic incentives. The agent knows her reward on the current arm, but faces a distribution of several rewards on any new arm. Hence, he will have a combination of switching times, one for each possible reward. Each of them is characterized by Lemma 2.

We will identify an opaque bonus scheme with its underlying reward distribution F hereafter. We consider the set of opaque bonus schemes with any finite-support distribution.[14] In this terminology, a transparent bonus scheme from above is characterized by a deterministic F.

Definition 1. An *opaque bonus scheme* is a probability distribution F over the bonus payment $b \in [0, \infty)$ with finite support.

Let $V_A(F)$ and $V_P(F)$ be the agent's and the principal's overall values when the principal uses an opaque bonus scheme F. These are the values of a brand new arm under the bonus scheme F, given the agent-optimal experimentation policy. Letting $f(b)$ be the probability distribution function of F, we can implicitly express these values as follows:

$$V_A(F) = -c + \sum_{b \in \text{supp } F} f(b) \cdot \left[\int_0^{t^A(b, V_A(F))} \pi_0 \lambda e^{-(r+\lambda)t} (y_A + b) dt + \left(1 - \pi_0 + \pi_0 e^{-\lambda t^A(b, V_A(F))}\right) e^{-r t^A(b, V_A(F))} V_A(F) \right],$$

$$V_P(F) = \sum_{b \in \text{supp } F} f(b) \cdot \left[\int_0^{t^A(b, V_A(F))} \pi_0 \lambda e^{-(r+\lambda)t} (y_P - b) dt + \left(1 - \pi_0 + \pi_0 e^{-\lambda t^A(b, V_A(F))}\right) e^{-r t^A(b, V_A(F))} V_P(F) \right].$$

[14]Assuming finite support is not restrictive, but simplifies the analysis.

It is useful to identify a special class of opaque bonus schemes—*binary*. A binary bonus scheme offers just two rewards to the agent — a "good" and a "bad" one.

Definition 2. A *binary bonus scheme* is an opaque bonus scheme with underlying reward distribution F such that $|\text{supp}(F)| = 2$.

Intuitively, if the agent draws a "bad" reward, this incentivizes him to switch earlier; a "good" reward pushes him to switch later. Since the principal prefers earlier switching time and paying less, she benefits when the agent gets unlucky and draws a low reward and "loses" when the agent draws a high reward. All opaque bonus schemes share this property. Binary bonus schemes take this idea to the extreme and effectively pool all rewards below the average and all rewards above the average into single points. As it turns out, we can restrict our analysis to binary bonus schemes without loss of generality.

Lemma 3. *Consider an arbitrary opaque bonus scheme F. There exists a binary bonus scheme F' such that $V_A(F') = V_A(F)$ and $V_P(F') \geq V_P(F)$.*

In words, this means that an opaque bonus scheme with a complicated reward distribution cannot be better for the principal than a simple binary bonus scheme. The proof shows that you can pick two values in the support of distribution F and construct a binary bonus scheme F' with support on those two values. The resulting bonus scheme keeps the agent as happy as before (thus maintaining participation) and makes the principal weakly better off.

Another way to interpret Lemma 3 is that if you take any opaque bonus scheme, you can weakly improve upon it with a binary bonus scheme. It is natural to ask whether the same logic can be applied to transparent bonus schemes with a deterministic reward. Given a transparent bonus scheme, is there a binary bonus scheme that strictly improves upon it? The answer is yes.

Theorem 1. *Consider any transparent bonus scheme with $b > 0$ and overall values V_A and V_P. There exists a binary opaque bonus scheme F such that $V_A(F) = \bar{V}_A$ and $V_P(F) > \bar{V}_P$.*

Given any bonus scheme that offers a certain bonus payment on success, the principal can do better by replacing that reward with an uncertain lottery between two rewards—low and high. Intuitively, the agent's best response is to switch early under the former and switch late under the latter. The principal benefits from an early switch but is harmed by a late switch. A binary bonus scheme has to balance these two effects while maintaining the agent's participation. Theorem 1 states that it is possible to design this bonus scheme in such a way that the agent is just as happy as under the transparent bonus scheme, and the principal is strictly better off.

To understand how an opaque reward scheme can be superior, let us examine how it affects the agent's dynamic incentives. The principal has three variables to work with: the low reward b_L, the high reward b_H, and the probability distribution. We can capture the latter with the probability of b_L, denoted by q. The distribution F can then be denoted by $F = (b_L, q;\, b_H, 1-q)$. Suppose that the principal wishes to keep the agent's initial value at some $\bar{V}_A \geq \bar{u}$. This pins down the value of q through the agent's value equation:

$$\begin{aligned}\bar{V}_A = -c + q \cdot \left[\int_0^{t_L^A} \pi_0 \lambda e^{-(r+\lambda)t}(y_A + b_L)dt + \left(1 - \pi_0 + \pi_0 e^{-\lambda t_L^A}\right) e^{-r t_L^A} \bar{V}_A\right] \\ + (1-q) \cdot \left[\int_0^{t_H^A} \pi_0 \lambda e^{-(r+\lambda)t}(y_A + b_H)dt + \left(1 - \pi_0 + \pi_0 e^{-\lambda t_H^A}\right) e^{-r t_H^A} \bar{V}_A\right],\end{aligned} \tag{1.1}$$

where $t_L^A = t^A(b_L, \bar{V}_A)$ and $t_H^A = t^A(b_H, \bar{V}_A)$, as determined by Lemma 2. Similarly, we can write the principal's value as

$$
\begin{aligned}
V_P = \qquad & q \cdot \left[\int_0^{t_L^A} \pi_0 \lambda e^{-(r+\lambda)t} (y_P - b_L) dt + \left(1 - \pi_0 + \pi_0 e^{-\lambda t_L^A}\right) e^{-r t_L^A} V_P \right] \\
& + (1-q) \cdot \left[\int_0^{t_H^A} \pi_0 \lambda e^{-(r+\lambda)t} (y_P - b_H) dt + \left(1 - \pi_0 + \pi_0 e^{-\lambda t_H^A}\right) e^{-r t_H^A} V_P \right].
\end{aligned} \tag{1.2}
$$

The principal's value can also be represented as $V_P = V^* - \bar{V}_A$, where V^* is the joint value of the two players defined as follows:

$$
\begin{aligned}
V^* = -c + q \cdot & \left[\int_0^{t_L^A} \pi_0 \lambda e^{-(r+\lambda)t} (y_A + y_P) dt + \left(1 - \pi_0 + \pi_0 e^{-\lambda t_L^A}\right) e^{-r t_L^A} V^* \right] \\
+ (1-q) \cdot & \left[\int_0^{t_H^A} \pi_0 \lambda e^{-(r+\lambda)t} (y_A + y_P) dt + \left(1 - \pi_0 + \pi_0 e^{-\lambda t_H^A}\right) e^{-r t_H^A} V^* \right].
\end{aligned} \tag{1.3}
$$

Note that the agent's switching time is increasing in the bonus b, which implies $t_L^A < t_H^A$. Since the principal prefers the agent to switch quicker, she prefers the case $b = b_L$ over the case $b = b_H$. If the principal could choose b_L, b_H and q independently of each other, she would maximize q and set b_L to a value that maximizes V_P through equation (1.2). However, q is implicitly determined by b_L and b_H through equation (1.1), with $q \in (0,1)$ for any finite b_L and b_H. Despite this constraint, the intuition of focusing on the the low-reward case and making it as likely as possible is key to making opaque schemes perform well.

Suppose that we fix b_L at a given level and try to maximize q. The only variable the principal can change is b_H. Increasing b_H increases the third term in equation (1.1), and thus allows the principal to redistribute some probability from $b = b_H$ to $b = b_L$. Intuitively, the principal is leveraging a larger high bonus to shift probability to the more beneficial case while keeping the agent indifferent.

Note that the agent's continuation value is increasing and unbounded in b_H.[15] Hence, the principal can increase q all the way to $q = 1$ by making b_H arbitrarily large. Figure 1.1 illustrates this by depicting the probability q for a range of b_L as we increase b_H further and further. In the limit, we have $q \to 1$ for any b_L.

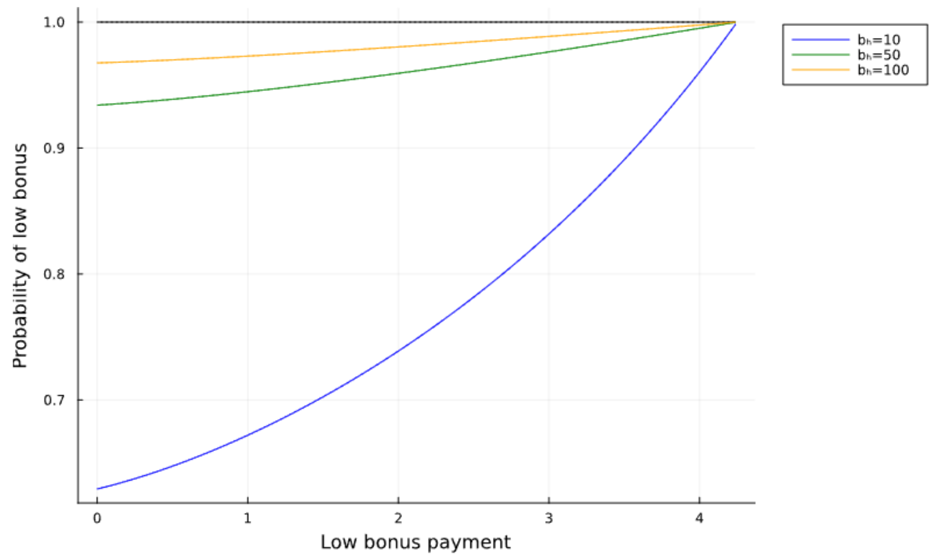
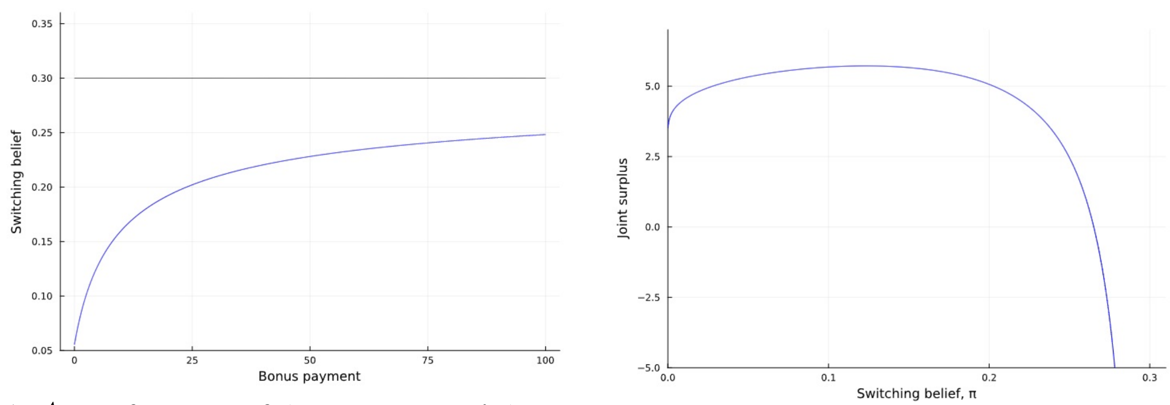


Figure 1.1: Probability of low bonus, q, as a function of b_L for various levels of b_H. The agent's value is fixed at $V_A = \bar{u}$. Parameters: $\pi_0 = 0.3$, $\lambda = r = 1$, $y_A = 5$, $y_P = 10$, $c = 0.5$, $\bar{u} = 1$.

Let us now examine the principal's value under an extreme opaque bonus scheme. Since the principal's value can be expressed as $V_P = V^* - \bar{V}_A$, maximizing the principal's value with a binary bonus scheme is equivalent to maximizing joint value V^* while holding the agent's value at $\bar{V}_A$. Note that joint value in equation (1.3) is unaffected by transfers between the players; it is only affected by the agent's switching time, t_L^A and t_H^A. To see the effect of b_H on these rules, we can directly apply Lemma 2. Intuitively, an increase of the bonus payment on the current arm (holding an unexplored arm's value fixed) will make sticking to the current arm more attractive. This prolongs experimentation on the current arm, increasing $t^A(b, \bar{V}_A)$.

[15]This would still be true if the agent was mildly risk-averse, e.g. by having a concave but unbounded von Neumann-Morgenstern utility function.

We illustrate the effect of bonus payment on experimentation graphically in Figure 1.2. Figure 1.2(a) depicts the agent's optimal switching belief given bonus payment b on the current arm and holding the overall value of the agent fixed at $V_A = \bar{u}$. Figure 1.2(b) depicts the joint surplus as a function of the agent's switching belief, which is affected by the bonus payment on the current arm. That is, the figure depicts $V_A + V_P$ given any possible switching belief $\pi^A \in [0, \pi_0]$,[16] which is affected by the current arm's bonus. Note that the joint value is single-peaked in the switching belief, with the peak corresponding to the efficient switching policy.



(a) π^A as a function of the current arm's bonus b, holding V_A (the value of a new arm) fixed.

(b) Joint surplus as a function of π^A.

Figure 1.2: Joint surplus and switching policy as a function of the bonus payment on the current arm, holding the value of a new arm fixed. Parameters: $\pi_0 = 0.3$, $\lambda = r = 1$, $y_A = 5$, $y_P = 10$, $c = 0.5$, $\bar{u} = 1$.

Figure 1.3 on the next page combines these two graphs and depicts the joint surplus (the blue curve) as a function of the bonus payment on the current arm. Note that the agent's overall value is kept at $V_A = \bar{u}$ in both the opaque and the transparent bonus schemes. In the graph, the transparent bonus scheme is denoted by the red line at the bonus payment $b = \bar{b}$. It is clear that the joint value is not maximized at that point, as the agent's switching time is too large. However, pushing b further down would violate

[16] $\pi^A > \pi_0$ is impossible due to the agent's posterior belief decreasing over time in absence of success.

the agent's participation constraint, since it both reduces the current arm's reward and the overall value V_A. The opaque bonus scheme can resolve this tension.

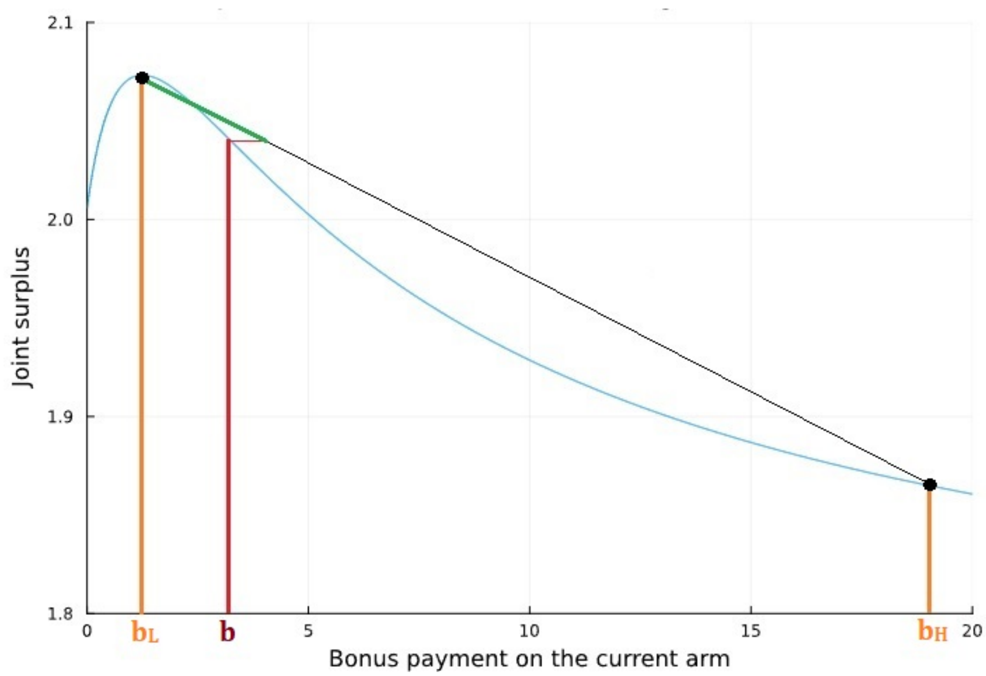


Figure 1.3: Simulation of the joint continuation value of a current arm. The orange lines describe an opaque bonus scheme with payments b_L and b_H. Parameters: $\pi_0 = 0.3$, $\lambda = r = 1$, $y_A = 5$, $y_P = 10$, $c = 0.5$, $\bar{u} = 1$.

Figure 1.3 illustrates one such bonus scheme with low reward b_L and high reward b_H. For any $b_L < \bar{b}$ and $b_H > \bar{b}$, there exists a probability distribution on $b \in \{b_L, b_H\}$ such that the agent's expected value is exactly equal to $\bar{u}$:

$$\bar{u} = -c + \mathbb{P}(b_L)V_A(b_L, \bar{u}) + (1 - \mathbb{P}(b_L))V_A(b_H, \bar{u}),$$

where

$$V_A(b, \bar{u}) = \int_0^{t^A(b,\bar{u})} \pi_0 \lambda e^{-(r+\lambda)t}(y_A + b_L)dt + \left(1 - \pi_0 + \pi_0 e^{-\lambda t_L^A}\right) e^{-rt_L^A} \bar{u}.$$

The principal can manipulate the values of b_L and b_H while keeping the agent indifferent between experimentation and outside option. The expected joint value of the opaque bonus scheme is somewhere between the continuation values at $b = b_L$ and $b = b_H$, which is depicted by a green and a black dashed line segments in Figure 1.3. The expected joint value is higher under the opaque bonus scheme than the transparent one as long as it is located on the green segment, i.e. there is enough probability placed on $b = b_L$.

Now suppose that the principal lets b_H get arbitrarily large, $b_H \to \infty$. The agent's continuation value $V_A(b_H, \bar{u})$ is increasing in b_H, which means that the principal keeps shifting more probability to $b = b_L$ to keep the agent's overall value at $\bar{u}$. In the limit, $V_A(b_H, \bar{u})$ becomes arbitrarily large, and thus the probability of b_L gets arbitrarily close to 1:

$$\mathbb{P}(b_L) = \frac{V_A(b_H, \bar{u}) - \bar{u} - c}{V_A(b_H, \bar{u}) - V_A(b_L, \bar{u})} \to 1 \text{ as } V_A(b_H, \bar{u}) \to \infty.$$

However, the joint continuation value $V_{joint}(b_H, \bar{u}) = V_A(b_H, \bar{u}) + V_P(b_H, \bar{u})$ approaches a finite number. To understand this, note that b_H is simply a transfer between the players that does not affect the joint value. The only way in which b_H affects the joint value is through the agent's switching time, $t^A(b, \bar{u})$. As $b_H \to \infty$, we have $t^A(b_H, \bar{u}) \to \infty$, and the joint continuation value under $b = b_H$ approaches

$$V_{joint}(b_H, \bar{u}) \to -c + \frac{\pi_0 \lambda}{\lambda + r}(y_A + y_P).$$

Together with $q \to 1$, this implies that, as $b_H \to \infty$, the expected joint value approaches $V_{joint}(b_L, \bar{u})$. Referring back to Figure 1.3, this limiting behavior means that the expected joint value eventually enters the green line segment. Since the agent's value is kept at $\bar{u}$, the principal's value approaches $V_{joint}(b_L, \bar{u}) - \bar{u}$. Intuitively, the principal's value in the limit is determined by the agent's outside option and the joint surplus given the low-bonus switching policy.

The next result describes the effect of increasing b_H on the principal's value more precisely.

Theorem 2. *Consider a binary bonus scheme* $F = (b_L, q;\, b_H, 1-q)$ *with fixed* b_L *and* $V_A(F) = \bar{V}_A$. *There exists* $B > b_L$ *such that* $V_P(F)$ *is increasing in* b_H *for any* $b_H \geq B$. *Moreover, the global maximum of* $V_P(F)$ *over* b_H *is attained when* $b_H \to \infty$.

Figure 1.4 illustrates this result by depicting the principal's value of an opaque bonus scheme as a function of b_H.[17] Bonus b_L is calibrated to maximize the joint continuation value, as in Figure 1.3.

We can see that the opaque scheme's value is initially decreasing. This is due to the joint surplus decreasing relatively rapidly (as can be seen in Figure 1.3). That is caused by the local concavity of the joint continuation value function. However, for sufficiently high b_H the value function becomes convex and flattens, which allows the effect of shifting probability from b_H to b_L to prevail. This shift of probabilities is responsible for the continual increase of the opaque scheme's value thereafter.

Figure 1.4 also shows that the induced principal's value is not convex in the bonus b. Increasing b_H while maintaining the agent's participation constraint effectively results in an increase of the bonus payment's variance. However, it can help or hurt the principal's value depending on how large the increase in b_H is. We can also see that an opaque bonus scheme improves upon the transparent bonus scheme only if the high bonus is sufficiently large.

1.3.4 Efficiency of opaque bonus schemes

We now turn attention towards the question of how well opaque bonus schemes perform in this setting. It is useful to compare them to the best outcome for the princi-

[17] Recall that for all values of b_H, probability q is adjusted to keep $V_A(F) = \bar{u}$.

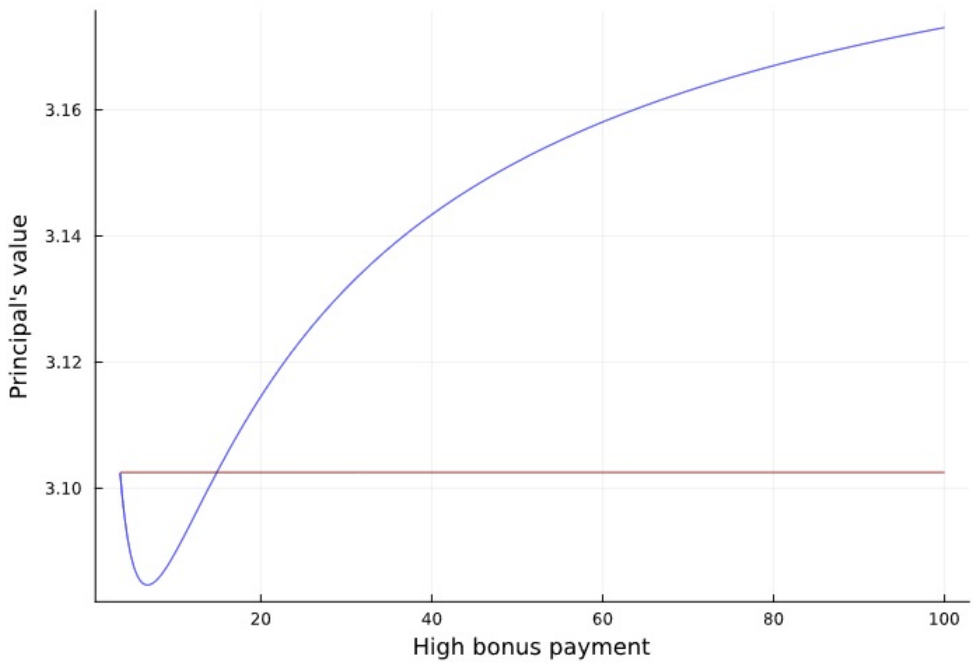


Figure 1.4: Simulation of the principal's value of a transparent scheme (red line) and an opaque bonus scheme (orange line), as a function of b_H. Parameters: $\pi_0 = 0.3$, $\lambda = r = 1$, $y_A = 5$, $y_P = 10$, $c = 0.5$, $\bar{u} = 1$, $b_L = 1.88$.

pal. This benchmark involves first-best switching policy (that maximizes joint surplus) and full surplus extraction from the agent (i.e., setting $V_A = \bar{u}$). Using Lemma 1, the efficient switching policy can be expressed in terms of a switching belief π^*:

$$\pi^* = \frac{m^*}{\lambda(y_A + y_P)}, \text{ where } m^* = \sup_{t \geq 0} \frac{-c + \frac{\pi_0 \lambda}{r+\lambda}(1 - e^{-(r+\lambda)t})(y_A + y_P)}{\frac{1}{r}\left(\frac{\pi_0 r}{r+\lambda}(1 - e^{-(r+\lambda)t}) + (1 - \pi_0)(1 - e^{-rt})\right)}.$$

This is a useful benchmark for bounding the best overall value that the principal can achieve. The agent's participation constraint requires him to receive at least $\bar{u} = \frac{s}{r}$ in order to engage in experimentation.[18] Therefore, the principal's overall value under any mechanism cannot exceed the joint value under the efficient switching policy minus the agent's outside value $\bar{u}$. Theorem 3 offers a condition under which this upper bound is attainable via an opaque bonus scheme.

[18]This is the discounted value of choosing the outside option forever.

Let V^* be the joint surplus under the first-best switching policy.

Theorem 3. *There exists an opaque bonus scheme with the principal's value arbitrarily close to $V_P = V^* - \bar{u}$ if and only if*

$$\bar{u} \geq \frac{\pi^* \lambda}{r + \pi^* \lambda} y_A$$

Otherwise, there exists an opaque bonus scheme with the principal's value arbitrarily close to $V_P = V^* - \frac{\pi^* \lambda}{r + \pi^* \lambda} y_A$.

Perhaps surprisingly, a key reason for this result is the non-negativity of b. To understand this, consider the agent's dynamic incentives. At every point in time, he compares the benefits associated with the current arm's success payoff $y_A + b$ and belief $\pi_t < \pi_0$ to a brand new arm with value V_A. As the agent's switching time is generally inefficiently high, the optimal bonus scheme aims to bring it down to the efficient t^* (i.e. bring π^A up to π^*). In a binary bonus scheme, this can be done by reducing b_L while keeping the agent's value fixed at $V_A = \bar{u}$. Recall that the agent's switching belief conditional on the low reward is

$$\pi^A = \frac{r \bar{u}}{\lambda (y_A + b_L - \bar{u})}.$$

However, if $\bar{u}$ is too low, the binary bonus scheme may be incapable of bringing the agent's switching time all the way to t^* while maintaining $V_A = \bar{u}$. The combined constraints of keeping $V_A(F) = \bar{u}$ and $b_L \geq 0$ may make it impossible for the principal to achieve first-best outcome, and thus lead to an inefficient switching policy.

If the principal could impose a negative transfer $b_L < 0$ (effectively taxing the agent's success), then Theorem 3 would become stronger. We provide the alternative result for this case in Section 1.5. In short, if $b_L < 0$ is possible, then binary bonus schemes can always approximate the efficient switching rule and extract all surplus.

Intuitively, this is because we remove the binding constraint that sometimes precludes this outcome in Theorem 3.

It is important to note that Theorem 3 does not imply that the principal prefers the agent to have a higher $\bar{u}$. All else equal, increase in $\bar{u}$ generally hurts the principal. Theorem 3 simply describes the condition under which the restriction to stationary bonus schemes does not prevent the principal from achieving her first-best outcome. When the agent's outside option is not sufficiently valuable, a time-contingent incentive mechanism (like a bonus payment with a deadline) can perform strictly better than the best stationary bonus scheme.

1.4 Application to YouTube

In this section, we will discuss how our results can be applied to YouTube and other online platforms that rely on decentralized content creation.

First, we need to establish that the problem faced by YouTube and its content creators is reminiscent of the model in Section 1.2. To begin, consider content creators—novice creators, to be specific. When these creators enter the platform, they have a practically unlimited number of content variations to explore. Initially, they do not know what they can be successful at and must discover this by trial and error. Once they start producing a particular kind of video, they may eventually strike success (e.g., generate a lot of views and gain a large audience), but the success is uncertain. Whenever creators switch to a new content variation, they must invest a bit of time and effort into figuring out how to make new videos. This involves purchasing new equipment if necessary, researching the existing creators in the new content niche, and generally learning how to properly make a new type of video. Overall, this supports modeling a YouTube content creator's problem as an infinite-armed bandit problem with switching costs.

Now let us turn attention to YouTube itself. The platform's main source of revenue is advertising, as explained on its website. The more consumers are on the platform, the more videos (and ads) they watch—the higher the platform's profit. Hence, YouTube is interested in having a lot of engaging content that can be supplied to consumers. For that, it needs content creators that can produce high-quality content and capture a large audience. In other words, the platform benefits from creators succeeding. However, the platform is not directly involved in the production of content and thus does not internalize a creator's switching cost. Regarding novice creators, it only cares about quickly getting them to the point where they achieve success and can provide engaging content for the consumers. This establishes a basic payoff structure that resembles the payoffs in our model. Both the platform and the creators benefit when success happens, but the creators face an additional switching cost.

YouTube can affect content creators' decisions by designing additional incentives, such as a bonus scheme. In practice, YouTube has two relevant systems that provide incentives for a novice—a recommendation algorithm and a pay-for-ads scheme. The recommendation algorithm is responsible for taking the creator's content and matching it to consumers that may be interested in watching it. All else equal, if a video is recommended more by the algorithm, the creator gets more views. The pay-for-ads scheme is the main mechanism that translates views into a monetary benefit, at least for novice creators.[19] If we consider generating (ad) views to be a measure of success on YouTube, then pay-for-ads is essentially a bonus for success that we analyze in our model. It can also be viewed as a standard pay-for-performance incentive scheme.

The previous three paragraphs establish that the model in Section 1.2, reasonably approximates the problem faced by YouTube and novice content creators. We can now

[19]Successful content creators can monetize their audiences through direct sponsorships from firms and donation services like Patreon. We focus on novice creators that do not have an audience to monetize because they are the creators whom YouTube would like to actively experiment with new content.

examine how the results in Section 1.3 apply to YouTube's problem. Taken in a normative way, they suggest that if YouTube wants to use a simple, stationary incentive scheme to motivate experimentation by novice content creators, it should employ a highly uncertain bonus scheme. This scheme should leverage a very unlikely but large reward against an overwhelmingly likely but low reward. The result is optimal experimentation by most content creators (who get unlucky with their bonus), which comes at the cost of a tiny minority of content creators that get a lucky draw of the bonus and thus remain stuck. This incentive scheme requires the platform to neither spend resources on monitoring each creator nor keep track of the content they have supplied in the past. We can also predict that this incentive scheme will result in nearly socially efficient experimentation as long as the content creators have sufficient valuable outside options.

We can also ask whether the optimal bonus scheme we derive descriptively matches the actual incentive schemes used by YouTube. Let us focus on the pay-for-ads scheme, as it is the closest match to the bonus schemes we analyze. It has been observed[20] that the payouts for ads on YouTube have a large variance based on the region where views are generated and on the content niche. Specific pay rates fall in the range of \$1 to \$12 per 1000 views, which is a substantial spread when considering a "successful" video garners hundreds of thousands of views. This uncertainty affects the incentives of novice content creators. However, there is little evidence suggesting that it has been designed that way by YouTube. The observed variability in payouts is mostly attributed to the outcomes of online advertisement auctions.[21] Some content keywords and some regions are more valuable to advertisers, which leads to higher bids on the corresponding ads. This generates variance in payouts that the creators receive.

In contrast, YouTube's bonus scheme is remarkably transparent and uniform.

[20]See this blog on Thinkific, for instance.

[21]For instance, Google's AdSense Calculator provides an ad revenue estimate based on a web page's content category and region.

Regardless of how much advertisers pay for views on a given video, YouTube gives a fixed share of that revenue to content creators. In a world where all content and all audiences generated the same ad revenue, the platform's bonus scheme would be equivalent to the deterministic bonus scheme in our model. As we show, such a bonus scheme does not provide optimal incentives for creative experimentation. Novice content creators will not try new content variations with efficient frequency because they do not fully incorporate the platform's payoff into their decision-making.

Our results suggest that YouTube can improve the incentives for creativity without relying on more complicated mechanisms, such as time-contingent bonuses or deadlines. A likely reason why YouTube employs a simple, flat pay-for-performance bonus scheme is that it is easy to scale. To provide better incentives for creativity, the platform can reduce the standard pay rate and compensate for it by offering a very high pay rate on selected rare content variations. Those content variations will be initially unknown to content creators. This uncertainty will put them in a situation where unexplored content has a higher pay rate (on average) than their current content. As a result, they will become more exploratory and find success more quickly, which will benefit the platform.

In this dissertation chapter, we have abstracted away from the fact that different projects may have different returns. For example, two content variations on YouTube may generate very different advertising revenues. Since YouTube likely knows which content is more profitable, it could use this knowledge in the design of its uncertain bonus scheme. For instance, it could skew high pay rates for creators towards content variations that typically deliver high advertising revenue. However, such micromanagement may require additional resources and costs for the platform. In addition, creators are already incentivized to seek content that offers a high advertising revenue by the virtue of receiving a fixed share of that larger amount. Skewing the uncertain bonus scheme towards high-paying content may thus be unnecessary.

1.5 Extensions

In this section, we will briefly analyze a few important extensions of the model. We will begin with considering the case where the principal can set a negative bonus payment b. We will then analyze a simple moral hazard extension of the model and examine a new form of incentive—switching subsidy.

1.5.1 Negative transfers.

In this section, we will briefly explore what happens if the principal can fix a negative bonus b. This can be interpreted as the principal taxing the agent's profit or being able to choose. Does relaxing the design of the bonus scheme in this way affect the outcome?

We can show that Theorem 1 still holds, i.e. opaque bonus schemes are better than transparent ones. The proof relies on the principal being able to offer a lower bonus than in the transparent scheme under consideration, which is still the case when we relax the constraint $b \geq 0$. Similarly, Lemma 3's proof is unaffected, which means that we can restrict attention to binary bonus schemes.

The main change is observed when we get to Theorem 3. The inequality constraining efficiency of the optimal binary bonus scheme exists precisely because the principal cannot tax the agent. If we remove this constraint, we get a stronger result.

Theorem 4. *There exists an opaque bonus scheme with the principal's value arbitrarily close to $V_P = V^* - \bar{u}$.*

With $b \geq 0$ constraint removed, the principal can always achieve the first-best outcome using opaque bonus schemes. In other words, if the principal can tax the agent's profit (even without causing them to go into debt), then she can extract all surplus. Notably, this result is maintained if we restrict $b \geq -y_A$, imposing a debt-free condition on the agent. The optimal bonus scheme does not involve the agent having to pay the principal at any point.

To understand this, recall that the agent-optimal switching belief with bonus b on the current arm is equal to

$$\pi^A = \frac{rV_A}{\lambda(y_A + b - V_A)},$$

where V_A is the value of an unexplored arm (given the optimal experimentation policy). Recall that we typically have $\pi^A < \pi^*$ due to the agent not fully accounting for the principal's success payoff. When the principal is designing the optimal opaque reward scheme, she is trying to get π^A up to π^* under the low bonus b_L, and at the same time get V_A to equal $\bar{u}$. If b_L is allowed to be negative (down to $-y_A$), the principal can make the current arm as unappealing for the agent as they want, which can push π^A as high as she wants. This ensures that the efficient switching under $b = b_L$ is achievable. With the optimal bonus scheme pushing the probability of b_L to 1, the result is full surplus extraction ($V_A = \bar{u}$) and maximization of joint surplus ($\pi^A = \pi^*$) with probability 1.

Theorem 4 once again highlights that the importance of the agent's outside option hinges upon the principal not being able to tax the agent's payoff. This can be a practical restriction; for instance, YouTube cannot tax the financial support creators get from ad sponsorships or services like Patreon. Achieving efficient experimentation by novice creators is thus dependent on the value of the creators' opportunity costs.

1.5.2 Moral hazard and switching subsidy

Suppose that the agent has access to projects (bandit arms) of two distinct types, *expensive* (type A) and *cheap* (type B). Expensive arms are the same as in the main model: they have prior probability π_0 of being good, deliver success at rate λ conditional on being good, and have switching cost $c_A > 0$. Cheap arms are different: they are bad with probability 1 and have switching cost $c_B > 0$, such that $c_B < c_A$. Intuitively, these are garbage variations that the agent can start working on with almost no effort, but this comes at the cost of success chance. The principal cannot distinguish between expensive and cheap arms, but the agent can. The agent still has access to an outside option that offers a safe flow payoff $s > 0$. At every time moment,a he now chooses between four actions instead of three: (i) continue work on the current arm; (ii) switch to an expensive arm; (iii) switch to a cheap arm; (iv) take the outside option.

We wish to investigate how the presence of cheap arms affects the optimal bonus scheme. We also want to see how they affect a new incentive—a switching subsidy. Suppose the principal offers a variation of the bonus scheme from Section 1.2. She promises a bonus payment w_b upon success *and* an additional payment w_s paid whenever the agent switches to a new arm. We will refer to this as *bonus scheme with switching subsidy*.

We will show that the presence of cheap arms constrains the use of the switching subsidy, but does not impede the effectiveness of opaque bonus schemes.

Lemma 4. *In equilibrium, we must have $w_s \leq c_B$.*

Note that if the principal sets $w_s > c_B$, the agent's best response will be to instantaneously switch to a cheap arm regardless of p. This offers him an immediate positive payoff (the net switching subsidy), which outweighs any flow expected payoff he can get from working on an arm, or even from taking the outside option. Instantaneous switching to cheap arms will result in infinitely negative expected value for the principal. Hence, the presence of cheap arms with a low switching cost constraints the principal in her reward design. Intuitively, this is because the switching subsidy incentivizes the agent to switch between arms *regardless* of their success probabilities, since the principal cannot tell the arms apart.

Given this constraint, we can ask whether opaque bonus schemes (with stochastic w_b) can still improve upon a transparent one (with deterministic w_b). The answer hinges on the values of the agent's outside option and the switching costs. Recall that the efficient switching belief π^* is defined by

$$\pi^* = \frac{m^*}{\lambda(y_A + y_P)}, \text{ where } m^* = \sup_{t \geq 0} \frac{-c_A + \frac{\pi_0 \lambda}{r+\lambda}(1 - e^{-(r+\lambda)t})(y_A + y_P)}{\frac{1}{r}\left(\frac{\pi_0 r}{r+\lambda}(1 - e^{-(r+\lambda)t}) + (1-\pi_0)(1 - e^{-rt})\right)}.$$

Further, let $\hat{w}_s$ and $\hat{w}_b$ be the transparent bonus scheme with switching subsidy that makes the agent indifferent between experimentation and outside option while minimizing bonus $\hat{w}_b$. That is, $\hat{w}_s$ and $\hat{w}_b$ solve $\min \hat{w}_b$ subject to $\hat{w}_b \geq 0$, $\hat{w}_s \leq c_B$ and $\hat{m} = s$, where index $\hat{m}$ is defined by

$$\hat{m} = \sup_{t \geq 0} \frac{-c_A + \hat{w}_s + \frac{\pi_0 \lambda}{r+\lambda}(1 - e^{-(r+\lambda)t})(y_A + \hat{w}_b)}{\frac{1}{r}\left(\frac{\pi_0 r}{r+\lambda}(1 - e^{-(r+\lambda)t}) + (1-\pi_0)(1 - e^{-rt})\right)}.$$

In essence, bonus scheme $(\hat{w}_s, \hat{w}_b)$ makes the agent indifferent while ensuring that the bonus payment is as low as possible. We then have the following result.

Theorem 5. *The optimal the bonus scheme with switching subsidy is opaque if and only if*

$$\hat{w}_b > 0 \text{ and } \pi^* > \frac{r\bar{u}}{\lambda(y_A + \hat{w}_b - \bar{u})}.$$

Otherwise, a transparent bonus scheme with switching subsidy $(\hat{w}_s, \hat{w}_b)$ *is also optimal.*

It is important to note that even when the transparent scheme $(\hat{w}_s, \hat{w}_b)$ is optimal, there exist opaque bonus schemes that can perform arbitrarily close to it. Hence, a transparent bonus with switching subsidy never outperforms opaque bonus schemes; however, it can sometimes perform worse. The key reason is the existence of cheap arms which bounds w_s from above. To understand this, consider the second inequality in Theorem 5. The right-hand side is the lowest switching belief that a transparent bonus scheme with switching subsidy can achieve. If $\pi^* > \frac{\bar{u}}{\lambda(y_A+\hat{w_b})}$, then the joint value (and the principal's as a result) can be improved by pushing the switching belief up. If $\hat{w_b} > 0$, opaque bonus schemes can induce a lower bonus, which pushes the switching belief up. Thus, transparent bonus schemes are suboptimal whenever both inequalities hold.

These inequalities are more likely to hold when the switching cost of cheap arms is lower. A lower c_B puts a stronger restriction on $\hat{w_s}$, which in turn makes $\hat{w_b}$ higher. This causes the switching belief $\frac{r\bar{u}}{\lambda(y_A+\hat{w_b})}$ to increase, making the first inequality more likely to hold. Hence, we should not expect to see switching subsidies in environments where the agent has access to very cheap projects that are extremely unlikely to succeed.

1.6 Conclusion

This dissertation chapter has studied the question of how one can motivate creativity when being creative is costly. Creativity (i.e., trying new ideas and variations) is at the core of many professions, from research to creation of art and, more recently, content in online platforms. Finding the optimal way to motivate creativity using simple reward schemes is at the heart of our analysis.

We study a principal-agent, continuous-time model of experimentation with switching costs and infinitely many exponential bandit arms. A switching cost is key to the setting as a way to model the cost of creativity (switching to new arms). However, the infinite set of bandit arms is more of a convenient assumption for modeling purposes. Our insights would remain largely unaffected if we considered a model with finitely many arms. Another important aspect of the model is the asymmetry of switching costs, which is what creates the conflict of incentives between the principal and the agent.

Our key finding is that the optimal bonus scheme uses uncertainty to motivate creativity. The principal can exploit an extremely unlikely but large reward in order to induce the optimal experimentation policy under the overwhelmingly likely low reward. This bonus scheme relies on the agent's continuation value being increasing and unbounded in the bonus. Despite the bonus scheme being time-independent and simple, it can achieve the principal's first-best outcome as long as the agent's outside option is sufficiently valuable. Moreover, if the principal can tax the agent's payoff from an arm's success, then the first-best is always achievable.

Our results can guide the design of rewards in online platforms and in other settings where stimulating creativity is important. If there is a cost to designing complicated time-contingent incentive schemes, creating highly uncertain stationary reward schemes is preferable. We also note that the existing uncertainty in the incentive schemes used by YouTube can motivate novice content creators to try new content variations more frequently. Regardless of intent, this uncertainty can contribute to the platform's end goal of increasing the supply of engaging content for its viewers.

Chapter 1 is currently being prepared for submission for publication of the material. The dissertation author, Danil Dmitriev, is the sole author of this chapter.

Chapter 2

Sabotage-Proof Mechanism Design

2.1 Introduction

Polls – especially those conducted online – are notorious for their lack of robustness to sabotage; the derailment of online polls by internet trolls[1] is a common and well-documented occurrence, often producing amusing news articles but also potentially large costs. Take for instance the 2015–2016 New Zealand flag referendums, where an online poll was used to crowdsource a replacement for the country's national flag. The public gallery of flag submissions quickly became inundated with ridiculous, unusable flags. This process took well over a year, cost approximately $NZ 26 million, and was ultimately fruitless in producing a new flag. Similar derailments have interfered with crowdsourcing in marketing campaigns (BBC, 2016), information-gathering during the 2020 US Presidential Election (Collins and Popken, 2019; Frenkel et al., 2020), and even prevented the government of North Macedonia from properly counting its own population (*The Economist*, 2020).

How does one optimally design voting mechanisms in the presence of saboteurs

[1] A "troll" in this context is someone who is deliberately trying to derail a poll or any other mechanism

or trolls?[2] This question can be split into two parts: how does one design the entry part of the mechanism to encourage normal agents and dissuade trolls, and how does one design the voting part of the mechanism given a fixed population of participants? We aim to answer both questions in this project. However, so far our results focus on the second part. It is a natural place to start the analysis and proceed to the entry part via backward induction. For now, we consider a situation where the entry has already occurred, and analyze the designer's problem given a fixed population of participants. In this framework, we focus on analyzing a few benchmark mechanisms and characterizing the optimal mechanism as completely as possible.

We start by analyzing a simple illustrative example with two genuine agents and one troll. Each genuine agent has a private type corresponding to a bliss point over actions the mechanism designer can take. The mechanism designer's objective is to maximize the welfare of the genuine agents. There is a common prior over their types. The designer gathers information through a poll and then takes an action. The troll's objective is to minimize the welfare of the genuine agents. We consider two specific mechanisms as reasonable baselines — "majority rule" and "average-of-votes" — for both methodological and empirical reasons. Choosing the outcome that was voted for on average is theoretically optimal assuming that trolls are absent, utility is quadratic, and that it is indeed possible to average over votes.[3] On the other hand, choosing the outcome with the most votes is by and large the most widely used in polls and elections. Our project offers insights into how suboptimal these two mechanisms are in the presence of trolls.

Analyzing the agents' equilibrium behavior in the illustration below, we derive the welfare implications for both mechanisms and compare with the benchmark of do-

[2] Here, we use "troll" and "saboteur" interchangeably.

[3] For instance, this is possible in the Weber (1929) Problem and in other facility location problems that followed.

ing no mechanism ("no-poll benchmark"). In the example, "majority rule" performs exactly as the no-poll benchmark, whereas "average-of-votes" (a more fine-tuned mechanism) performs better when ex-ante uncertainty over types is high, but worse when it is low.

We then consider the finite analysis of our two baseline mechanisms, "majority rule" and "average-of-votes", in a more general setting with a finite number of agents and trolls. We derive similar predictions as in the Illustration, and show that when ex-ante uncertainty is low, the "average-of-votes" mechanism will perform poorly compared to a blind mechanism. Intuitively, this happens because under low ex-ante uncertainty there is little potential gain from gathering information from the agents, but the negative impact of trolls is still present in its full force. This suggests that even under a fixed population of agents and trolls, a blind mechanism may still be better than running an informative mechanism when there is little information the designer can potentially gain.

Additionally, we derive some properties of the optimal mechanism given a fixed population of participants. First, we show that the optimal mechanism has to satisfy a *quasi-monotonicity* property, which basically requires that more votes for a given type result in an action that is closer to that type's optimal action. Interestingly, this property is always violated by the "majority rule" mechanism, which indicates that it is never optimal. Next, we show that the optimal mechanism also has to satisfy an *indifference* property, which requires the trolls to be indifferent between their messages. This allows us to reduce the search for the optimal mechanism to a narrow family of mechanisms that fulfill the property. It also allows us to generally rule out the "average-of-votes" as the optimal mechanism.

Finally, we consider the limit case where the number of trolls is arbitrarily large and derive a worst-case result applicable to any (continuous) mechanism. This can model a situation where trolls have very low costs of entry (and possibly submitting multiple votes). We show that given any continuous mechanism, trolls can achieve the worst-case outcome under it if they are sufficiently numerous. A natural corollary follows: if number of trolls is potentially unlimited, the best mechanism for the designer to implement is a *blind* (or no-poll) mechanism, which ignores messages from agents and chooses the ex-ante best outcome. This speaks to the observed tendency of online polls being shut down or cancelled when a large influx of trolls occurs, and suggests that such action may indeed be optimal in these circumstances.

At the end of the paper, we discuss these insights in more detail and outline future extensions. Most notably, we plan to consider the entry part of the designer's problem and see how the presence of trolls affects the optimal population of agents that the designer wants to attract. We also plan to analyze environments where the designer's preferences are not perfectly aligned with the genuine agents'. In this scenario, it is plausible that the trolls' desire to hurt the designer could inadvertently improve the genuine agents' welfare. For example, in an auction setting, saboteurs could decrease the auctioneer's expected revenue, which may be beneficial to the genuine bidders.

Literature Review Our focus in this paper is on the design of *polling* mechanisms in the presence of *adversarial saboteurs* among the voters. To our knowledge, there are papers that incorporate a strict subset of these considerations, but not all of them. For instance Chorppath and Alpcan (2011), Liu et al. (2017), Yang et al. (2017), Brahma et al. (2022), and Jiang et al. (2022) focus on mechanism design with malicious/adversarial agents in *non*-voting settings. In the literature on electoral competition, (Invernizzi, 2020) studies sabotage *within* parties and (Hirsch and Kastellec, 2022) studies studies sabotage between parties. We instead focus on the *other* side of the ballot-box (voters) in *polls* (i.e. without strategic candidates).

Our paper is also related to the literature on the design of false name-proof[4] mechanisms when *anonymous* agents can participate more than once (e.g. by creating multiple identifiers, botting, etc.). Such mechanisms were originally studied in *combinatorial-auction* settings (Yokoo (2003, 2008), Yokoo et al. (2001, 2004, 2006), and Rastegari et al. (2007)) and more recently in voting games albeit without saboteurs (Conitzer (2008), Bachrach and Elkind (2008), Aziz et al. (2011), Elkind et al. (2011), and Fioravanti and Massó (2022)).

Our paper is also related to Lambert and Shoham (2008), who studies how to design a *survey* mechanism that elicits truthful opinions, and to Gary-Bobo and Jaaidane (2000), who study a polling mechanism design problem. However, neither paper allows for saboteurs amongst the voting population as in this paper. Sabotage is a consideration more often seen in (dynamic and static) contests (discussed in Chowdhury and Gürtler (2015)). Most relevant is Ishida (2012), which considers the problem of designing sabotage-proof dynamic contests.

2.2 Illustration

Suppose there are three agents, $i \in \{1,2,3\}$. Agents 1 and 2 are "genuine" or "normal" (interchangeable) agents. Each of them has a type $\theta_i \in \{\gamma_1, \gamma_2\}$, with an i.i.d. prior distribution characterized by $\mathbb{P}(\theta_i = \gamma_1) = p$, where $p \in (0;1)$. For simplicity, assume $\gamma_1 = 1$ and $\gamma_2 = 2$. Agent 3 is a "troll" or "saboteur" (interchangeable) and will be described below.

There is a mechanism designer who wants to maximize the well-being of agents

[4] I.e. mechanisms where agents do not have an incentive to participate more than once, even if they are able to do so.

1 and 2. The utility function of an agent of type θ_i is given by

$$u_i(a,\theta_i) = -(a-\theta_i)^2,$$

where $a \in \mathbb{R}$ is the *expected* action taken by the designer.[5] The objective function of the designer is given by

$$V(a) = \mathbb{E}\left[\sum_{i=1}^{2} u_i(a,\theta_i)\right].$$

Agent 3 is a "troll", or a "saboteur", whose goal is to reduce the well-being of agents 1 and 2 as much as possible. He does not know the agents' types, but knows the prior distribution, just as the mechanism designer. In a sense, his objective is entirely opposite of the designer's objective.[6] Hence, no matter what mechanism the designer creates, agent 3 will participate in a way that ex-ante minimizes $V(a)$. Agents 1 and 2 are aware of this and can take it into account when choosing how to behave.

In order to maximize the well-being of agents 1 and 2, the designer can create a *mechanism*, which consists of a message set M and an outcome rule $x: M^3 \to \mathbb{R}$. We limit our attention to direct mechanisms in which agents report their types, i.e. $M = \{1,2\}$. The choice of a mechanism then boils down to choosing the outcome function.

There are two baseline mechanisms we will consider. The first will be referred to as "majority rule", where the designer implements the action equal to the mode of observed messages (and randomizes in case of a tie). It closely matches the design of online polls discussed in the Introduction. The second mechanism will be referred to as "average-of-messages rule", where the designer implements the action equal to the average of observed messages.[7] This mechanism is optimal in the absence of trolls,

[5] Notice that these agents are risk *neutral*. If a had instead represented the designer's *realized* action, these agents would be risk *averse*.

[6] Note that the designer does not care about the troll's well-being. One way to interpret this assumption is that the troll comes from outside the population of agents that the designer cares about, e.g. a foreigner participating in a poll about purely domestic matters.

[7] Recall that $\gamma_1 = 1$ and $\gamma_2 = 2$, so messages are just real numbers.

given the quadratic-loss utility function of the normal agents.

As we will see, both mechanisms have their comparative strengths and weaknesses when it comes to solving the designer's problem. In short, the majority rule is less susceptible to the influence of trolls, since they have to be pivotal in order to affect the outcome. On the other hand, the average-of-votes rule incorporates more information from the normal agents and has the potential to better match the average type of the agents. However, that potential can be limited by the increased influence of trolls, who no longer need to be pivotal in order to affect the outcome. In fact, we show that this mechanism might perform worse than doing no mechanism at all — a "blind mechanism" benchmark, where the designer always takes the ex-ante best action. This occurs when there is little information that can be gained through the poll to begin with, in which case the negative influence of trolls outweighs the positive gain from more information.

The next two subsections outline a detailed analysis of the two baseline mechanisms.

Majority rule

As implied by its name, this outcome rule selects the *mode* of the received messages when the mode is unique; otherwise, we assume that the outcome rule uniformly randomizes between the choices tied for first. Formally, if $\mathbf{m}$ represents the vector of observed messages, then

$$x(\mathbf{m}) \equiv \mathcal{U}\{\text{mode}(\mathbf{m})\}. \tag{2.1}$$

How do the genuine agents and the troll behave in this mechanism?

Lemma 5. *Fix mechanism $M = \{1, 2\}$ and $x(\mathbf{m}) \equiv \mathcal{U}\{\text{mode}(\mathbf{m})\}$, and assume that genuine agents always break indifference in favor of telling the truth. Then any BNE of the resulting* *game is for genuine agents to tell the truth and for the troll to tell either $m_3 = 1$ or $m_3 = 2$ (he is indifferent).*

The proof of this lemma can be found in Appendix A.2.1. Intuitively, any agent's vote matters only when that agent is pivotal. For genuine agents, that means that their decision matters only when the other agent and the troll split votes, in which case the agent strictly prefers to tell the truth.[8] For the troll, his decision matters only when the genuine agents are (truthfully) splitting the vote, in which case he is indifferent between saying 1 (and hurting type 2) or saying 2 (and hurting type 1). Importantly, this is not the case when there is more than one troll, and we plan to consider this case in subsequent work.

Let us now consider the welfare implications of the "majority rule" mechanism. Given the BNE described in Lemma 5 (for argument's sake, assume $m_3 = 2$), the ex-ante welfare of agents is equal to

$$V_{MVW} = p^2 \cdot 0 + 2p(1-p) \cdot \left(-(0)^2 - (1)^2\right) + (1-p)^2 \cdot 0 = -2p(1-p) \tag{2.2}$$

Our benchmark for welfare is the no-poll scenario, under which the designer does not create any mechanism and simply takes the ex-ante best action. That action should maximize the objective function,

$$V_{NP}(a) = -p^2 \cdot 2(a-1)^2 - 2p(1-p) \cdot \left((a-1)^2 + (a-2)^2\right) - (1-p)^2 \cdot 2(a-2)^2 .$$

which has the corresponding first order condition,

$$-4p^2(a-1) - 4p(1-p) \cdot (2a-3) - 4(1-p)^2(a-2) = 0$$

[8] Our assumption about the indifference-breaking rule eliminates nonsensical equilibria where all agents always say the same message.

Solving this first order condition for a yields $a = 2 - p$ as the designer's ex-ante best action.[9]

Under this action, we can show that the agents' welfare (after a few algebraic simplifications) is given by

$$
\begin{aligned}
V_{NP} &= -2p^2(1-p)^2 - 2p(1-p)((1-p)^2 + p^2) - 2(1-p)^2 p^2 \\
&= -2p(1-p).
\end{aligned}
$$

Note that this is exactly the same as V_{MVW}. This implies that running a "majority rule" mechanism leads to the same welfare as running no mechanism at all! The presence of the troll completely nullifies the effectiveness of the mechanism in conveying information to the designer.

Average-of-votes

This outcome rule simply takes the average of the messages received by the mechanism designer. Formally,

$$x(\mathbf{m}) \equiv \frac{1}{N}\sum_{i=1}^{N} m_i \tag{2.3}$$

Similarly to the "majority rule" discussion, let us find the equilibrium in the resulting game between the genuine agents and the troll.

Lemma 6. *Fix mechanism $M = \{1, 2\}$ and $x(\mathbf{m}) \equiv \frac{1}{N}\sum_{i=1}^{N} m_i$, and assume that genuine agents always break indifference in favor of telling the truth. Then the unique BNE of the resulting*

[9] Note that $\frac{\partial^2}{\partial a^2} V_{NP}(A) = -4p^2 - 8(1-p)p - 4(1-p)^2$

game involves the genuine players telling the truth and the troll playing the following strategy:

$$m_{3T}^{*} = \begin{cases} 1, \ \textit{if}\ p < \frac{1}{2} \\ \{1,2\}, \ \textit{if}\ p = \frac{1}{2} \\ 2, \ \textit{if}\ p > \frac{1}{2}. \end{cases}$$

The proof of this lemma can be found in Appendix A.2.2. Intuition behind it is similar to that of Lemma 5, with the exception that now the troll is not indifferent between messages because he is able to affect the outcome in all cases (as opposed to only the cases where he is pivotal in the "majority rule" mechanism).

Let us now consider the welfare implications of the "average-of-votes" mechanism. Assume $p > \frac{1}{2}$ (to pin down the exact message of the troll). Given the BNE described in Lemma 6 , the ex-ante welfare of the genuine agents is equal to

$$\begin{aligned} V_{AM} &= p^2 \cdot 2\left(-\left(\frac{1}{3}\right)^2\right) + 2p(1-p)\cdot\left(-\left(\frac{2}{3}\right)^2 - \left(-\frac{1}{3}\right)^2\right) + (1-p)^2 \cdot 0 \\ &= -\frac{2}{9}p^2 - \frac{10}{9}p(1-p) = -\frac{2}{9}p\,(5-4p)\,. \end{aligned}$$

Let us now compare it to the no-poll benchmark. We already know that under it the designer's problem is exactly the same as the one already considered in the "majority rule" discussion, so the optimal action is $a = 2 - p$ and the attained welfare is

$$V_{NP} = -2p(1-p).$$

When is this outcome better than the one provided by "average-of-votes" mechanism? Note:

$$V_{AM} < V_{NP} \ \Leftrightarrow\ -\frac{2}{9}p\,(5-4p) < -2p(1-p) \ \Leftrightarrow\ p > \frac{4}{5}.$$

That is, if p is sufficiently high, the no-poll benchmark (as well as the "majority rule" mechanism) provides higher welfare than the "average-of-votes" mechanism. Intuitively, this happens because there is little ex-ante uncertainty over the distribution of types of the genuine agents, meaning that the no-poll benchmark performs relatively well. This also means that there is little information that could potentially be gained from running the "average-of-votes" mechanism, while the negative effect of the troll's presence still remains in full force. Therefore, if ex-ante uncertainty over the type distribution is sufficiently low, the "average-of-votes" mechanism loses to the no-poll benchmark and to the "majority rule" mechanism. On the other hand, if the ex-ante uncertainty is sufficiently high, there is a lot of information to be gained from the "average-of-votes" mechanism, so it is worth choosing it over the considered alternatives.

2.3 Model

In this section, we introduce a general model of voting mechanism design with a finite number of voters and trolls (or saboteurs). We focus on the case of two types in the interest of clearly presenting our results.[10] We will partially characterize the optimal mechanism by showing that it must satisfy a particular "indifference property" for the saboteurs. We will also analyze the performance of two benchmark mechanisms — *majority rule* and *average-of-votes rule*. Finally, we will finish by discussing how these benchmark mechanisms can be improved in simple ways to account for the presence of the saboteurs.

[10]Our main results in subsections 2.3.2 and 2.3.3 can be straightforwardly extended to the case with an arbitrary, finite number of types.

2.3.1 Setting

There is a designer that can take a public action $a \in \mathbb{R}$ and N "genuine"/"normal" voters. Each voter i has a type $\theta_i \in \{\gamma_1, \gamma_2\} =: \Gamma$ that is i.i.d. with $\mathbb{P}(\theta_i = \gamma_1) = p$. This is common knowledge. A voter of type θ has a standard quadratic-loss utility function

$$u(a, \theta) = -(a - \theta)^2.$$

We make the assumption of a specific functional form for tractability of analysis. In general, we can assume any single-peaked utility function (so that it has a bliss point).

The designer's objective function is to maximize the expected aggregate welfare of the genuine voters voters:

$$U(a) = \mathbb{E}\left[\sum_{i=1}^{N} u(a, \theta_i)\right].$$

In addition to the N genuine agents, the voting population also contains T trolls (or saboteurs). Each troll agent has objective function that is diametrically opposed to that of the designer:

$$u_T(a) \equiv -\mathbb{E}\left[\sum_{i=1}^{N} u(a, \theta_i))\right] = -U(a).$$

Hence, their goal is to minimize the designer's objective function.

In order to choose a, the designer picks a voting mechanism which specifies a message set M and an outcome rule $g : M^{N+T} \to \mathbb{R}$. We will focus on *direct* mechanisms that allow voters to submit a report of their type. Formally, $M = \{\gamma_1, \gamma_2\}$. The messages of voters and trolls are indistinguishable to the designer, but she knows that there are N voters and T trolls. Given that there are two types, we can express a mechanism's outcome rule as a mapping from the number of votes for $\theta = 1$ into an outcome $a \in$

$[1,2]$:[11]

$$g : \{0,1,\ldots,N+T\} \to [1,2].$$

The timing of the model is as follows. Nature draws the types of the voters, $\{\theta_i\}_{i=1}^N$. The designer announces and commits to a mechanism g. Voters and trolls submit messages to the mechanism. The outcome is picked according to g and submitted messages, and payoffs realize.

Before we proceed to the analysis of the optimal mechanism, it is useful to analyze a benchmark case where there are no trolls. Suppose $T = 0$. The designer then faces a straightforward problem of eliciting types of the voters and picking the best action. Given that the voters vote sincerely, the designer will observe $\{\theta_i\}_{i=1}^N$. Maximizing aggregate welfare:

$$\max_{a\in[1,2]} -\sum_{i=1}^{N}(a-\theta_i)^2 \implies a^* = \frac{1}{N}\sum_{i=1}^{N}\theta_i.$$

In other words, the optimal mechanism without the trolls is the average-of-votes. Hereafter, we will describe a mechanism by its outcome rule $g(k)$ for all $k \in \{0,1,\ldots,N+T\}$, where k is the number of votes for the lower type, γ_1.

Lemma 7. *If $T = 0$, the optimal mechanism is the average-of-votes rule, i.e. the outcome rule is*

$$g(k) = \frac{k}{N}\gamma_1 + \frac{N-k}{N}\gamma_2.$$

It is useful to examine what happens to the performance of the average-of-votes mechanism (denoted by g_{av}). We will also compare its performance to that of a "blind mechanism" (g_b), which ignores the votes and always picks the ex-ante best action. Us-

[11] We allow for "compromise" outcomes, in which the designer picks action $a \in (\gamma_1,\gamma_2)$.

ing the prior, we can find that action to be

$$g_b(k) = \arg\max_a \mathbb{E}_{\theta_i}\left[\sum_{i=1}^{N} -(a-\theta_i)^2\right] \implies g_b(k) = p\gamma_1 + (1-p)\gamma_2.$$

Suppose there are some trolls, $T \geq 1$. The average-of-votes mechanism is now defined as

$$g_{av}(k) = \frac{k}{N+T}\gamma_1 + \frac{N+T-k}{N+T}\gamma_2.$$

We can show that the trolls' best strategy under this mechanism is to vote for the ex-ante *less likely* type. For concreteness, assume $p > \frac{1}{2}$, which makes γ_2 the less likely type.

Lemma 8. *Assume $p > \frac{1}{2}$. Under g_{av}, the trolls optimally vote for $\theta = \gamma_2$.*

Since the average-of-votes mechanism does not account for the trolls' presence, the designer's expected welfare is smaller than when $T = 0$. However, sometimes trolls not only reduce the effectiveness of the mechanism, but can completely overturn any welfare improvement that it generates in their absence.

Lemma 9. *The expected welfare under g_{av} is strictly lower than under the blind mechanism g_b if and only if*

$$p > \frac{N+2T}{N+2T+T^2} =: \bar{p}.$$

Note that if $T = 0$, this inequality turns into $p > 1$. This is expected: without trolls, the average-of-votes mechanism will always outperform the blind mechanism. The lemma also sheds light on the circumstances when ignoring information from the voters can be beneficial. When p is large (close to 1), there is little ex-ante uncertainty about the average type in the voter population. As a result, there is little benefit in gathering information about voters' types in the first place. This leads a mechanism that doesn't put any weight on the prior to perform worse than picking the ex-ante best action.

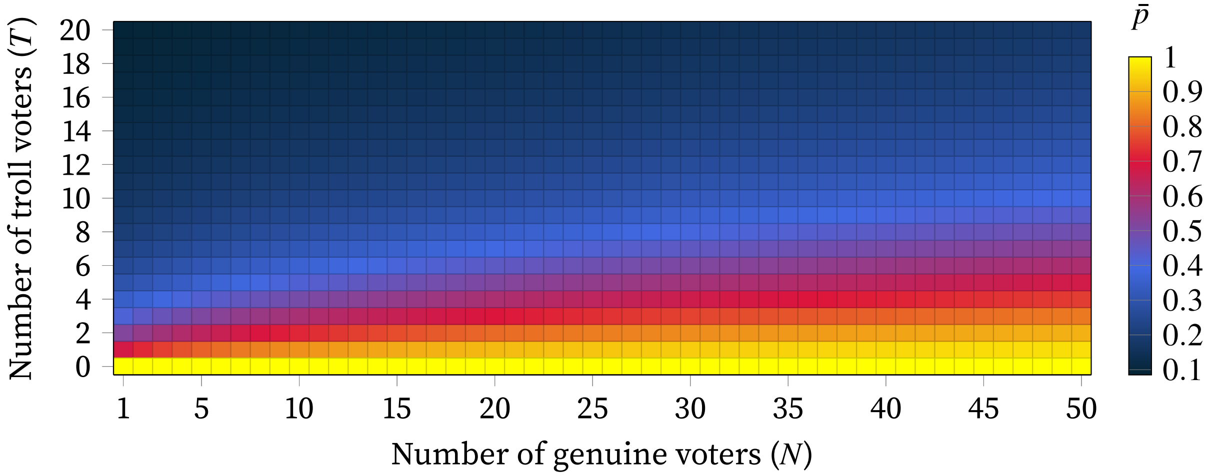


Figure 2.1: How $\bar{p} := \frac{N+2T}{N+2T+T^2}$ from Lemma 9 varies with N and T

Figure 2.1 demonstrates how $\bar{p}$ depends on N and T. If we focus on a specific level curve, we can observe that when T needs to increase at a slower rate than N to maintain the same level of $\bar{p}$. Technically, T needs to grow at the rate proportional to $\sqrt{N}$.

2.3.2 Properties of the Optimal Mechanism

This section derives two properties of the optimal mechanism in the given setting — *quasi-monotonicity* and an *indifference condition*. The first property puts reasonable bounds on the optimal mechanism and allows us to rule out the majority rule as the optimal mechanism. The second property puts a strict restriction on the optimal mechanism that generally rules out the "average-of-votes" mechanism. Additionally, we can use the insights of the indifference condition to suggest simple improvements to both of these benchmark mechanisms.

Suppose that the designer observes k votes given for type $\theta = \gamma_1$. Given that there are T trolls in total, the true number of genuine agents of type $\theta = \gamma_1$ can be from $\max\{k-T, 0\}$ to $\min\{k+T, N\}$. This allows us to put some reasonable bounds on the optimal mechanism's outcome rule. The following proposition describes them formally.

Proposition 1. *If g is an optimal mechanism, then for any $k \in \{0, 1, \ldots, N+T\}$*

$$\frac{k}{N}\gamma_1 + \frac{N-k}{N}\gamma_2 \;\leq\; g(k) \;\leq\; \frac{\max\{k-T, 0\}}{N}\gamma_1 + \frac{\min\{N-k+T, N\}}{N}\gamma_2.$$

This property can be described as quasi-monotonicity with respect to votes. Roughly speaking, the optimal mechanism's action should be increasing in the number of votes that are cast for type $\theta = \gamma_2$. There may be local non-monotonicity, but overall the outcome rule should fall within the given bounds. This property is clearly satisfied by the "average-of-votes" mechanism, but it is not satisfied by the "majority rule" mechanism. Therefore, we can conclude that the latter mechanism is not optimal under any prior p and any number of voters or trolls.

Can the "average-of-votes" mechanism be optimal, then? The next property of the optimal mechanism sheds some light on this.

Proposition 2. *Under the optimal mechanism, the trolls are indifferent between sending $m = \gamma_1$ and $m = \gamma_2$.*

We refer to this as the indifference property, since the optimal mechanism has to keep the trolls indifferent between all messages. Intuitively, when the trolls are not indifferent, the designer can slightly adjust the outcome rule under the mechanism without changing the trolls' best reply. This allows the designer to improve aggregate welfare under that strategy (and potentially worsen it under other strategies). This bit-by-bit optimization remains possible until the designer reaches a mechanism in which the trolls are indifferent between sending either message. Notably, this result can be readily generalized to a setting with more than two types; then, the optimal mechanism must keep the trolls indifferent between *all* messages.

The indifference property has two main benefits. First, it allows us to rule out a lot of potential mechanisms and focus on a narrow family of those that keep the trolls indifferent. "average-of-votes" mechanism generally does not satisfy the property. The only situation where it does is where $p = 0.5$. In that case, the trolls are indifferent between sending either messages, and it turns out that the "average-of-votes" is the optimal mechanism in that case. However, in any other situation the mechanism is not optimal because it violates the indifference property.

Second, the indifference property is a useful tool for reducing the computational complexity of searching for the optimal mechanism. In a general setting with N voters, T trolls and k types, a mechanism has to specify k^{N+T} outcomes. Proposition 2 puts $\frac{k(k-1)}{2}$ equations that restrict these variables. As a result, it reduces the dimensionality of the set of mechanisms that one needs to search through.

The impact of the indifference property can be seen visually in Figure 2.2, which depicts the designer's expected utility for the case $N = 2$ and $T = 1$ and for various priors p. Recall that in this case, a mechanism is characterized by $\{g(0), g(1), g(2), g(3)\}$, where $g(k)$ is the outcome conditional on observing k votes for γ_1. Correspondingly, the axes in each subfigure capture $g(1)$ and $g(2)$ through the weights placed on γ_1.[12]

We can observe that the designer's utility is generally higher along a dotted line in each subfigure. This dotted line depicts the set of mechanisms that satisfy the indifference property from Proposition 2. As can be seen, the designer's utility is generally higher the closer the mechanism is to the dotted line. Intuitively, the closer a mechanism is to making trolls indifferent, the better (on average) it is. We can also see the computational impact of the indifference property, which reduces the dimensionality of the set of candidate mechanisms from 2 (a square) to 1 (a line).

We are currently investigating the designer's problem under the indifference re-

[12]For instance, if the weight placed on γ_1 under 1 vote for γ_1 is equal to 0.4, that means $g(1) = 0.4\gamma_1 + 0.6\gamma_2$.

striction in order to see whether we can explicitly derive the optimal mechanism.

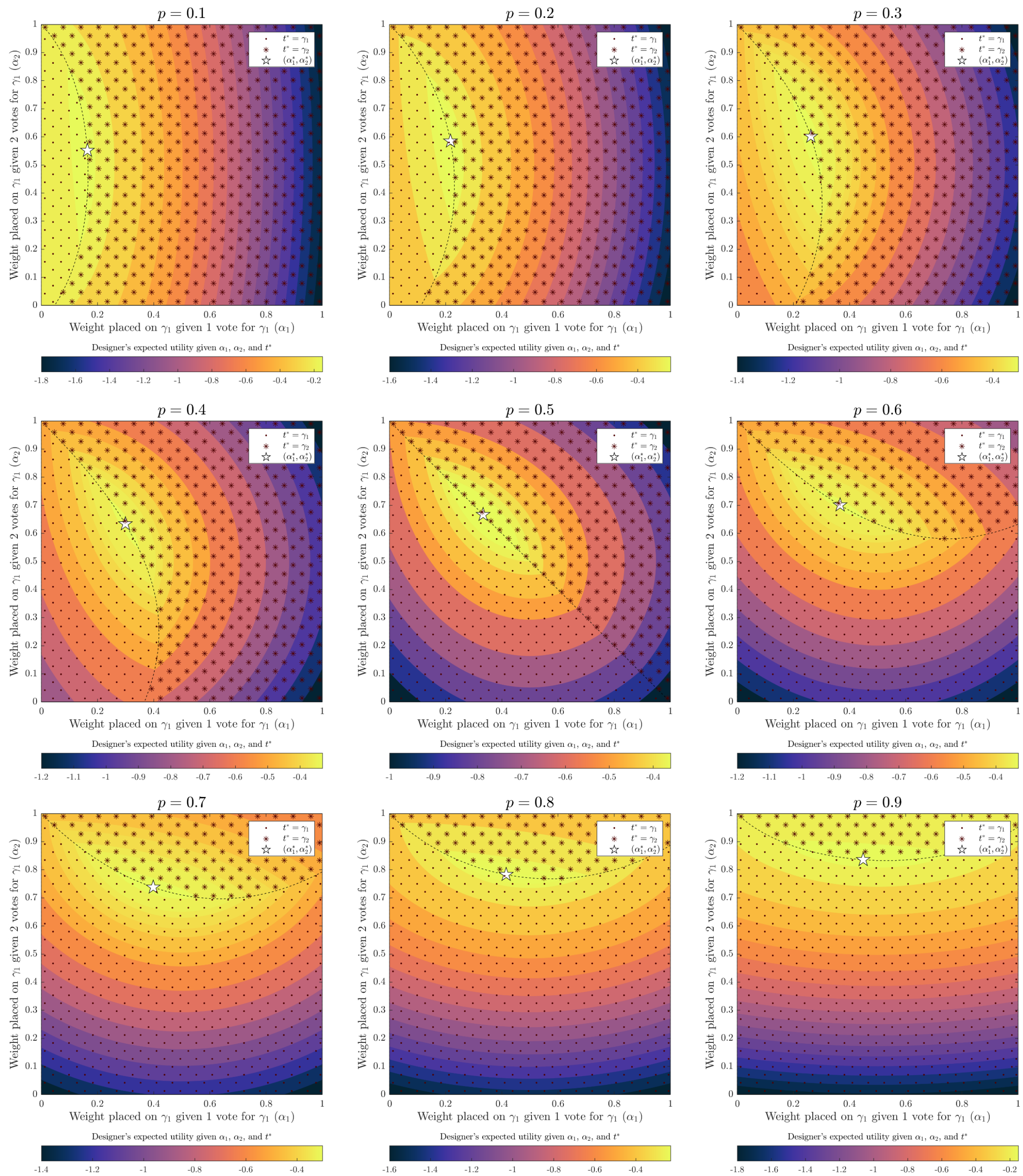


Figure 2.2: Graphical illustration of the optimal mechanism for the case $N = 2$ and $T = 1$, for various levels of p. The white star denotes the optimal mechanism. The dotted line that the star is situated on is the family of mechanisms that make the troll indifferent between messages.

2.3.3 Improving Benchmark Mechanisms

In this section, we will propose ways to improve two benchmark mechanisms — majority rule and average-of-votes rule — by using the indifference property.

Under both mechanisms, the trolls' optimal strategy is the same. They can vote for the more likely type γ_1[13] or the less likely type γ_2. To decrease the expected aggregate welfare, trolls should vote for the *less likely* type. This introduce a *bias against the prior* into the mechanisms' outcomes. In order to improve aggregate welfare, the designer should tweak the mechanisms in a way that introduces *bias towards the prior*. We will show an intuitive way to do that for both benchmark mechanisms.

First, consider the majority rule. A common modification of this mechanism is a *supermajority* rule, which makes it harder to affect the outcome of the vote by small deviations in the vote distribution. One needs to also specify what happens if the supermajority is not reached, which we refer to as the *default option*. This is where the designer can put some bias towards the prior and offset the influence of the trolls. We will show that modifying the majority rule in this way can improve its expected welfare.

Formally, let g_{mr} be the majority-rule mechanism and $g_{smr}^{\alpha,x}$ be an α-supermajority rule with default outcome x. The outcome x is implemented if neither γ_1 nor γ_2 gets enough votes to meet the threshold α. Naturally, we only consider $\alpha > \frac{1}{2}$. Before we proceed to the result, recall that $p \geq \frac{1}{2}$, i.e., type $\theta = \gamma_1$ is ex-ante more likely in the voter population. The result will assume that there are at least 3 trolls in order to avoid a trivial case where changing the supermajority rule does not change the expected outcome.

Proposition 3. *Suppose $T \geq 3$. There exists $\overline{\alpha} > \frac{1}{2}$ such that the expected welfare under mechanism $g_{smr}^{\alpha,\gamma_1}$ is strictly larger than the expected welfare under g_{mr} for any $\alpha \in \left(\frac{1}{2}, \overline{\alpha}\right)$.*

Intuitively, the supermajority rule limits the trolls' influence in two ways. First,

[13] Recall that we assume $\mathbb{P}(\theta_i = \gamma_1) = p > \frac{1}{2}$.

it makes it less likely that their vote is pivotal. Second, it incorporates a bias towards the prior in its default option. Note that in the Proposition, the supermajority rule picks $a = \gamma_1$ (the more likely type from ex-ante perspective) as its default option. This works to counteract the trolls' influence in the cases where they were previously pivotal.

Now we turn our attention to the average-of-votes mechanism. The mechanism's performance suffers in a similar fashion to the majority rule—trolls vote for the less likely type and bias the outcome against the prior. One natural way to adjust the mechanism is by changing the weights assigned to each vote. In the benchmark mechanism, votes for γ_1 and γ_2 receive equal vote in determining the outcome. The designer can introduce a bias towards the prior by assigning a larger weight in the outcome to the vote for the more likely type, γ_1.

Formally, let $g_{am}(\beta)$ be the weighted-average-of-votes rule in which votes for γ_1 receive weight β and weights for γ_2 receive weight 1. For example, if there are k votes for γ_1 and $N+T-k$ votes for γ_2, the outcome under g_{am}^{β} would be

$$g_{am}^{\beta}(k) = \frac{k\beta}{k\beta + (N+T-k)}\gamma_1 + \frac{N+T-k}{k\beta + (N+T-k)}\gamma_2.$$

Note that $\beta = 1$ corresponds to the benchmark average-of-votes rule.

Proposition 4. *There exists $\overline{\beta} > 1$ such that the expected welfare under g_{am}^{β} is strictly higher than the expected welfare under g_{am}^{1} for any $\beta \in \left(1, \overline{\beta}\right)$.*

The weighted-average-of-votes rule counters the trolls' influence in a direct way—by decreasing the weight of votes for the troll-preferred option in determining the outcome.

2.4 Limit Environment: Trolls Ruin Everything

In this section, we will consider the effect of the trolls on a mechanism's outcome when the number of trolls becomes large. For this purpose, we will restrict attention to direct mechanisms that map distribution of votes into a distribution over outcomes. Formally, consider a direct mechanism that is characterized by the outcome rule $g : \Delta\Gamma \to \Delta\Gamma$.[14] Note that the argument of the mechanism is a distribution over votes (which are types due to directness of the mechanism), whereas the output of the mechanism is a distribution over outcomes (which are types by a reasonable assumption).[15]

Let $g(\Delta\Gamma)$ denote the set of possible distributions over outcomes that mechanism g can generate. Let

$$t = \min_{x \in g(\Delta\Gamma)} V(x) = \min_{x \in g(\Delta\Gamma)} \mathbb{E}\left[\sum_{i=1}^{N} u_i(x, \theta_i)\right]$$

be the worst utility (from ex-ante perspective) that the mechanism may generate for the designer. Let π_t be the distribution of votes that produces that outcomes, i.e. $V(g(\pi_t)) = t$. This is the ideal scenario for the trolls, given that their aim is to minimize the designer's objective function. The following result focuses on trolls' ability to manipulate the mechanism to produce that scenario.

Before we introduce the result, let us introduce some notations. We will call a mechanism g *continuous* if $g : \Delta\Gamma \to \Delta\Gamma$ is a continuous mapping. Let $p(\theta, T)$ denote a distribution of votes that is produced when normal agents have types $\theta = (\theta_1, \ldots, \theta_N)$ and trolls' vote distribution is $\pi(T)$, which is defined as follows:

$$\pi(T) = \min_{\pi \in F(T)} |\pi - \pi_t|,$$

[14] Recall that $\Gamma = \{\gamma_1, \gamma_2\}$.

[15] Alternative way to view this is to map a distribution of votes into an outcome that is in the span of Γ, when such an outcome can be defined. This is the case for the "average-of-votes" mechanism.

where $F(T) = \{\pi \in \Delta\Gamma \mid \forall \gamma_i,\ \pi(\gamma_i) = \frac{k}{T} \text{ for some } k \in \mathbb{N}\}$.

Given these notations, we have the following result:

Proposition 5. *Fix a continuous mechanism $g : \Delta\Gamma \to \Delta\Gamma$. Then for any θ and any $\varepsilon > 0$, there exists $\bar{T}$ such that if $T > \bar{T}$, we have $|V(g(\pi(\theta, T))) - t| < \varepsilon$.*

This result can be interpreted as follows. Fixing any continuous mechanism, there will be a distribution of votes π_t that produces the (ex-ante) worst outcome under this mechanism. If trolls are sufficiently numerous, they can get the actual distribution of votes arbitrarily close to π_t no matter the distribution of normal agents' types. And hence, due to continuity of the mechanism, they can get its outcome $g(\pi(\theta, T))$ arbitrarily close to $g(\pi_t)$.

Letting T grow to potentially unlimited extent may seem implausible at first, since it relies on trolls being able to enter in large numbers at little to no cost.[16] However, this is a common feature of open-access online polls that have been discussed in the Introduction as part of our motivation. Several of those polls have clear signs of "botting", which is a practice of creating dozens and hundreds fake accounts or entries in order to participate in a poll. In these circumstances, having T grow arbitrarily large is not a strange assumption, and may have actually contributed to the organizers' decision to shut down those polls, as we will see below.

Note that Proposition 5 places the designer into worst-case analysis territory of mechanism design. She knows that trolls, provided that they are sufficiently numerous, may get the outcome of any mechanism arbitrarily close to the worst case of that mechanism. Given this, she may evaluate mechanisms based on their worst case alone and pick the best mechanism based on that evaluation.

One option that is always open to the designer is to simply pick a distribution over outcomes that is best from the ex-ante perspective. We will refer to this as a *blind*

[16] In fact, it's only a Google search away (https://tinyurl.com/PollBotGoogleSearch).

mechanism, since it does not gather any information from the agents and picks a distribution over outcomes based on the prior alone. Formally, a blind mechanism is characterized by function $g_b : \Delta\Gamma \to \Delta\Gamma$ such that $\forall \pi \in \Delta\Gamma,\ g_b(\pi) \in \arg\max_{x\in\Delta\Gamma} V(x)$.

Lemma 10. *Let g_b be a blind mechanism and g be any other continuous mechanism. Let π_t be the worst-case distribution of votes for g. Then $V(g(\pi_t)) \leq V(g_b(\cdot))$.*

This result easily follows from the definition of a blind mechanism. Since g_b always maps any distribution of votes into $\arg\max_{x\in\Delta\Gamma} V(x)$, and since $g(p_t)$ is the worst-case outcome for g, it must be that $V(g(p_t)) \leq \max_{x\in\Delta\Gamma} V(x)$. This implies $V(g(p_t)) \leq V(g_b(\cdot))$.

Lemma 10 indicates that when the designer expects too many trolls to participate in her mechanism, her best option might be to "shut down" the mechanism and run a blind one. This relates our analysis to motivational examples in the Introduction, where most online polls that were infiltrated by trolls were shut down by organizers. Our analysis provides theoretical rationale for such a decision, and also suggests when it is optimal. In particular, it is optimal to run a blind mechanism when costs of entering the mechanism are very low for trolls, i.e. in situations where trolls will be sufficiently numerous to bias the outcome towards the worst-case scenario.

2.5 Discussion and Next Steps

We have studied two baseline mechanisms of gathering information in the presence of trolls: "average-of-votes" and "majority rule". Moreover, the "majority rule" mechanism performs exactly the same as the no-poll benchmark, meaning that it does not give the designer any additional information that he can use. We have also shown that it is not clear which mechanism is comparatively better: if the ex-ante uncertainty over the best action is high, a more flexible option ("average-of-votes") is better, and vice versa. This suggests that more simple, rigid mechanisms may perform better in the presence of trolls when there is relatively little ex-ante uncertainty over what the best action is. And if the opposite is the case, then more flexible mechanisms may perform better. We plan to study this conjecture in a more general setting and determine its truth value.

Another important question to ask is what role commitment plays in the presence of trolls. In classical papers on limited commitment, it has been shown that a mechanism designer generally performs better when he has access to full commitment. However, in the environment where some agents actively try to sabotage his mechanisms, the power to commit may actually *hurt* the designer's objective. If he commits to a certain mechanism, then trolls will be able to take full advantage of it. They might not be able to do so if the designer does not fully commit to a mechanism. This discussion can be related to the examples from Introduction, where many online polls were rejected or shut down after it was discovered that trolls had a major influence over the outcome. If the designer picks a mechanism where he has to best respond to his beliefs about the agents' types, it imposes an interesting constraint on the trolls — they have to conceal themselves among legitimate messages in order to avoid being "detected". We plan to formalize this analysis and determine whether it is indeed true that full-commitment mechanisms perform worse (on some metric) than limited-commitment mechanisms.

It is also important to consider environments where the designer's objective is not aligned with that of the normal agents, e.g. auction or classical principal-agent setting. Apart from doing the same analysis of trolls' influence on the welfare of normal agents, we could also now disentangle two alternative models of trolls' preferences — anti-agents and anti-designer. In the first case, the trolls wish to hurt the aggregate welfare of the normal agents, which is similar to our current example. In the second case, trolls wish to hurt the designer as much as possible, which could have non-trivial effects on the welfare of normal agents.

To sum up, our future analysis includes studying full-commitment mechanisms more generally and check whether the insights we have gained so far hold there. In addition, we plan to compare full-commitment mechanisms with limited-commitment mechanisms and see whether the latter generally perform better against trolls' influence than the former. We also plan to extend analysis to the case where the designer's preferences are not aligned with preferences of normal agents.

Chapter 2 is currently being planned for submission for publication of the material. Frederick Papazyan and the dissertation author, Danil Dmitriev, are principal co-authors of this chapter.

Chapter 3

Dynamic Inconsistency and Convex Commitment Devices

3.1 Introduction

Many researchers studying dynamically inconsistent preferences have treated the demand for costly commitment devices as smoking gun evidence of present-biased dynamic inconsistency (e.g., O'Donoghue and Rabin (1999)). Commitment restricts future choice sets, which makes it easier for the decision-maker to avoid undesirable behavior. Commitment demand, therefore, indicates that the decision-maker believes she will take undesirable future actions. For the same reason, commitment devices are considered the gold standard for mitigating welfare loss from dynamic inconsistency. However, the welfare and policy value of commitment devices depends not just upon the degree of dynamic inconsistency, but also upon the degree of *perceived* dynamic inconsistency and upon the interaction between actual and perceived dynamic inconsistency. Recent studies have shown conflicting evidence regarding how sophisticated people are about their dynamic inconsistency, ranging from full naivete to partial sophistication to overly pessimistic perceptions.

In this chapter, we present a novel experimental design where we offer individuals convex commitment contracts during the performance of real-effort tasks. Similar to Augenblick and Rabin (2019), we ask participants to choose how much effort to supply for various wages. For a given wage and session, they make a labor supply decision in a previous session and at the current session. For each decision made in a previous session, we ask with what probability (between 10% and 90%) they would like to commit to the early decision rather than the late one, with higher commitment being more costly. We can then estimate structural time preference parameters from labor supply decisions and participants' implied perceptions of these parameters from commitment decisions.

We recruited 42 [1] undergraduate students at UC San Diego and ran two sets of experimental sessions, each one consisting of four decision-making periods. Despite commitment demand being costly, we observe demand from a substantial number of participants and substantial depth of demand. The labor supply decisions, in contrast, seem to be relatively time-consistent. This indicates that many of our participants were overly pessimistic about their future labor supply decisions. We quantify this by showing that both the ex-ante and ex-post welfare effects of commitment usage are negative and increasing in magnitude with commitment level, with much of this effect coming from the explicit cost of commitment.

In order to support these reduced-form results, we consider a structural model of decision-making in our experiment consisting of the standard quasi-hyperbolic discounting model (O'Donoghue and Rabin (1999)) and a convex effort cost function (as in Augenblick and Rabin (2019)). We use the structural model to obtain aggregate estimates of the present bias, exponential discount factor, disutility of labor parameters,

[1] Our goal is to recruit between 100 and 150 participants, but we have been delayed by the costs incurred to conduct the experiment.

as well as the implied perception of their present bias. We find that our estimate of the perception of present bias is significantly lower than the estimate of the present bias itself, implying that the participants on average significantly overestimate the degree of their present bias.

The results of our study have implications for future experimental design and welfare analysis. Our use of a convex commitment device appears to elicit commitment demand more precisely than binary devices. It also allows our use of a structural model to infer perceptions of dynamic inconsistency and compare them against actual inconsistency displayed by the participants. The use of a convex, rather than binary, commitment device may also help alleviate a possible experimenter demand effect. The welfare implications of our analysis are straightforward: when people are overly pessimistic about their dynamic inconsistency, they demand too much commitment and lose out on potential earnings. Future studies can extend this analysis to other fields where dynamic inconsistency and commitment demand have been observed.

The rest of the chapter is structured as follows. Section 3.2 contextualizes our contributions within the existing literature. Section 3.3 explains our experimental design. Section 3.4 provides a standard theoretical model of quasi-hyperbolic discounting which we use for structural estimation. Section 3.5 provides reduced-form results that were summarized above. Section 3.6 provides structural estimates of discounting parameters. Section 3.7 provides the welfare analysis.

3.2 Literature Review and Contributions

Empirical investigations of dynamic inconsistency have largely focused on measuring the degree of dynamic inconsistency, and have often found evidence of dynamic inconsistency and demand for costly commitment devices in many settings. However, there has been relatively little work on measuring perceptions of dynamic inconsistency. The few papers that attempt to do so yield conflicting conclusions. Augenblick et al. (2015) present a test of dynamic inconsistency in real-effort tasks and investigate the demand of a binding commitment device in an experimental setting. The authors find evidence of dynamic inconsistency in the real-effort tasks, and find that dynamic inconsistency in effort tasks predicts demand for the binding commitment device—indicating that dynamically inconsistent participants are at least somewhat aware of their inconsistency.

Augenblick and Rabin (2019) present another test of dynamic inconsistency in task performance, combining the real-effort tasks with incentivized belief elicitation to estimate perceived dynamic inconsistency. Somewhat surprisingly, the data indicate little to no aggregate awareness of dynamic inconsistency despite significant presence of dynamic inconsistency. In contrast to Augenblick and Rabin (2019), Carrera et al. (2019) find substantial evidence of partial (but not full) awareness of dynamic inconsistency. In a field experiment on gym attendance, the authors offer participants both commitment contracts and anticommitment contracts, documenting demand for both types. The results suggest caution when interpreting commitment demand as reduced-form evidence of awareness of dynamic inconsistency, with the authors warning that experimenter demand effects and noisy valuations could also be significant drivers of commitment demand rather than a desire to change future behavior.

Other recent studies involving commitment demand paint contrasting pictures of how dynamic inconsistency relates to commitment demand. Sadoff et al. (2020) present a field test of dynamic inconsistency in food choice, finding both a substantial degree of dynamic inconsistency and a substantial demand of commitment. Notably, however, the participants demanding commitment were less likely to exhibit actual dynamic inconsistency. Toussaert (2018) attempts to distinguish between commitment demand caused by awareness of dynamic inconsistency, such as in O'Donoghue and Rabin (2001), and commitment demand caused by the presence of temptation and self-control costs, such as in Gul and Pesendorfer (2003). Toussaert (2018) classifies a substantial number of experimental participants as demanding commitment because of temptation costs, while finding few subjects' behavior consistent with awareness of dynamic inconsistency.

Our project addresses the literature in several ways. First, our use of a structural model of commitment demand, rather than a reduced-form approach, allows us to interpret how much of commitment demand is coming from a desire to change future behavior and how much is coming from noise. Second, our use of a convex design also means we are robust to experimenter demand effects on commitment choice—participants may feel the experimenter wants them to select the commitment option when offered only binary choices. While we offer more commitment options than under binary choice, using multiple commitment options obscures possible inference about experimenter demands. We believe that this, in conjunction with our use of a structural model, addresses several concerns about commitment demand raised by Carrera et al. (2019).

Another contribution of this project is to address the large disconnect between the results in Augenblick and Rabin (2019) and previous literature on commitment demand. While a substantial body of previous literature finds that participants frequently demand commitment, the results in Augenblick and Rabin (2019) suggest that people have little understanding of their dynamic inconsistency, and therefore should not demand commitment. Our study attempts to connect these results through the use of the convex commitment device, which allows us to measure perceptions of dynamic inconsistency directly from commitment demand (in contrast to the belief elicitation technique used in Augenblick and Rabin (2019)). In addition, the results of our measured relationship between actual and perceived dynamic inconsistency could have an impact in model selection and welfare analysis. Both Toussaert (2018) and Sadoff et al. (2020) find a negative relationship between the degree of dynamic inconsistency and commitment demand, but differ on their assessment of the welfare effects of commitment devices by attributing commitment demand to different factors. We also find a negative relationship between the actual and perceived dynamic inconsistency parameters, but our welfare analysis is more in line with the Sadoff et al. (2020) interpretation that commitment offerings should be carefully tailored to the individuals involved to avoid welfare losses.

3.3 Experimental Design

We recruited 42 undergraduate students from the UCSD Economics Laboratory. Participants' instructions can be found in the Appendix.

Each session was scheduled on a Monday morning at 10:30, to avoid time or day effects. In addition to their earnings from task completion, participants were paid a $5 sign-up fee plus a $15 completion bonus, both upon exit to avoid income effects during the experiment. We focus on a sub-sample of 27 participants who attended all sessions and made all decisions.[2]

The experiment involves participants making labor supply decisions for piece-rate wages across time. Labor consists of transcribing strings of alphanumeric characters, with each correct string counting as one unit of labor. All participants received their payments only during the final session of the experiment, but were required to supply labor on multiple sessions during the experiment.

[2]4 participants failed to attend all sessions, and 11 attended all sessions but were unable to make all decisions due to a computer error. Including the data available from these participants does not change any structural estimates or conclusions, but it does make the non-parametric analysis unnecessarily difficult to interpret.

The experiment consisted of 4 sessions, each one week apart. In sessions 1-3, the participants made make decisions about how much labor to supply (i.e. how many strings to translate) in the following session. Each participant was be randomly shown 8 wages ranging between $0.01 and $0.31 per task performed, and asked to report how many tasks they would like to perform during the following session at each of these wages.

In sessions 1-3, after reporting their desired labor supply, the participants were also be asked to report how likely they want this decision to become the decision-that-counts in the following session. For each wage-labor decision they have just made, they were be asked to make a decision about the probability (between 10% and 90%) this decision becomes the decision-that-counts in the following session. The complementary probability is the likelihood that the decision they make in the following period session becomes the decision-that-counts in that period. The probability choices are costly, with a higher probability of committing to the current session's decision coming at a higher cost.

In sessions 2-4, the participants made decisions about how much labor to supply in the current session. They were again be shown the same 8 wages they were shown in the previous sessions, and asked how many tasks they would like to perform in the current session at each of those wages.

In sessions 2-4, after reporting their desired labor supply, we select uniformly at random one of the wages each participant faced. We then randomly select between their labor supply decision made in the current session and their decision made in the previous session, according to the probability the participant reported in the previous session. This decision is the decision-that-counts, and the participant was then asked to perform that many tasks in exchange for payment of the selected wage at the end of the experiment [3].

We would like to briefly discuss how this experimental design elicits both actual and perceived dynamic inconsistency. To measure the actual degree of present bias, we compare the labor supply decisions made on different dates about labor supplied in the same session. Decisions about labor supply in the future for wage payments in the future are not affected by present-biased dynamic inconsistency, since *both* the costs and benefits of the labor supply decision are experienced in the future. In contrast, decisions about labor supply in the present for wage payments in the future are affected by present-biased dynamic inconsistency, since labor costs are experienced immediately but wage payments are earned in the future. Since dynamic inconsistency affects only the labor supply decisions made on the date the labor work is performed, observing labor supply decisions both on the date of work and in advance allows us to measure how dynamic inconsistency affects the labor supply decisions.

To measure the *perceived* degree of dynamic inconsistency, we use the commitment decisions from sessions 1-3. The more dynamically inconsistent a participant believes they are, the further they believe their future behavior will be from where they would currently like it to be. The larger this difference in outcomes, the larger the welfare loss the participant expects from dynamically inconsistent behavior. Therefore, participants who believe they are more dynamically inconsistent are willing to pay more to commit themselves to their earlier decisions, and therefore choose a higher implementation probability. Note that binary commitment devices (only offering the options to commit or not commit, used previously in the literature), do not provide sufficiently fine choice data to reveal the degree of perceived dynamic inconsistency. Our convex commitment device allows more precise elicitation, as we show in the next section.

[3] All decisions were made using the oTree platform created by Chen et al. (2016)

3.4 Theoretical Model

Consider a decision maker with preferences over outcome streams $x = \{x_t\}_{t=1}^{T}$ given by the $(\beta, \hat{\beta}, \delta)$ preferences in O'Donoghue and Rabin (2001). In each period τ, the decision maker has preferences given by the utility function

$$U_\tau(x) = u(x_\tau) + \beta \sum_{t=\tau+1}^{T} \delta^t u(x_t).$$

The decision maker, however, believes in period τ that in period $\tau + k$ she will have preferences $\hat{U}_{\tau+k}(x) = u(x_{\tau+k}) + \hat{\beta} \sum_{t=\tau+k+1}^{T} \delta^t u(x_t)$. The parameter β is the decision maker's degree of present bias, the parameter $\hat{\beta}$ is the agent's belief about their degree of present bias, and the parameter δ is the agent's long run impatience.

The decision maker will receive monetary payments in exchange for providing labor, so $x_t = (M_t, L_t)$. [4] We assume that her Bernoulli function $u(x_t)$ is quasilinear [5], so

$$u(x_t) = M_t - C(L_t). \tag{3.1}$$

We assume that $C(0) = 0$, and that $C' > 0$ and $C'' > 0$. Following the timing above, at time τ the decision maker solves for how much labor she plans to supply during period $t > \tau$ for payment received in period T. This is given by

$$L^*_{\tau,t,\neq} = \text{argmax } \beta \cdot \delta^{T-t} \cdot L \cdot w - \beta \cdot \delta^{\tau-t} \cdot C(L). \tag{3.2}$$

During period $t = \tau$, the decision maker must decide how much labor to supply

[4] Since the decision maker recieves a linear wage payment for labor provision, note that $M_t = w \cdot L$ for some L.

[5] We must assume that the Bernoulli function is additively separable, since our experiment hinges on the costs and rewards of labor supply being in different periods and that the global utility function is additive separable across time. We use a quasilinear specification for simplicity and for structural estimation, but any additively separable specification will do.

during the current period. This is given by [6]

$$L^*_{\tau,t,=} = \text{argmax } \beta \cdot \delta^{T-t} \cdot L \cdot w - C(L). \tag{3.3}$$

At time τ the decision maker believes at time $t > \tau$ the actual amount of labor that she will supply will *not* by given by $L^*_{\tau,t,\neq}$. This is because the decision maker believes her preferences at time t are given by $\hat{U}_t$, rather than U_t. Therefore, the decision maker believes her labor supply at t will be chosen to solve

$$L^*_{\tau,t,p} = \text{argmax } \hat{\beta} \cdot \delta^{T-t} \cdot L \cdot w - \delta^{\tau-t} C(L). \tag{3.4}$$

Note that (3.4) is similar to (3.3), but rewards are weighted by $\hat{\beta}$ rather than β; this is because the agent believes that their degree of present bias is $\hat{\beta}$ rather than β. The decision maker then selects a probability $p \in [\underline{p}, \overline{p}]$ of having $L^*_{\tau,t,\neq}$ implemented at time t, for which the decision maker has to pay a cost $X(p)$ [7] at time T, where $X(0) = X'(0) = 0, X'(p) > 0$, and $X''(p) > 0$. Assuming that preferences over p are given by an expected utility function, the optimal choice of p is given by

$$\begin{aligned} p^*_{\tau,t,\neq} = \text{argmax}_p \; & p(\beta \cdot \delta^{T-t} \cdot w \cdot L^*_{\tau,t,\neq} - \beta \cdot \delta^{\tau-t} \cdot C(L^*_{\tau,t,\neq})) \\ & + (1-p)(\beta \cdot \delta^{T-t} \cdot w \cdot L^*_{\tau,t,p} - \beta \cdot \delta^{\tau-t} \cdot C(L^*_{\tau,t,p})) - \beta \cdot \delta^{T-t} \cdot X(p). \end{aligned} \tag{3.5}$$

We will now characterize $L^*_{\tau,t,\neq}, L^*_{\tau,t,=}, L^*_{\tau,t,p}$ and $p^*_{\tau,t}$. Under the assumptions on C, C' has an inverse function which we denote D. The solutions to (3.2)-(3.4) are then $L^*_{\tau,t,\neq} =$

[6]Note that in equation (3.2), both cost and payouts are weighted by β because both effort costs and monetary rewards are experienced on a future date. In equation (3.3), monetary rewards are weighted by β since those are experienced on future date while costs are not weighted by β since those are experienced immediately.

[7]The entire model can be redefined using $\tilde{p} = 1 - p$ and making $-X(\tilde{p}) = X(p)$ a payment for not committing rather than a cost for committing. This is a treatment we are interested in conducting during future sessions.

$D(\delta^{T-\tau} \cdot w)$, $L^*_{\tau,t,=} = D(\beta \cdot \delta^{T-\tau} \cdot w)$, and $L^*_{\tau,t,p} = D(\hat{\beta} \cdot \delta^{T-\tau} \cdot w)$. Taking the derivative with respect to p in equation 3.5 and substituting in the solutions to (3.2) and (3.4) produces

$$\begin{aligned}\text{derivative } &= \beta \cdot \big(\delta^{T-t} \cdot w \cdot D(w \cdot \delta^{T-\tau}) - \delta^{\tau-t} C(D(w \cdot \delta^{T-\tau})) - w \cdot \delta^{T-t} \cdot D(w \cdot \hat{\beta} \cdot \delta^{T-\tau}) \\ &\quad + \delta^{\tau-t} C(D(w \cdot \hat{\beta} \cdot \delta^{T-\tau})) - \delta^{T-t} X'(p)\big).\end{aligned}$$

If $\hat{\beta} = 1$, this is simply $-X^{'}(p)$ which implies $p^* = \underline{p}$. This is consistent with the idea that people who believe they have no present bias (time consistent and fully naive decision makers) would not choose costly commitment devices. When $\hat{\beta} < 1$, (6) is decreasing in $\hat{\beta}$ so the agent would choose $p^* > \underline{p}$. For choices of $p^* < \overline{p}$, equation (6) is set equal to 0. Thus, given w and C, $\hat{\beta}$ then uniquely determines p^*.

3.5 Reduced-form Results

The following table summarizes our observations of commitment demand. Note that the total sample size is 28 people, which is the number of people who participated in all sessions.

Category	Share of sample
Total in sample	100%
Pay for commitment at least once	93%
Pay at least $1 for commitment at least once	79%
Pay the maximal amount for commitment at least once	46%
Pay for commitment on at least 50% of decisions	68%
Pay for commitment on at least 75% of decisions	46%
Pay for commitment on every decision	11%

Commitment demand is widespread: many people commit frequently and substantially. This stands in stark contrast to a plethora of laboratory studies (Augenblick et al. (2015)) where commitment demand is often limited. We are hesitant to claim we can fully explain the difference given our moderate sample size, but we conjecture that this difference comes from three features of our design.

First, our convex device is mechanically more flexible than the binary commitment devices used in the literature. Decision-makers who value both commitment and flexibility would be generally less interested in commitment demand than those without a preference for flexibility. For example, a decision-maker that chooses to buy a moderate amount of commitment in our experiment would be inclined to not commit at all in a situation where her only choices were the extremes—commit fully or not at all.

Another factor we believe is responsible for the difference is that the design in Augenblick et al. (2015) puts the cost of commitment in different units (dollars) than the gain from commitment (units of effort), whereas our design has cost and benefits of commitment in the same unit (dollars). It is conceivable that participants find it easier to make decisions when benefits and costs are expressed in the same units. This tendency would explain why there was a large spike in commitment demand at the price of 0 in Augenblick et al. (2015) while there is almost no demand at positive prices.

Finally, our participants were allowed to make commitment decisions *conditional* upon the wage whereas in Augenblick et al. (2015) the participants made a single commitment decisions for all interest rates they faced. This implies that the participants had to average the benefits of commitment across different rates, which could result in them demanding less commitment. The benefits of commitment are tightly linked to the trade-off rates in the economic environment, so designers of commitment devices should take care to allow for flexibility when observable[8] variables change in the environment.

[8]While there are obvious benefits to designing commitment that is flexible in response to changes not observed by the designer, the costs of disentangling changing background variables from dynamic inconsistency can be high. Observability of a factor mitigates this cost, since the designer can easily verify a change in the environment.

Next, we will examine the relationship between the per-task wage and the commitment demand. As Figure 3.1 shows, there is an overall positive correlation between wage and commitment choice, and thus between labor supply and commitment choice. However, when we control for wages, labor supply does not differ significantly across periods, across time until labor is provided, or across commitment levels. We also observe that when comparing the first and the last session, commitment demand drops mildly over time for most wages. This suggests the participants may be learning about their (mis)perceptions of their present bias, and attempting to correct their behavior.

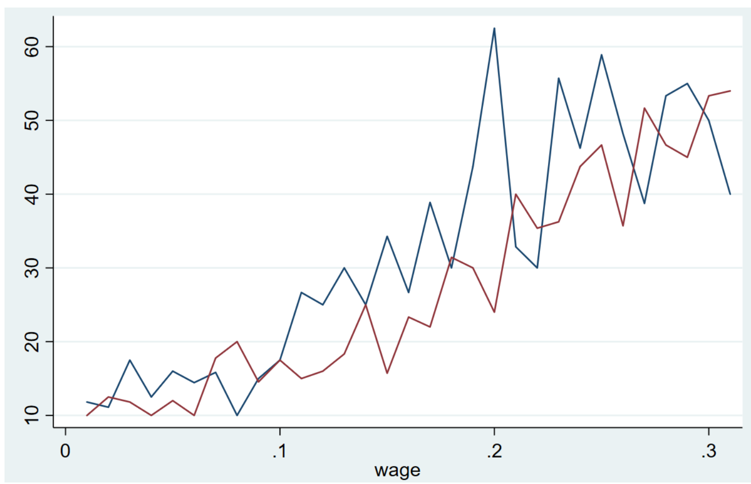


Figure 3.1: The blue line is average commitment demanded during period 1. The red line is the average commitment demanded during period 3 (the last period in which commitment is offered).

Figure 3.2 shows that labor supply decisions are increasing with the wage and are mostly consistent across time. The first observation is consistent with an increasing labor supply, as predicted by the theoretical model. The fact that the average labor supply is the same regardless of whether labor is performed in the current or in a future session suggests that participants generally do not suffer from dynamic inconsistency. We also find that participants' preferences over money and effort do not seem to change over time. This is beneficial for the structural analysis we will do in the next section.

Figure 3.2: The blue line is average labor supply chosen before the period in which labor is performed. The red line is the average labor supply chosen during the period in which labor is performed.

3.6 Structural Results

We make the structural assumption that $C(L) = \frac{1}{\varphi\gamma}L^{\gamma}$. We set $X'(p) = \frac{1}{10}(p - 0.1)$, $\overline{p} = 0.9$, and $\underline{p} = 0.1$ in our experiment. Therefore,

$$L^*_{\tau,t,\neq} = (w\delta^{T-\tau}\varphi)^{\frac{1}{\gamma-1}}, \quad L^*_{\tau,t,=} = (\beta\delta^{T-\tau}w\varphi)^{\frac{1}{\gamma-1}},$$

and

$$p^* = X'^{-1}\left(w(w\delta^{T-\tau}\varphi)^{\frac{1}{\gamma-1}} - \frac{1}{\varphi\gamma}(w\delta^{T-\tau}\varphi)^{\frac{\gamma}{\gamma-1}} - w(\hat{\beta}\delta^{T-\tau}w\varphi)^{\frac{1}{\gamma-1}} + \frac{1}{\varphi\gamma}(\hat{\beta}\delta^{T-\tau}w\varphi)^{\frac{\gamma}{\gamma-1}}\right).$$

Figure 3.1 shows the aggregate estimates from the experimental data. We see that $\beta \approx 1$ and $\delta \approx 1$, as we predicted from our graphs of the labor supply decisions. The estimate of δ is consistent with much of the previous literature, and the literature for β has some conflicting results. The estimate of $\gamma \approx 2$ is the same as Augenblick and Rabin (2019). Parameter ϕ does not appear to be well-measured, likely due to heterogeneity of

Table 3.1: Aggregate estimates of the model parameters.

			Number of obs	=	2,088
			Wald chi2(0)	=	.
Log pseudolikelihood = **-7289.0083**			Prob > chi2	=	.

(Std. Err. adjusted for **29** clusters in participant)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
sigma						
_cons	39.61456	4.812773	8.23	0.000	30.1817	49.04742
beta						
_cons	1.006873	.0485881	20.72	0.000	.911642	1.102104
delta						
_cons	1.000607	.0452341	22.12	0.000	.9119494	1.089264
phi						
_cons	2011.81	2166.194	0.93	0.353	-2233.853	6257.472
gamma						
_cons	2.401765	.2419683	9.93	0.000	1.927516	2.876014
betah						
_cons	.8015742	.0250235	32.03	0.000	.752529	.8506193

labor supply functions among participants.

The key measurement we want to focus on is $\hat{\beta}$ (which is called β_h to distinguish from the empirical estimate of β). The only empirical estimate of β_h that we are aware of is in Augenblick and Rabin (2019), who find $\beta_h \approx 1$. Our estimate of β_h is significantly lower than 1. We would like to attribute this difference in estimates to the difference in experimental designs, with Augenblick and Rabin (2019) asking participants for their predicted labor supply and our experiment using convex commitment choices. However, given our small sample, we cannot support this assertion without more data. Our estimate for β_h is also quite below the predicted lower bound of β in O'Donoghue and Rabin (2001).

The fact that β_h is estimated to be below 1 and even below β has a few implications. Our participants mostly seem to believe that they are dynamically inconsistent. However, they on average overestimate the degree of their present bias. This is also mostly true when you consider individual estimates of β and β_h, however imprecise they may be due to small number of observations. Overall, the discrepancy between β and β_h have led participants to over-demand commitment, with negative consequences for their welfare. Section 7 considers these consequences in more detail.

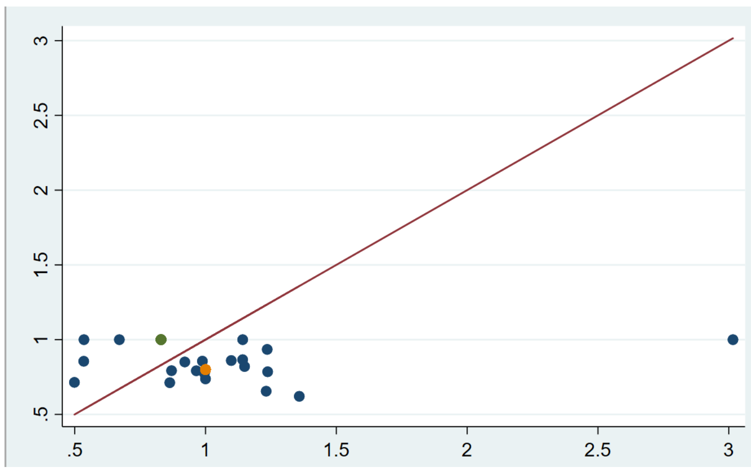


Figure 3.3: Blue dots are individual measures of β (horizontal axis) and $\hat{\beta}$ (vertical axis). The orange dot is our aggregate estimate, and green dot is the aggregate estimate from Augenblick and Rabin (2019), and the red line is a reference line for $\beta = \hat{\beta}$

Figure 3.3 shows the individual structural estimates of β and β_h for our participants. The red line shows $\beta = \beta_h$, and as we can see, many participants lie below that line [9]. Indeed, many of our participants have estimates for β_h that are below the assumed lower bound of β in O'Donoghue and Rabin (2001). This is consistent with our aggregate structural estimate that $\hat{\beta}$ is significantly lower than β.

Figure 3.4: The ex-ante (red line) and ex-post (blue line) average earnings difference as a function of commitment probability. The green line is a reference line displaying the costs of commitment

3.7 Welfare Analysis

Figure 3.4 demonstrates the welfare effects of commitment demand. The green line shows the cost of commitment. The red line considers the difference between the expected earnings under the given commitment choice and the expected earnings under the lowest, cheapest commitment choice of $p = 0.1$. The blue line simply takes the difference between the earnings from the participant's labor choice the period before labor and the earnings from the labor choice the period labor is performed, minus the cost of commitment they chose. Both red and blue lines are showing averages for the corresponding commitment levels across wages and participants.

Both welfare calculations indicate losses from commitment, and these losses are increasing in the commitment probability and closely follow the slope of the commitment cost. To put it simply, people lose from commitment—the more they commit, the more they lose. If commitment helped the participants improve their welfare, both the red and the blue lines would be above the green line. What we see instead is that both

[9] Removing the outlier above of $\beta \approx 3$ does not affect the aggregate estimates of β or $\hat{\beta}$.

lines follow the commitment cost very closely in both level and slope. This indicates that commitment does not provide any meaningful improvement to the participants' welfare. This contrasts sharply with John (2020), who documented losses at low levels of commitment and gains at high levels of commitment.

3.8 Conclusion

Commitment devices currently are the gold-standard treatment for mitigating the effects of dynamic inconsistency, but this position rests on the assumption that individuals only perceive dynamic inconsistency that they experience. Despite a moderate sample size, we document many individuals who are incorrectly pessimistic about their future selves. They demand costly commitment despite not needing it, which results in ex-post earnings losses that closely follow the cost of commitment. This result suggests that one important direction for future work is designing interventions other than commitment devices that can improve outcomes under dynamic inconsistency. Another direction is investigating other settings where dynamic inconsistency has been documented and experimentally checking whether people have correct perceptions of their inconsistency.

As discussed in Section 3.5, we find much more commitment demand than similar studies in the past (particularly Augenblick et al. (2015)). We provide several conjectures that could explain this difference. A primary factor is the fact that in our study, both the benefits and the costs of commitment are expressed in the same units (dollars). Dealing in the same units may be easier for participants and might make them more likely to commit. Another possible reason is that participants value flexibility in the commitment device, which makes our convex commitment device more attractive than a hypothetical binary device with two extreme options (fully commit or not commit at all). Finally, our participants were allowed to make commitment decisions *conditional* upon the wage whereas in Augenblick et al. (2015) the participants made a single commitment decision for all interest rates they faced. The ability to fine-tune the level of commitment to the relevant parameters in the environment could be responsible for higher eagerness to commit in our sample.

Chapter 3 is currently being prepared for submission for publication of the material. Adrian Wolanski and the dissertation author, Danil Dmitriev, are principal co-authors of this chapter.

Appendix A

Supplemental Material

A.1 Appendix of Chapter 1

A.1.1 Proof of Lemma 1

We rely on Bergemann and Välimäki (2001) in showing the applicability of Gittins index to our problem. While their model is in discrete time and ours is in continuous time, we share a key property—stationarity of the agent's experimentation problem. Since there is always another arm available, the agent never goes back to previously used arms. Optimality of an index policy follows from a similar argument as Bergemann and Välimäki (2001) use to prove their Theorem 1.

Let us start by compute the Gittins index of an arm with posterior belief π. Gittins index for this problem is the highest number $m(\pi)$ that satisfies

$$0 = \sup_{t\geq 0} \pi \int_0^t \lambda e^{-\lambda\tau} \left(e^{-r\tau} y_A - \int_0^\tau e^{-rk} m(\pi) dk \right) d\tau - (1 - \pi + \pi e^{-\lambda t}) \int_0^t e^{-r\tau} m(\pi) d\tau$$

$$0 = \sup_{t\geq 0} \pi \int_0^t \lambda e^{-\lambda\tau} \left(e^{-r\tau} y_A - \frac{1}{r}(1 - e^{-r\tau}) m(\pi) \right) d\tau - (1 - \pi + \pi e^{-\lambda t}) \frac{1}{r}(1 - e^{-rt}) m(\pi)$$

For any fixed t, this turns into

$$
\begin{aligned}
m(\pi,t) &= \frac{\pi \int_0^t \lambda e^{-(r+\lambda)\tau} y_A d\tau}{\pi \int_0^t \lambda e^{-\lambda\tau} \frac{1}{r}(1-e^{-r\tau})d\tau + (1-\pi+\pi e^{-\lambda t})\frac{1}{r}(1-e^{-rt})} \\
&= \frac{\frac{\pi\lambda}{r+\lambda}(1-e^{-(r+\lambda)t})y_A}{\frac{1}{r}\left(\pi(1-e^{-\lambda t}) - \frac{\pi\lambda}{r+\lambda}(1-e^{-(r+\lambda)t}) + (1-\pi+\pi e^{-\lambda t})(1-e^{-rt})\right)} \\
&= \frac{\frac{\pi\lambda}{r+\lambda}(1-e^{-(r+\lambda)t})y_A}{\frac{1}{r}\left(\pi - \frac{\pi\lambda}{r+\lambda} - \left(\pi - \frac{\pi\lambda}{r+\lambda}\right)e^{-(r+\lambda)t} + (1-\pi)(1-e^{-rt})\right)} \\
&= \frac{\frac{\pi\lambda}{r+\lambda}(1-e^{-(r+\lambda)t})y_A}{\frac{1}{r}\left(\frac{\pi r}{r+\lambda}(1-e^{-(r+\lambda)t}) + (1-\pi)(1-e^{-rt})\right)} \\
&= \frac{\pi r\lambda(1-e^{-(r+\lambda)t})y_A}{\pi r(1-e^{-(r+\lambda)t}) + (1-\pi)(r+\lambda)(1-e^{-rt})} \\
&= \frac{\pi r\lambda y_A}{\pi r + (1-\pi)(r+\lambda)\frac{1-e^{-rt}}{1-e^{-(r+\lambda)t}}}
\end{aligned}
$$

The index is then given by

$$
m(\pi) = \sup_{t\geq 0} \frac{\pi r\lambda y_A}{\pi r + (1-\pi)(r+\lambda)\frac{1-e^{-rt}}{1-e^{-(r+\lambda)t}}}
$$

Note that this is maximized as $t \to 0$. The index is essentially the flow payoff on a hypothetical safe arm that makes the agent indifferent between switching immediately and working on the risky arm for some time and then switching to the safe arm. The risky arm gets worse the instant one works on it, so the optimal time to switch to the safe arm is instantaneous too. Hence, the index is given by

$$
\begin{aligned}
m(\pi) &= \lim_{t \to 0} \frac{\pi r \lambda y_A}{\pi r + (1-\pi)(r+\lambda)\frac{1-e^{-rt}}{1-e^{-(r+\lambda)t}}} \\
&= \frac{\pi r \lambda y_A}{\pi r + (1-\pi)(r+\lambda)\lim_{t\to 0}\frac{re^{-rt}}{(r+\lambda)e^{-(r+\lambda)t}}} \\
&= \frac{\pi r \lambda y_A}{\pi r + (1-\pi) r} \\
&= \pi \lambda y_A.
\end{aligned}
$$

Let us also compute the Gittins index of an unexplored arm. Recall that it incurs a cost c for switching to it immediately. Thus, its index $\hat{m}$ must solve

$$
0 = \sup_{t \geq 0} -c + \pi_0 \int_0^t \lambda e^{-\lambda \tau} \left(e^{-r\tau} y_A - \int_0^\tau e^{-rk} \hat{m} dk \right) d\tau - (1 - \pi_0 + \pi_0 e^{-\lambda t}) \int_0^t e^{-r\tau} \hat{m} d\tau
$$

Following the same steps as above,

$$
\hat{m} = \sup_{t \geq 0} \frac{-c + \frac{\pi_0 \lambda}{r+\lambda}(1 - e^{-(r+\lambda)t}) y_A}{\frac{1}{r}\left(\frac{\pi_0 r}{r+\lambda}(1 - e^{-(r+\lambda)t}) + (1-\pi_0)(1 - e^{-rt}) \right)}.
$$

This is well-defined and achieves supremum for some $t > 0$. The agent will experiment on each arm until its index, $p\lambda y_A$, drifts below the index $\hat{m}$.

A.1.2 Proof of Lemma 2

During a time interval dt, the agent succeeds with probability $\pi\lambda dt$, and with probability 0 otherwise. If success does not occur, the agent updates her belief to the following:

$$
\pi + d\pi = \frac{\pi(1 - \lambda dt)}{1 - \pi \lambda dt} \approx \pi + (-\pi\lambda + \pi^2 \lambda) dt,
$$

where the latter part is an approximation with higher-order terms of dt dropped. Simplifying:

$$d\pi \approx -\lambda\pi(1-\pi)dt.$$

Whenever the agent is working on an arm with current belief π, he can choose to continue working "for a little bit" (for time $dt \to 0$) or switch to a new arm. Let V_A be the value of an unexplored arm. Then, the agent's continuation payoff satisfies the following Bellman equation:

$$u_A(\pi) = \max\{V_A,\ e^{-rdt}\mathbb{E}[u(\pi+d\pi)|\pi]\}.$$

Note that if the agent continues working, they succeed with probability $\pi\lambda dt$, which yields a payoff of $y_A + b$. Alternatively, the arm fails, and so the continuation payoff will be equal to $u(\pi+d\pi) = u_A(\pi) + \frac{\partial}{\partial\pi}u_A(\pi)d\pi$. Thus, we can express the expectation $\mathbb{E}[u(\pi+d\pi)|\pi]$ in the following way:

$$\begin{aligned}
\mathbb{E}[u(\pi+d\pi)|\pi] &= \pi\lambda(y_A+b)dt + (1-\pi\lambda dt)\left(u_A(\pi) + \frac{\partial}{\partial\pi}u_A(\pi)d\pi\right) \\
&\approx \pi\lambda(y_A+b)dt + (1-\pi\lambda dt)\left(u_A(\pi) - \lambda\pi(1-\pi)\frac{\partial}{\partial\pi}u_A(\pi)dt\right)
\end{aligned}$$

Note that we can approximate $e^{-rdt} \approx 1 - rdt$. Thus, if the agent's belief is above the switching belief $\underline{\pi}$, the continuation payoff must satisfy

$$\begin{aligned}
u_A(\pi) &\approx \pi\lambda(1-rdt)(y_A+b)dt + (1-rdt)(1-\pi\lambda dt)\left(u_A(\pi) - \lambda\pi(1-\pi)\frac{\partial}{\partial\pi}u_A(\pi)dt\right) \\
&\approx \pi\lambda(y_A+b)dt + (1-(r+\pi\lambda)dt)u_A(\pi) - \lambda\pi(1-\pi)\frac{\partial}{\partial\pi}u_A(\pi)dt.
\end{aligned}$$

The last step removes higher-order terms of dt, as is standard. Rearranging the equation and substituting back into the Bellman equation above gives us

$$u_A(\pi) = \max\left\{V_A,\ \frac{\pi\lambda}{r+\pi\lambda}(y_A+b) - \frac{\lambda\pi(1-\pi)}{r+\pi\lambda}\frac{\partial}{\partial\pi}u_A(\pi)\right\}.$$

This also gives us a differential equation on $u_A(\pi)$ for beliefs above the switching belief $\underline{\pi}$:

$$\frac{\partial}{\partial\pi}u_A(\pi) + \frac{r+\pi\lambda}{\lambda\pi(1-\pi)}u_A(\pi) = \frac{1}{1-\pi}(y_A+b) \tag{A.1}$$

This first-order linear differential equation can be solved using integrating factor $G(\pi)$:

$$G(\pi) = e^{\int \frac{r+\pi\lambda}{\lambda\pi(1-\pi)}d\pi} = e^{\frac{r}{\lambda}(\ln(\pi)-\ln(1-\pi))-\ln(1-\pi)} = \left(\frac{\pi}{1-\pi}\right)^{\frac{r}{\lambda}}\frac{1}{1-\pi}$$

Equation (A.1) then simplifies to

$$\frac{\partial}{\partial\pi}(G(\pi)u_A(\pi)) = \frac{1}{(1-\pi)}(y_A-c+V_A)G(\pi).$$

Thus,

$$\begin{aligned}
u_A(\pi) &= \frac{1}{G(\pi)}\left(\int \frac{y_A+b}{1-\pi}G(\pi)d\pi + C\right) \\
&= \frac{1}{G(\pi)}(y_A+b)\int\left(\frac{\pi}{1-\pi}\right)^{\frac{r}{\lambda}}\frac{1}{(1-\pi)^2}d\pi + \frac{C}{G(\pi)} \\
&= \frac{1}{G(\pi)}(y_A+b)\frac{\lambda}{r+\lambda}\left(\frac{\pi}{1-\pi}\right)^{\frac{r+\lambda}{\lambda}} + \frac{C}{G(\pi)} \\
&= \frac{\pi\lambda}{r+\lambda}(y_A+b) + \left(\frac{1-\pi}{\pi}\right)^{\frac{r}{\lambda}}(1-\pi)C,
\end{aligned}$$

where C is an integration constant that will be pinned down by the boundary conditions. Specifically, the continuation payoff function above needs to satisfy value matching and smooth pasting. Let $\underline{\pi}$ be the belief at which the agent chooses to switch to a new arm.

Then:

$$u_A(\underline{\pi}) = V_A, \text{ and } \frac{\partial}{\partial \pi} u_A(\underline{\pi}) = 0.$$

The smooth pasting condition is equivalent to:

$$\begin{aligned} \frac{\lambda}{r+\lambda}(y_A + b) - \left(\frac{1-\underline{\pi}}{\underline{\pi}}\right)^{\frac{r}{\lambda}} C - \frac{r}{\lambda}\frac{1}{\underline{\pi}^2}\left(\frac{1-\underline{\pi}}{\underline{\pi}}\right)^{\frac{r-\lambda}{\lambda}}(1-\underline{\pi})C = 0 \\ \frac{\lambda}{r+\lambda}(y_A + b) - \frac{\underline{\pi}\lambda + r}{\underline{\pi}\lambda}\left(\frac{1-\underline{\pi}}{\underline{\pi}}\right)^{\frac{r}{\lambda}} C = 0 \end{aligned}$$

Hence,

$$C = \frac{\lambda}{r+\lambda} \cdot \frac{\underline{\pi}\lambda}{\underline{\pi}\lambda + r}\left(\frac{\underline{\pi}}{1-\underline{\pi}}\right)^{\frac{r}{\lambda}}(y_A + b).$$

Substituting this into the value matching condition:

$$\begin{aligned} \left(\frac{\underline{\pi}\lambda}{r+\lambda} + \frac{\lambda}{r+\lambda} \cdot \frac{\lambda\underline{\pi}(1-\underline{\pi})}{\underline{\pi}\lambda + r}\right)(y_A + b) = V_A \\ \frac{\underline{\pi}\lambda(\underline{\pi}\lambda + r) + \underline{\pi}\lambda^2 - \underline{\pi}^2\lambda^2}{(r+\lambda)(\underline{\pi}\lambda + r)}(y_A + b) = V_A \\ \frac{\underline{\pi}\lambda}{\underline{\pi}\lambda + r}(y_A + b) = V_A \end{aligned}$$

Rearranging:

$$(\underline{\pi}\lambda + r)V_A = \underline{\pi}\lambda(y_A + b) \implies \underline{\pi} = \frac{rV_A}{\lambda(y_A + b - V_A)}$$

The agent will switch to a new arm if and only if $\pi < \underline{\pi}$.

We can deduce the agent's waiting time from this switching belief. Let $t^A(b, V_A)$ be the agent's optimal switching time given bonus payment b and new arm's value V_A. Note:

$$\underline{\pi} = \frac{\pi_0 e^{-\lambda t^*(y_A)}}{1 - \pi_0 + \pi_0 e^{-\lambda t^*(y_A)}}.$$

Hence,

$$t^A(b,V_A) = -\frac{1}{\lambda}\ln\left(\frac{(1-\pi_0)\underline{\pi}}{\pi_0(1-\underline{\pi})}\right)$$
$$= -\frac{1}{\lambda}\ln\left(\frac{(1-\pi_0)rV_A}{\pi_0(\lambda(y_A+b-V_A)-rV_A)}\right).$$

A.1.3 Proof of Lemma 3

Consider an arbitrary opaque bonus scheme with a reward distribution F. Let $\{b_1,b_2,\ldots,b_n\}$ be the support of F, and let f_i be the probability of $b=b_i$. Let $V_A(F)=\bar{V}_A$ be the agent's initial value, i.e. the value of an unexplored arm, given the agent-optimal experimentation policy under F (described in Lemma 2). Let $V_A(b,x)$ be the agent's continuation value immediately after drawing reward b when the value of an unexplored arm is x. Similarly, denote the principal's continuation value by $V_P(b,x)$.[1]

Suppose the principal wants to construct a better bonus scheme $\hat{F}$ such that the supports of F and $\hat{F}$ coincide and $V_A(F')=V_A(F)=\bar{V}_A$. The principal solves the following problem:

$$\begin{aligned} \max_{\hat{f}_1,\hat{f}_2,\ldots,\hat{f}_n\geq 0} \quad & \sum_{i=1}^{n}\hat{f}_i V_P(b_i,\bar{V}_A), \\ \text{s.t.} \quad & \sum_{i=1}^{n}\hat{f}_i V_A(b_i,\bar{V}_A)=\bar{V}_A, \\ & \sum_{i=1}^{n}\hat{f}_i = 1. \end{aligned}$$

Expressing $\hat{f}_1$ through the first constraint and substituting into the objective function and the second constraint morphs this problem into maximizing a linear objective function subject to one linear constraint, with $n-1$ variables. As is standard, there exists a "corner solution", where $n-2$ variables equal zero. Therefore, there exists a solution

[1]The principal's continuation value depends on x because the agent's experimentation policy depends on both b and x.

to the original problem that has only two positive probabilities, which corresponds to a binary bonus scheme. Denote this scheme by F'. Given that F is a feasible point in the optimization problem above, it follows that $V_A(F') = V_A(F)$ and $V_P(F') \geq V_P(F)$.

A.1.4 Proof of Theorem 1

Consider a deterministic bonus scheme with bonus payment $b = \bar{b}$, and consider a binary opaque bonus scheme F. Let $b_L < \bar{b}$ and $b_H > \bar{b}$ be the two offered payments, and let q be the probability of the low payment. Suppose that the principal keeps the agent's value the same as under the deterministic bonus scheme, $V_A(F) = V_A(\bar{b})$.

The agent's switching time is given by:

$$t_L^* = -\frac{1}{\lambda}\ln\left(\frac{(1-\pi_0)r\bar{u}}{\pi_0(\lambda(y_A+b_L-\bar{u})-\bar{u})}\right)$$
$$t_H^* = -\frac{1}{\lambda}\ln\left(\frac{(1-\pi_0)r\bar{u}}{\pi_0(\lambda(y_A+b_H-\bar{u})-\bar{u})}\right)$$

The agent's continuation value under each payment is given by:

$$V_A(b_L) = \frac{\pi_0\lambda}{r+\lambda}(y_A+b_L)\left(1-e^{-(r+\lambda)t_L^*}\right)+\left(1-\pi_0+\pi_0 e^{-\lambda t_L^*}\right)e^{-rt_L^*}\bar{u}$$
$$V_A(b_H) = \frac{\pi_0\lambda}{r+\lambda}(y_A+b_H)\left(1-e^{-(r+\lambda)t_H^*}\right)+\left(1-\pi_0+\pi_0 e^{-\lambda t_H^*}\right)e^{-rt_H^*}\bar{u}$$

Since the principal wants to keep the agent's overall expected value at $\bar{u}$, the bonus scheme's probability q is pinned down by the following:

$$\bar{u} = -c + qV_A(b_L) + (1-q)V_A(b_H) \implies q = \frac{V_A(b_H)-c-\bar{u}}{V_A(b_H)-V_A(b_L)}.$$

The principal's overall value (as a function of b_H) is then given by

$$V_P = q \cdot \left(\frac{\pi_0 \lambda}{r+\lambda}(y_P - b_L)\left(1 - e^{-(r+\lambda)t_L^*}\right) + \left(1 - \pi_0 + \pi_0 e^{-\lambda t_L^*}\right) e^{-rt_L^*} V_P \right) +$$
$$+ (1-q) \cdot \left(\frac{\pi_0 \lambda}{r+\lambda}(y_P - b_H)\left(1 - e^{-(r+\lambda)t_H^*}\right) + \left(1 - \pi_0 + \pi_0 e^{-\lambda t_H^*}\right) e^{-rt_H^*} V_P \right)$$

We can solve for V_P in closed form:

$$V_P = \frac{\pi_0 \lambda \left(q(y_P - b_L)\left(1 - e^{-(r+\lambda)t_L^*}\right) + (1-q)(y_P - b_H)\left(1 - e^{-(r+\lambda)t_H^*}\right)\right)}{(r+\lambda)\left(1 - q(1 - \pi_0 + \pi_0 e^{-\lambda t_L^*})e^{-rt_L^*} - (1-q)(1-\pi_0 + \pi_0 e^{-\lambda t_H^*})e^{-rt_H^*}\right)}$$

Suppose that the principal increases the spread of payments by letting $b_H \to \infty$. As $b_H \to \infty$, we have $q \to 1$ and $t_H^* \to \infty$, which simplifies the expression, in particular the denominator:

$$V_P \to \frac{\pi_0 \lambda (y_P - b_L)\left(1 - e^{-(r+\lambda)t_L^*}\right) - \pi_0 \lambda \lim_{b_H \to \infty}(1-q) b_H}{(r+\lambda)\left(1 - (1 - \pi_0 + \pi_0 e^{-\lambda t_L^*})e^{-rt_L^*}\right)}$$

Note that

$$\begin{aligned}
\lim_{b_H \to \infty}(1-q)b_H &= \lim_{b_H \to \infty}(\bar{u} + c - V_A(b_L)) \frac{b_H}{V_A(b_H) - V_A(b_L)} \\
&= (\bar{u} + c - V_A(b_L)) \lim_{b_H \to \infty} \frac{b_H}{\frac{\pi_0 \lambda}{r+\lambda} b_H \left(1 - e^{-(r+\lambda)t_H^*}\right) + o(b_H)} \\
&= (\bar{u} + c - V_A(b_L)) \cdot \frac{r+\lambda}{\pi_0 \lambda}
\end{aligned}$$

Hence,

$$
\begin{aligned}
V_P \to & \frac{\frac{\pi_0\lambda}{r+\lambda}(y_P - b_L)\left(1 - e^{-(r+\lambda)t_L^*}\right) - (\bar{u} + c - V_A(b_L))}{1 - (1 - \pi_0 + \pi_0 e^{-\lambda t_L^*})e^{-rt_L^*}} \\
= & \frac{\frac{\pi_0\lambda}{r+\lambda}(y_P + y_A + \bar{u})\left(1 - e^{-(r+\lambda)t_L^*}\right) + \left(1 - \pi_0 + \pi_0 e^{-\lambda t_L^*}\right)e^{-rt_L^*}\bar{u} - \bar{u} - c}{1 - (1 - \pi_0 + \pi_0 e^{-\lambda t_L^*})e^{-rt_L^*}} \\
= & \frac{-c + \frac{\pi_0\lambda}{r+\lambda}(y_P + y_A)\left(1 - e^{-(r+\lambda)t_L^*}\right)}{1 - (1 - \pi_0 + \pi_0 e^{-\lambda t_L^*})e^{-rt_L^*}} - \bar{u}
\end{aligned}
$$

Note that this limit is the joint value under the switching time t_L^*, minus the agent's reservation utility. Let t^* be the switching time that maximizes the joint value (recall Section 3.3). We know that $t^A(\bar{b}, \bar{u}) > t^*$, since the transparent bonus scheme is certainly inefficient. Thus, given that $b_L < \bar{b}$ and that $t^A(b, \bar{u})$ is increasing in w, we can conclude that the principal's value is higher under the limiting opaque bonus scheme than under the corresponding transparent bonus scheme for any $b_L < \bar{b}$.

A.1.5 Proof of Theorem 2

Consider a binary opaque bonus scheme with low reward b_L and high reward b_H, with probability of low reward being q. Let $\bar{V}_A = q \cdot V_A(b_L) + (1-q) \cdot V_A(b_H)$ be the agent's expected value under it. Note that this pins down q.

Note that the agent's continuation value of a given arm with reward s, $u_A(\pi_0, s, \bar{V}_A)$, is increasing in s. This implies that q is increasing in b_H. Moreover, $\lim_{s\to\infty} u_A(\pi_0, s, \bar{V}_A) = \infty$, as this continuation value is given by

$$
u_A(\pi_0, s, \bar{V}_A) = \max_{\tau \geq 0} \frac{\pi_0\lambda}{r+\lambda}(y_A + s + \bar{V}_A)\left(1 - e^{-(r+\lambda)\tau}\right) + \left(1 - \pi_0 + \pi_0 e^{-\lambda\tau}\right)e^{-r\tau}\bar{V}_A,
$$

where the objective function goes to ∞ for any fixed value of τ. Additionally, note that the joint continuation value of a given arm with reward s, $u_{joint}(\pi_0, s, \bar{V}_A)$ is bounded.

Moreover,

$$\lim_{s\to\infty} u_{joint}(\pi_0, s, \bar{V}_A) = \frac{\frac{\pi_0\lambda}{\lambda+r}(y_A + y_P - c)}{1 - \frac{\pi_0\lambda}{\lambda+r}},$$

since the agent's switching time goes to ∞ as $s \to \infty$.

Consider the opaque bonus scheme with b_L and b_H once more. Using the implications above, we can conclude that

$$q \cdot u_{joint}(\pi_0, b_L, \bar{V}_A) + (1-q) \cdot u_{joint}(\pi_0, b_H, \bar{V}_A) > q \cdot u_{joint}(\pi_0, b_L, \bar{V}_A) + (1-q) \cdot \frac{\frac{\pi_0\lambda}{\lambda+r}(y_A + y_P - c)}{1 - \frac{\pi_0\lambda}{\lambda+r}},$$

where the right-hand function is strictly increasing in b_H due to q being increasing in b_H. Thus, there exists $\bar{b}$ such that for any $b_H > \bar{b}$, the left-hand function is also increasing in b_H.

As $b_H \to \infty$, the joint expected value converges to $u_{joint}(\pi_0, b_L, \bar{V}_A)$, similar to the proof of Theorem 1. Since $t^A(b_H, \bar{V}_A) > t^A(b_L, \bar{V}_A) > t^*$, it follows that $u_{joint}(\pi_0, b_L, \bar{V}_A) > u_{joint}(\pi_0, b_H, \bar{V}_A)$ for any $0 \leq b_L < b_H$. Therefore,

$$u_{joint}(\pi_0, b_L, \bar{V}_A) > q \cdot u_{joint}(\pi_0, b_L, \bar{V}_A) + (1-q) \cdot u_{joint}(\pi_0, b_H, \bar{V}_A)$$

for any $q < 1$. This implies that the joint expected value attains its global maximum when $b_H \to \infty$.

A.1.6 Proof of Theorem 3

Recall from the proofs of Theorem 1 and Theorem 2 that the optimal binary bonus scheme F^* lets $b_H \to \infty$ while keeping the agent's participation constraint satisfied, $V_A(F) = \bar{u}$. In particular, the principal's expected value of this bonus scheme is given by

$$V_P(F^*) = u_{joint}(\pi_0, b_L, \bar{u}) - c - \bar{u}.$$

The optimal binary bonus scheme should choose b_L to maximize the joint continuation value $u_{joint}(\pi_0, b_L, \bar{u})$. Since the transfer does not enter the value function except through the agent's switching time, this is equivalent to choosing b_L so that

$$t^A(b_L, \bar{u}) = t^*,$$

or as close to t^* as possible. Recall that $t^A(b_L, \bar{u})$ is increasing in b_L. This implies that the principal can match the efficient switching time t^* as long as it is possible under the lowest possible payment, $b_L = 0$.

Recall from Lemma 2 that the agent's switching rule satisfies

$$\frac{r\bar{u}}{\lambda(y_A + b_L - \bar{u})} = \frac{\pi_0 e^{-\lambda t^A(b_L, \bar{u})}}{1 - \pi_0 + \pi_0 e^{-\lambda t^A(b_L, \bar{u})}},$$

which is equivalent to

$$\bar{u} = \frac{\lambda(y_A + b_L)}{\lambda + r \cdot \frac{1 - \pi_0 + \pi_0 e^{-\lambda t^A(b_L, \bar{u})}}{\pi_0 e^{-\lambda t^A(b_L, \bar{u})}}}.$$

Thus, the principal can choose b_L to achieve $t^A(b_L, \bar{u}) = t^*$ if and only if

$$\bar{u} \geq \frac{\lambda y_A}{\lambda + r \cdot \frac{1 - \pi_0 + \pi_0 e^{-\lambda t^*}}{\pi_0 e^{-\lambda t^*}}}.$$

A.1.7 Proof of Theorem 4

Adopt most of Theorem 3's proof. Let p^* be the efficient switching belief. Under the limiting opaque bonus scheme from Theorems 1 and 3, the agent's switching belief is given by

$$\pi^A(b_L) = \frac{r\bar{u}}{\lambda(y_A + b_L - \bar{u})},$$

where b_L is the low reward in the bonus scheme under consideration. Note that $\pi^A(b_L)$ is monotonically decreasing in b_L, and $\lim_{b_L \to -y_A} \pi_A(b_L) = +\infty$. Thus, there exists $b_L > -y_A$ such that $\pi^A(b_L) = \pi^*$. This implies that the optimal bonus scheme achieves the efficient switching policy with probability 1 (probability of b_L) and extracts all surplus by holding the agent's participation constraint with equality.

A.1.8 Proof of Theorem 5

First, suppose that $\hat{w}_b = 0$. This means that there exists a switching subsidy $\hat{w}_s \leq c_B$ such that the agent is indifferent between the outside option and the experimentation. By Lemma 2, the agent's switching belief in this case is $\pi^A = \frac{r\bar{u}}{\lambda y_A}$. This is the highest switching belief that the principal can possibly achieve using any stationary bonus scheme (transparent or opaque) while maintaining the agent's value at $\bar{u}$. Hence, the transparent bonus scheme $(\hat{w}_s, 0)$ is optimal. If $\hat{w}_b > 0$ but $\pi^* \leq \frac{r\bar{u}}{\lambda(y_A+\hat{w}_b)}$, then there exists a transparent bonus scheme (w_s, w_b) such that $w_s \leq \hat{w}_s \leq c_B$, $w_b > \hat{w}_b$, and the agent's switching belief is $\pi^A = \pi^*$. Hence, scheme (w_s, w_b) is optimal.

Now suppose that we have $\hat{w}_b > 0$ and $\pi^* > \frac{r\bar{u}}{\lambda(y_A+\hat{w}_b)}$. The transparent bonus scheme $(\hat{w}_s, \hat{w}_b)$ induces the switching belief $\pi^A = \frac{r\bar{u}}{\lambda(y_A+\hat{w}_b)}$. Increasing the agent's switching belief to get it closer to π^* would require reducing $\hat{w}_b$. To maintain the agent's value at $\bar{u}$, the principal would have to simultaneously increase $\hat{w}_s$. However, $\hat{w}_b > 0$ implies that $\hat{w}_s = c_B$; the principal thus cannot increase $\hat{w}_s$ any further without inducing the agent to constantly switch to cheap arms. In this case, no transparent bonus scheme achieves the principal's first-best outcome.

Consider designing a binary opaque bonus scheme with zero switching subsidy, as in Theorems 1 and 2. In the limit as the high reward becomes arbitrarily large, the opaque bonus scheme achieves the switching belief $\pi^A = \frac{r\bar{u}}{\lambda(y_A+b_L)}$ with probability 1. By choosing $b_L < \hat{w}_b$, we will get an opaque bonus scheme that has a more efficient switching belief while maintaining the agent's value at $\bar{u}$. The optimal bonus scheme is thus opaque, and not any of the transparent bonus schemes with a switching subsidy.

A.2 Appendix of Chapter 2

A.2.1 Proof of Lemma 5

Genuine Players' ($i = 1,2$) Problems

Given the vector of messages $\mathbf{m} \in \{L,H\}^3$ and types $(\theta_1, \theta_2) \in \{L,H\}^2$, (genuine) player $i \in \{1,2\}$ has the following utility function:

$$u_i(x(\mathbf{m}), \theta_i) = -(x(\mathbf{m}) - \theta_i)^2, \tag{A.2}$$

where $x(\mathbf{m}) \equiv \mathcal{U}\{\text{mode}(\mathbf{m})\}$. Notice that under this simple setup, $x(\cdot)$ simplifies to

$$x(\mathbf{m}) \equiv \text{sgn}(m_1 + m_2 + m_3) \tag{A.3}$$

because of the facts that there are 3 voters and that $(L,H) = (-1,1)$.

Let m_{k,θ_k} denote the message of player $k \in \{1,2,3\}$ who is of type $\theta_k \in \{L,H\} \cup \{T\}$. Then the expected utilities of players 1 and 2 are given by

$$\mathbb{E}_{\theta_2} u_1(m_{1\theta_1}, (m_{2\theta_2}, m_{3T}), \theta_1) = -p[x(m_{1\theta_1}, m_{2L}, m_{3T}) - \theta_1]^2 - (1-p)[x(m_{1\theta_1}, m_{2H}, m_{3T}) - \theta_1]^2$$

and

$$\mathbb{E}_{\theta_1} u_2(m_{2\theta_2}, (m_{1\theta_1}, m_{3T}), \theta_2) = -p[x(m_{1L}, m_{2\theta_2}, m_{3T}) - \theta_2]^2 - (1-p)[x(m_{1H}, m_{2\theta_2}, m_{3T}) - \theta_2]^2,$$

respectively.

To facilitate performing the calculations below, let $\mathrm{EU}_{1\theta_1}(m_{1\theta_1},(m_{2L},m_{2H};m_{3T})) := \mathbb{E}_{\theta_2} u_1(m_{1\theta_1},(m_{2\theta_2},m_{3T}),\theta_1)$ and define $\mathrm{EU}_{1\theta_1}(m_{1\theta_1},(m_{2L},m_{2H};m_{3T}))$ similarly.

Let $i = 1$ and $\theta_1 = L = -1$. To characterize player 1's optimal strategy, I consider the following 8 cases, summarized in the table, below:

	1	2	3	4	5	6	7	8
m_{2L}	L	L	H	H	L	L	H	H
m_{2H}	H	H	L	L	L	L	H	H
m_{3T}	L	H	L	H	L	H	L	H

Going through all of these cases exhausts all the possible (m_{2L},m_{2H},m_{3T}).

Case 1: $(m_{2L},m_{2H};m_{3T}) = (L,H;L)$

$EU_{1L}(-1,(-1,1;-1)) = 0 > -4(1-p) = EU_{1L}(1,(-1,1;-1)) \Rightarrow m^*_{1L}(-1,1;-1) = -1.$

Case 2: $(m_{2L},m_{2H};m_{3T}) = (L,H;H)$

$EU_{1L}(-1,(-1,1;1)) = -4(1-p) > -4 = EU_{1L}(1,(-1,1;1)) \Rightarrow m^*_{1L}(-1,1;1) = -1.$

Case 3: $(m_{2L},m_{2H};m_{3T}) = (H,L;L)$

$EU_{1L}(-1,(1,-1;-1)) = 0 > -4 = EU_{1L}(1,(1,-1;-1)) \Rightarrow m^*_{1L}(1,-1;-1) = -1.$

Case 4: $(m_{2L},m_{2H};m_{3T}) = (H,L;H)$

$EU_{1L}(-1,(1,-1;1)) = -4p > -4 = EU_{1L}(1,(1,-1;1)) \Rightarrow m^*_{1L}(1,-1;1) = -1.$

Case 5: $(m_{2L},m_{2H};m_{3T}) = (L,L;L)$

$EU_{1L}(-1,(-1,-1;-1)) = 0 = EU_{1L}(1,(-1,-1;-1)) \Rightarrow m^*_{1L}(-1,-1;-1) = \{-1,1\}.$

Case 6: $(m_{2L},m_{2H};m_{3T}) = (L,L,H)$

$EU_{1L}(-1,(-1,-1;1)) = 0 > -4 = EU_{1L}(1,(-1,-1;1)) \Rightarrow m^*_{1L}(-1,-1;1) = -1.$

Case 7: $(m_{2L},m_{2H};m_{3T}) = (H,H;L)$

$EU_{1L}(-1,(1,1;-1)) = 0 > -4 = EU_{1L}(1,(1,1;-1)) \Rightarrow m^*_{1L}(1,1;-1) = -1.$

Case 8: $(m_{2L},m_{2H};m_{3T}) = (H,H;H)$

$EU_{1L}(-1,(1,1;1)) = 0 = EU_{1L}(1,(1,1;1)) \Rightarrow m^*_{1L}(1,1;1) = \{-1,1\}.$

Summarizing the above,

$$m^*_{1L}(m_{2L}, m_{2H}; m_{3T}) = \begin{cases} \{-1,1\} & \text{, if } m_{2L} = m_{2H} = m_{3T} \\ -1 & \text{, otherwise} \end{cases}$$

The derivations for $m^*_{2L}(\cdot)$, $m^*_{1H}(\cdot)$, and $m^*_{2H}(\cdot)$ are very similar and are omitted. Let $i, j \in \{1,2\}$ ($i \neq j$) and $\theta_i \in \{L, H\}$. Then,

$$m^*_{i\theta_i}(m_{jL}, m_{jH}, m_{3T}) = \begin{cases} \{-1,1\} & \text{, if } m_{2L} = m_{2H} = m_{3T} \\ \theta_i & \text{, otherwise.} \end{cases}$$

For simplicity, we impose that when $\text{EU}_{i\theta_i}(m_{i\theta_i}, (m_{jL}, m_{jH}; m_{3T})) = \text{EU}_{i\theta_i}(-m_{i\theta_i}, (m_{jL}, m_{jH}; m_{3T}))$, player i (of type θ_i) will vote θ_i. Imposing this "indifference-breaking condition" yields

$$m^*_{i\theta_i}(m_{jL}, m_{jH}, m_{3T}) \equiv \theta_i \quad \text{(sincere voting by genuine types)}$$

Saboteur's Problem (Player 3)

Given θ_{-3} and $\mathbf{m}_{-3,\theta_{-3}}$,

$$u_3(m_{3T}, \mathbf{m}_{-3,\theta_{-3}}, \theta_{-3}) = -\sum_{j \neq 3} u_j(m_{j\theta_j}, \mathbf{m}_{-j,\theta_{-j}}, \theta_j)$$

Given the (common) prior, player 3's expected utility is as follows:

$$
\begin{aligned}
\mathbb{E}_{\theta_{-3}}[u_3(m_{3T}, \mathbf{m}_{-3,\theta_{-3}}, \theta_{-3})] \quad &= \quad \sum_{\tau \in \{L,H\}^2} \mathbb{P}\{\theta_{-3} = \tau\} \cdot u_3(m_{3T}, \mathbf{m}_{-3,\tau}, \tau) \\
&\equiv \quad p^2\big[-u_1(m_{1L}, (m_{2L}, m_{3T}), L) - u_2(m_{2L}, (m_{1L}, m_{3T}), L)\big] \\
&\quad + \; p(1-p)\big[-u_1(m_{1L}, (m_{2H}, m_{3T}), L) - u_2(m_{2H}, (m_{1L}, m_{3T}), H)\big] \\
&\quad + \; (1-p)(p)\big[-u_1(m_{1H}, (m_{2L}, m_{3T}), H) - u_2(m_{2L}, (m_{1H}, m_{3T}), L)\big] \\
&\quad + \; (1-p)^2\big[-u_1(m_{1H}, (m_{2H}, m_{3T}), H) - u_2(m_{2H}, (m_{1H}, m_{3T}), H)\big]
\end{aligned}
$$

Now, note the following:

(i) If $\theta_1 = \theta = \theta_2$, then player 3 is *not* pivotal, so

$$
u_i(\theta, (\theta, m_{3T}), \theta) = 0 \;\; \forall m_{3T} \;\; (i = 1,2; \; \theta = -1, 1)
$$

(ii) If $\theta_i = L, \; \theta_j = H \; (i, j \in \{1,2\}, \; i \neq j)$, then player 3 *is* pivotal, so

$$
u_i(-1, (1, m_{3T}), -1) = \begin{cases} 0 & \text{, if } m_{3T} = -1 \\ -4 & \text{, if } m_{3T} = 1 \end{cases} , \quad u_i(-1, (1, m_{3T}), 1) = \begin{cases} -4 & \text{, if } m_{3T} = -1 \\ 0 & \text{, if } m_{3T} = 1 \end{cases} .
$$

It then follows that

$$
\mathbb{E}_{\theta_{-3}}[u_3(m_{3T}, \mathbf{m}^*_{-3,\theta_{-3}}, \theta_{-3})] = 8p(1-p) \;\forall m_{3T} \in \{-1, 1\}
$$

Hence, player 3 is indifferent between choosing $m_{3T} = -1$ and $m_{3T} = 1$ for all $p \in (0,1)$. ■

A.2.2 Proof of Lemma 6

The structure of this derivation is *very* similar to that of the previous section. We first solve the genuine players' ($i = 1, 2$) problems using an exhaustive method. Afterward, we compare $\mathbb{E}_{\theta_{-3}}[u_3(-1,\mathbf{m}^*_{-3,\theta_{-3}},\theta_{-3})]$ and $\mathbb{E}_{\theta_{-3}}[u_3(1,\mathbf{m}^*_{-3,\theta_{-3}},\theta_{-3})]$ to derive $m^*_{3T}(\cdot)$. The only difference in this derivation is that now $x(\cdot)$ is given by (2.3) instead of (2.1).

Genuine Players' ($i=1,2$) Problems

<u>**Case 1:**</u> $EU_{1L}((-1,(-1,1;-1)) - EU_{1L}((1,(-1,1;-1)) = \frac{4}{3} - \frac{8p}{9} > 0$

$\Rightarrow m^*_{1L}((1,(-1,1;-1))) = -1,$

<u>**Case 2:**</u> $EU_{1L}((-1,(-1,1;1)) - EU_{1L}((1,(-1,1;1)) = \frac{20}{9} - \frac{8p}{9} > 0$

$\Rightarrow m^*_{1L}((1,(-1,1;1))) = -1,$

<u>**Case 3:**</u> $EU_{1L}((-1,(1,-1;-1)) - EU_{1L}((1,(1,-1;-1)) = \frac{8p}{9} + \frac{4}{9} > 0$

$\Rightarrow m^*_{1L}((1,(1,-1;-1))) = -1,$

<u>**Case 4:**</u> $EU_{1L}((-1,(1,-1;1)) - EU_{1L}((1,(1,-1;1)) = \frac{8p}{9} + \frac{4}{3} > 0$

$\Rightarrow m^*_{1L}((1,(1,-1;1))) = -1,$

<u>**Case 5:**</u> $EU_{1L}((-1,(-1,-1;-1)) - EU_{1L}((1,(-1,-1;-1)) = \frac{4}{9} > 0$

$\Rightarrow m^*_{1L}((1,(-1,-1;-1))) = -1,$

<u>**Case 6:**</u> $EU_{1L}((-1,(-1,-1;1)) - EU_{1L}((1,(-1,-1;1)) = \frac{4}{3} > 0$

$\Rightarrow m^*_{1L}((1,(-1,-1;1))) = -1,$

<u>**Case 7:**</u> $EU_{1L}((-1,(1,1;-1)) - EU_{1L}((1,(1,1;-1)) = \frac{4}{3} > 0$

$\Rightarrow m^*_{1L}((1,(1,1;-1))) = -1,$

<u>**Case 8:**</u> $EU_{1L}((-1,(1,1;1)) - EU_{1L}((1,(1,1;1)) = \frac{20}{9} > 0$

$\Rightarrow m^*_{1L}((1,(1,1;1))) = -1,$

Hence, player 1 (of type L) votes sincerely (i.e., $m^*_{1L}(\cdot) \equiv -1$. As before, the derivations for player 2 and for high types are omitted since they are *very* similar. Players 1 and 2 of either type vote sincerely (i.e., $m^*_{i\theta_i}(\cdot) \equiv \theta_i,\ \forall i \in \{1,2\}, \forall \theta_i \in \{-1,1\}$).

Saboteur's Problem (Player 3)

Solving player 3's problem is much simpler this time:

$$\mathbb{E}_{\theta_{-3}}[u_3(-1,\mathbf{m}^*_{-3,\theta_{-3}},\theta_{-3})]-\mathbb{E}_{\theta_{-3}}[u_3(1,\mathbf{m}^*_{-3,\theta_{-3}},\theta_{-3})]=\frac{8\,(p-1)^2}{9}-\frac{8\,p^2}{9}=\frac{8}{9}-\frac{16\,p}{9}$$

which is positive if and only if $p>\frac{1}{2}$. Hence,

$$m^*_{3T}\equiv\begin{cases}-1 & \text{, if } p<\frac{1}{2}\\ \{-1,1\} & \text{, if } p=\frac{1}{2}\\ 1 & \text{, if } p>\frac{1}{2}\end{cases}$$

■

A.2.3 Trolls' Behavior Under "Majority Rule"

When there are N agents, 2 types (γ_1 and γ_2), T trolls, and the voting mechanism is "Majority Rule", trolls will always vote for the less likely type.

Let $n_i(\theta)\in\{0,1,...,N\}$ denote the number of agents that are type γ_i $(i=1,2)$, given the realization $\theta\in\{\gamma_1,\gamma_2\}^N$. Since there are two types, $n_2(\theta)\equiv N-n_1(\theta)$.[2] Let

$$\varphi(n_1):=\binom{N}{n_1}p^{n_1}(1-p)^{(N-n_1)} \tag{A.4}$$

denote the probability that n_1 players are of the first type.

Notice that trolls are pivotal if and only if $|n_1-n_2|\leq T$. This is equivalent to

$$\frac{N-T}{2}\leq n_1\leq\frac{T+N}{2}.$$

For simplicity, assume that $\frac{N-T}{2},\frac{N+T}{2}\notin\mathbb{N}$ and that $T<N$.[3] Trolls have two strategies to

[2]The input for n_1 will be suppressed throughout.

[3]These assumptions are by no means necessary. In a more exhaustive proof, where $N,T\in\mathbb{N}$, we would need to use some tie-breaking rule for when the $n_1=\frac{N-T}{2}$ and $n_1=\frac{N+T}{2}$. The assumption $T<N$ is just so

compare: vote for $\theta = \gamma_1$ or for $\theta = \gamma_2$. We will show that trolls optimally vote for γ_2 iff $p > \frac{1}{2}$.

Denote the trolls' message by m_T and let $EU_T(m_T)$ denote their expected utility from voting m_T.[4] Then, the trolls will optimally pick $m_T^* = \gamma_2$ iff $EU_T(\gamma_1) - EU_T(\gamma_2) < 0$. After some (omitted) algebraic simplification, we get

$$\begin{aligned}
EU_T(\gamma_1) - EU_T(\gamma_2) &= \sum_{n_1=\lfloor\frac{N-T}{2}\rfloor+1}^{\lfloor\frac{N+T}{2}\rfloor} \varphi(n_1)\left[(N-n_1)(\gamma_1-\gamma_2)^2 - n_1(\gamma_1-\gamma_2)^2\right] \\
&= \sum_{n_1=\lfloor\frac{N-T}{2}\rfloor+1}^{\lfloor\frac{N+T}{2}\rfloor} \varphi(n_1)(N-2n_1)(\gamma_1-\gamma_2)^2 \\
&= \sum_{n_1=\lfloor\frac{N-T}{2}\rfloor+1}^{\lfloor\frac{N}{2}\rfloor} \left(\varphi(n_1)-\varphi(N-n_1)\right)(N-2n_1)(\gamma_1-\gamma_2)^2,
\end{aligned}$$

where the last equality is due to the symmetry of $N - 2n_1$ around $n_1 = \frac{N}{2}$. Next, note that due to the symmetry of binomial coefficients, we have

$$\varphi(n_1) - \varphi(N-n_1) = \binom{N}{n_1}\left(p^{n_1}(1-p)^{N-n_1} - p^{N-n_1}(1-p)^{n_1}\right),$$

which is negative for any $n_1 < \frac{N}{2}$ iff $p > \frac{1}{2}$. Thus, if $p > \frac{1}{2}$, we have $EU_T(\gamma_1) - EU_T(\gamma_2) < 0$, and the trolls' optimal strategy is to vote for the less likely type, $m = \gamma_2$.

"■"

that we focus on the less trivial part of the complete proof. When $T \geq N$, trolls are always pivotal.

[4]Here, trolls will always cast the same votes as one another, and hence can be treated as one player in this proof.

A.2.4 Proof of Lemma 8

Proof. The difference between the designer's expected utility when trolls imitate γ_1 and γ_2 can be shown to satisfy

$$\Delta EU = -\sum_{k=0}^{N} \binom{N}{k} p^k (1-p)^{N-k} \left(k \left(\frac{N-k}{N+T} \right)^2 + (N-k) \left(\frac{T+k}{N+T} \right)^2 \right) (\gamma_1 - \gamma_2)^2$$
$$+ \sum_{k=0}^{N} \binom{N}{k} p^k (1-p)^{N-k} \left(k \left(\frac{T+N-k}{N+T} \right)^2 + (N-k) \left(\frac{k}{N+T} \right)^2 \right) (\gamma_1 - \gamma_2)^2$$
$$\sim \sum_{k=0}^{N} \binom{N}{k} \left(k \left(\frac{N-k}{N+T} \right)^2 + (N-k) \left(\frac{T+k}{N+T} \right)^2 \right) \left((1-p)^k p^{N-k} - p^k (1-p)^{N-k} \right)$$

The last line is due to us removing $(\gamma_1 - \gamma_2)^2$ to simplify the expression (it is a common positive term). Assuming N is odd[5], we can represent the sum above as follows:

$$\Delta EU \sim \sum_{k=0}^{\frac{N-1}{2}} \binom{N}{k} \left((1-p)^k p^{N-k} - p^k (1-p)^{N-k} \right) \left[k \left(\frac{N-k}{N+T} \right)^2 + (N-k) \left(\frac{T+k}{N+T} \right)^2 - \right.$$
$$\left. - (N-k) \left(\frac{k}{N+T} \right)^2 - k \left(\frac{T+N-k}{N+T} \right)^2 \right].$$

Note that for $k < \frac{N}{2}$, we have $(1-p)^k p^{N-k} - p^k (1-p)^{N-k} > 0$, since $p > \frac{1}{2}$. Also note:

$$k \left(\frac{N-k}{N+T} \right)^2 + (N-k) \left(\frac{T+k}{N+T} \right)^2 - (N-k) \left(\frac{k}{N+T} \right)^2 - k \left(\frac{T+N-k}{N+T} \right)^2 =$$
$$= (N-2k) \frac{T^2}{(N+T)^2} + 0 > 0,$$

since $k < \frac{N}{2}$.

[5]The argument follows very similarly if N is even.

Therefore, the difference ΔEU is strictly positive term by term. This implies that the designer achieves higher utility when trolls report more likely type ($\gamma = 2$ since we have $p > \frac{1}{2}$). Hence, trolls will optimize by picking the less likely type ($\gamma = 1$) to report. ■

A.2.5 Proof of Lemma 9

Proof. It can be shown that the expected utility of the designer (ex-ante) is given by

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{N} u_i(g(\theta,T),\theta_i)\right] &= -\sum_{k=0}^{N}\binom{N}{k} p^k(1-p)^{N-k}\left(k\left(\frac{T+N-k}{N+T}\right)^2 + (N-k)\left(\frac{k}{N+T}\right)^2\right)\\ &= -\frac{Np\left(N^2(1-p)+N(1-p)(2T-1)+T(2p+T-2)\right)}{(N+T)^2}\end{aligned}$$

This sum is hard to interpret, but we can check comparative statics of it with respect to T and p. Straightforward algebraic calculations show that

$$\frac{\partial}{\partial T}\mathbb{E}\left[\sum_{i=1}^{N} u_i(g(\theta,T),\theta_i)\right] = -\frac{2Np(NT+p^2-3p+2)}{(N+T)^3} < 0,$$

which implies that having more trolls will strictly reduce the designer's welfare. This is expected, since more trolls will be able to bias the result of the mechanism in a more drastic way.

For the blind mechanism, the designer's expected utility under is given by

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{N} u_i(p,\theta_i)\right] &= -\sum_{k=0}^{N}\binom{N}{k} p^k(1-p)^{N-k}\left(k\,(1-p)^2 + (N-k)\,(2-p)^2\right)\\ &= -Np(1-p).\end{aligned}$$

The blind mechanism performs better than the average-of-votes mechanism if and only if

$$-Np(1-p) > -Np\frac{N^2(1-p)+N(1-p)(2T-1)+T(2p+T-2)}{(N+T)^2}.$$

Rearranging and simplifying:

$$\begin{aligned}
-(1-p) &> -\frac{N^2(1-p)+N(1-p)(2T-1)+T(2p+T-2)}{(N+T)^2} \\
p-1 &> -\frac{N^2(1-p)+N(1-p)(2T-1)+T(2p+T-2)}{(N+T)^2} \\
p &> \frac{pN^2+2pNT+(1-p)N+2(1-p)T}{(N+T)^2} \\
(N+T)^2 p &> (N^2+2NT-N-2T)p+N+2T \\
p &> \frac{N+2T}{N+2T+T^2}.
\end{aligned}$$

■

A.2.6 Proof of Proposition 1

Proof. Suppose that the observed number of votes for type γ_1 is $k \in \{0,1,\ldots,N+T\}$. The lowest number of trolls that can be in this number is 0 (if all of them voted for type γ_2), and the highest number is T (if all of them voted for type γ_1). Therefore, the highest possible number of genuine agents with type γ_1 is k and the lowest possible number is $\max\{k-T,0\}$. The true distribution of genuine types is anything in the range from $(\max\{k-T,0\}, \min\{N-k+T,N\})$ to $(k, N-k)$.

Note that the optimal outcome is decreasing in the number of genuine agents of type γ_1. Hence, it follows that the largest outcome that could be optimal is

$$\underline{b} = \frac{\max\{k-T,0\}}{N}\gamma_1 + \frac{\min\{N-k+T,N\}}{N}\gamma_2, \text{ (if number of } \gamma_1 \text{ types is lowest)}$$

and the smallest outcome that could be optimal is

$$\bar{b} = \frac{k}{N}\gamma_1 + \frac{N-k}{N}\gamma_2. \text{ (if number of } \gamma_1 \text{ types is highest)}$$

Let $g(k)$ be the mechanism's outcome under the mechanism g. If $g(k) < \underline{b}$, the expected welfare under the mechanism can be improved if we set $g(k) = \underline{b}$. Similarly, if $g(k) > \bar{b}$, the expected welfare can be improved if we set $g(k) = \bar{b}$.

■

A.2.7 Proof of Proposition 2

Proof. Define a mechanism as an outcome rule $g : \{0, 1, \ldots, N+T\} \to [\gamma_1, \gamma_2]$, where the argument is the number of votes for γ_1. The designer chooses this rule to maximize the ex-ante utility subject to the trolls' best response.

Suppose that for a given mechanism g' the trolls are not indifferent between $m = \gamma_1$ and $m = \gamma_2$. For concreteness, assume that they prefer $m = \gamma_1$. Fixing the trolls' action, the designer's ex-ante utility is continuous in $g(0)$, $g(1)$, …, $g(N+T)$. Therefore, in a neighborhood[6] of g' the trolls' strategy can be treated as a constant in that it does not change when the designer slightly adjusts $g'(0)$, $g'(1)$, …, $g'(N+T)$. There are two possibilities: either g' is a local (and global[7]) maximum, or the designer can improve upon it. We will prove that the former option is not possible.

If we fix the trolls' strategy at $m = \gamma_1$, the designer's best reply is to essentially "subtract" the trolls' votes from the total. That is, if the designer observes k votes for γ_1, she then knows that $k - T$ genuine agents have this type and $N + T - k$ genuine agents have the

[6]I.e. a set of mechanisms g such that $\left|\left|\big(g(0), g(1), \ldots, g(N+T)\big) - \big(g'(0), g'(1), \ldots, g'(N+T)\big)\right|\right| < \varepsilon$ for some $\varepsilon > 0$, where $||\cdot||$ is the Euclidean norm.

[7]This is due to the concavity of the designer's utility function.

other type. Then the designer's optimal mechanism is g_1, where

$$\begin{aligned}
&\max_{g_1(k)} -(k-T)\left(g_1(k)-\gamma_1\right)^2-(N+T-k)\left(g_1(k)-\gamma_2\right)^2 \\
&\implies g_1(k)=\frac{k-T}{N}\gamma_1+\frac{N+T-k}{N}\gamma_2.
\end{aligned}$$

Note that g_1 completely neutralizes the trolls' influence and achieves the same utility level as under perfect information. This, however, cannot be the equilibrium, since the trolls can benefit by switching some of their votes to $m=\gamma_2$. In that case, the mechanism will not take optimal action given any distribution of votes, and the designer's ex-ante utility will be lower. Hence, the trolls would prefer to deviate from $m=\gamma_1$.
This implies that the designer's optimal mechanism cannot be g_1, or any mechanism for which the trolls strictly prefer message $m=\gamma_1$. A similar proof can be done for the mechanisms for which the trolls prefer $m=\gamma_2$. Hence, the optimal mechanism must make the trolls indifferent between the messages. ■

A.2.8 Proof of Proposition 3

Proof. Assume $N+T$ is odd. The proof below can be easily adapted to the case where $N+T$ is even.

We will show this result by proving that changing the majority rule g_{mr} to a supermajority rule $g_{smr}^{\hat{\alpha},\gamma_1}$ with $\hat{\alpha}=\frac{1}{2}+\frac{1}{N+T}$ is always strictly beneficial for the designer. This is true regardless of how the trolls respond to the change of the mechanism.

First, suppose that the trolls' strategy remains the same. Recall that in the majority rule, the trolls strictly prefer to vote for $\theta=\gamma_2$. Thus, the mechanism $g_{smr}^{\alpha,\gamma_1}$ will differ in its outcome from g_{mr} in only one instance: when there are $\hat{k}=\left[\frac{N+T}{2}\right]+1$ votes for γ_2. The majority rule has $g_{mr}\left(\hat{k}\right)=\gamma_2$, but the supermajority rule has $g_{smr}^{\hat{\alpha},\gamma_1}\left(\hat{k}\right)=\gamma_1$. This is a beneficial change because in this instance there are more voters with type $\theta=\gamma_1$ than

$\theta = \gamma_2$. Hence, the change to $g_{smr}^{\hat{\alpha},\gamma_1}$ leads to a higher expected welfare.

We will now verify that this improvement remains if the trolls switch their strategy from voting for $\theta = \gamma_2$ to voting for $\theta = \gamma_1$. In this case, $g_{smr}^{\hat{\alpha},\gamma_1}$ will have the same outcomes (and the same expected welfare) as g_{mr} if the trolls voted for $\theta = \gamma_1$ instead of $\theta = \gamma_2$. The expected welfare under g_{mr} is strictly higher if the trolls vote for $\theta = \gamma_1$ than if they vote for $\theta = \gamma_2$. This implies that the change to $g_{smr}^{\hat{\alpha},\gamma_1}$ leads to a higher expected welfare. ■

A.2.9 Proof of Proposition 4

Proof. Recall that for $\beta = 1$, trolls prefer voting for $\theta = \gamma_2$, since $p > \frac{1}{2}$. For β close to 1, this will remain true due to the expected welfare's continuity in β (see below). The weighted-average-of-votes mechanism's expected welfare is

$$V(g_{am}^{\beta}) = -\sum_{i=0}^{N} \binom{N}{i} p^i (1-p)^{N-i} \left(i \left(\frac{N+T-i}{i\beta + (N+T-i)} \right)^2 + \right.$$
$$\left. + (N-i) \left(\frac{i\beta}{i\beta + (N+T-i)} \right)^2 \right) (\gamma_1 - \gamma_2)^2.$$

To show the result, we will take the derivative with respect to β and show that it is negative at $\beta = 1$. We will drop the (positive) term $(\gamma_1 - \gamma_2)^2$ to simplify the calculations.

$$\frac{\partial}{\partial \beta} V(g_{am}^{\beta}) \sim -\sum_{i=0}^{N} \binom{N}{i} p^i (1-p)^{N-i} \left(-\frac{2i^2 (N+T-i)^2}{(i\beta + (N+T-i))^3} + \right.$$
$$\left. + (N-i) \frac{2i^2 \beta (i\beta + (N+T-i)) - 2i^3 \beta^2}{(i\beta + (N+T-i))^3} \right)$$
$$= -\sum_{i=0}^{N} \binom{N}{i} p^i (1-p)^{N-i} 2i^2 (N+T-i) \frac{(N+T-i) + (N-i)\beta}{(i\beta + (N+T-i))^3}.$$

All terms are positive except for the negative sign at the front, so $\frac{\partial}{\partial\beta}V(g^{\beta}_{am}) < 0$ at $\beta = 1$. Due to continuity of the expected utility, the trolls' strategy will remain the same for a range $\beta \in (\bar{\beta}, 1)$ for some $\bar{\beta}$. Hence, the designer can improve upon the mechanism's expected welfare by reducing β. ■

A.2.10 Proof of Proposition 5

Proof of Proposition 5. We know that g is continuous, i.e.

$$\forall \varepsilon > 0\ \exists \delta > 0 \text{ s.t. } \forall p, p' \in \Delta\Gamma,\ |p - p'| < \delta \ \Rightarrow\ |g(p) - g(p')| < \varepsilon,$$

where $|\cdot|$ denotes the standard Euclidean norm.

Define $t = \min_{x \in g(\Delta\Gamma)} \mathbb{E}\left[\sum_{i=1}^{N} u_i(x, \theta_i)\right]$ and $p_t = \arg\min_{x \in g(\Delta\Gamma)} \mathbb{E}\left[\sum_{i=1}^{N} u_i(x, \theta_i)\right]$. Due to continuity of g, there exists a neighborhood of p_t where the mechanism's outcomes are close to t, the trolls' ideal outcome. We want to prove that as $T \to \infty$, the trolls will be able to get into that neighborhood no matter the distribution of other agents' types.

For T trolls, let $p(T)$ be a distribution of their votes (excluding normal agents) that is closest to p_t. Formally, let $F(T) = \{p \in \Delta\Gamma \mid \forall \gamma_i, p(\gamma_i) = \frac{k}{T} \text{ for some } k \in \mathbb{N}\}$, and define $p(T)$ as

$$p(T) = \min_{p \in F(T)} |p - p_t|.$$

Note that $F(T)$ for different $T \in \mathbb{N}$ can comprise points with arbitrary rational coordinates. Since rational numbers are dense in $\mathbb{R}$, it follows that we can always find sufficiently large T so that $|p(T) - p_t|$ is arbitrarily close to 0. Formally, for any $\delta > 0$ there exists $\hat{T}_1 \in \mathbb{N}$ such that

$$T > \hat{T}_1 \ \Rightarrow\ |p(T) - p_t| < \frac{\delta}{2}.$$

This will be useful later on.

$p(T)$ is the distribution of votes that trolls generate on their own. Now consider possible variations that can be introduced to $p(T)$ due to normal agents' (sincere) votes. Let $\theta = (\theta_1, \ldots, \theta_N)$ be vector of normal agents' types, and let $p(\theta, T)$ be the distribution of votes with normal agents and trolls included. That means

$$p(\theta,T)(\gamma_i) = \frac{T \cdot p(T)(\gamma_i) + \sum_{j=1}^{N} \mathbb{1}\{\theta_j = \gamma_i\}}{N+T}.$$

The distance between total distribution of votes and trolls' distribution of votes is given by

$$\begin{aligned} |p(T) - p(\theta,T)| &= \sqrt{\sum_{\gamma_i \in \Gamma} \Big(p(T)(\gamma_i) - p(\theta,T)(\gamma_i) \Big)^2} \\ &= \sqrt{\sum_{\gamma_i \in \Gamma} \left(\frac{1}{N+T} \cdot \Big(N \cdot p(T)(\gamma_i) - \sum_{j=1}^{N} \{\theta_j = \gamma_i\} \Big) \right)^2}. \end{aligned}$$

Due to quadratic nature of the norm, it will achieve its maximum when all normal agents are of the same type. Thus, the distance will take form

$$\begin{aligned} |p(T) - p(\theta,T)| &= \sqrt{\left(\frac{N}{N+T} \cdot (p(T)(\gamma) - 1) \right)^2 + \sum_{\gamma_i \neq \gamma} \left(\frac{N}{N+T} p(T)\gamma_i \right)^2} \\ &= \frac{N}{N+T} \cdot \sqrt{\Big(p(T)(\gamma) - 1 \Big)^2 + \sum_{\gamma_i \neq \gamma} \Big(p(T)\gamma_i \Big)^2} \end{aligned}$$

for some $\gamma \in \Gamma$. Clearly, $\lim_{T\to\infty} |p(T) - p(\theta,T)| = 0$, since the expression under the square root is bounded by $|\Gamma|$, which is finite. Thus, for any $\delta > 0$ there exists $\hat{T}_2 \in \mathbb{N}$ such that

$$T > \hat{T}_2 \;\Rightarrow\; |p(T) - p(\theta,T)| < \frac{\delta}{2} \text{ for any } \theta.$$

Now we can take the maximum of $\hat{T}_1$ and $\hat{T}_2$ that will ensure

$$T > \hat{T} = \max\{\hat{T}_1, \hat{T}_2\} \;\Rightarrow\; |p(T) - p_t| < \frac{\delta}{2} \text{ and } |p(T) - p(\theta, T)| < \frac{\delta}{2}.$$

Finally, note that

$$|p(\theta, T) - p_t| \leq |p(\theta, T) - p(T)| + |p(T) - p_t|.$$

Therefore, we can conclude that

$$T > \hat{T} \;\Rightarrow\; |p(\theta, T) - p_t| < \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

Hence,

$$\forall \varepsilon > 0 \; \exists \hat{T} \in \mathbb{N} \text{ such that } T > \hat{T} \;\; |g(p(\theta, T)) - g(p_t)| < \varepsilon.$$

If T is large enough, trolls can guarantee that the outcome of mechanism g is arbitrarily close to the ex-ante worst-case outcome $g(p_t)$. This finishes the proof.

■

A.3 Appendix of Chapter 3

A.3.1 Instructions in the Experiment

The Study

This study examines the people's decisions about doing work for monetary payments. For this study, we have designed a task that you can choose to do for different wages. The tasks are of no value to us, beyond understanding these decisions.

We are interested in these work decisions at different points in time. For example, you might be asked on April 1 how many tasks you want to do on April 15 for various wages. When April 15 arrives, you will be asked how many tasks you want to do on that day for various wages. Note that your decisions about future work do not have to be the same as the eventual decisions about work when the day arrives. **You will actually have to do the tasks specified in one of these decisions, so treat each choice you make as if it will be the one that will determine your tasks.** We refer to the tasks that will actually be done as the **decision-that-counts**, and elaborate below on how exactly this decision is chosen.

We are also interested in people's preferences about which decisions are implemented. For example, you might be asked on April 1 how many tasks you want to do on April 15 for different wages *as well as* how much you prefer to actually do your April 1 decisions about April 15 tasks over your April 15 decisions about April 15 tasks. That is, we will ask you how you would like to determine the decision-that-counts.

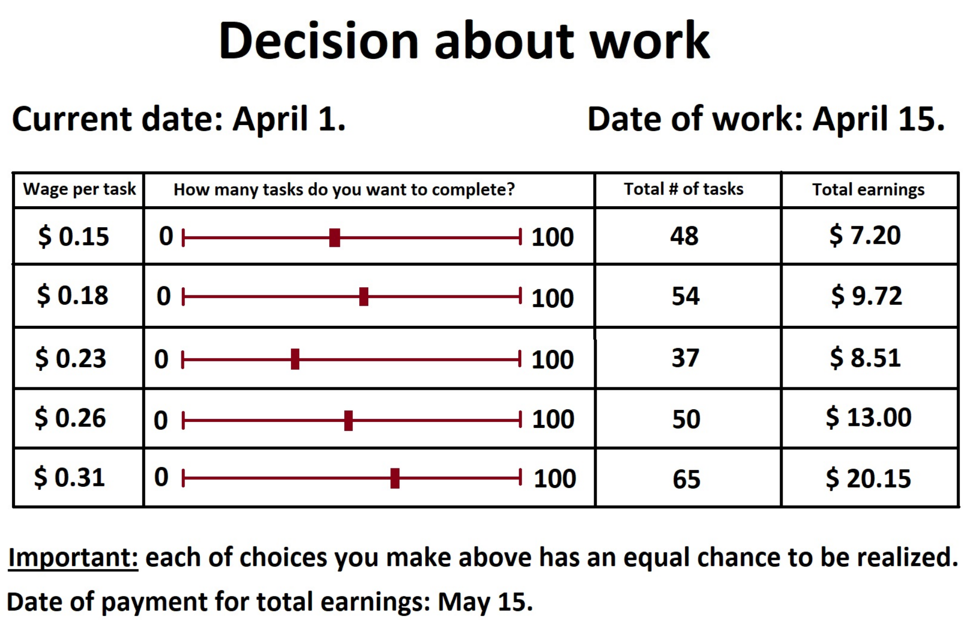
Work Decisions

Recall that we are interested in people's decisions about doing work for monetary payments. In each working session, you will have to complete 10 mandatory tasks, and any amount of supplemental tasks that you choose. You will be asked a series of questions regarding your preferences for completing these supplementary tasks.

We are interested in these decisions at different points in time. You will be asked questions about how many tasks you want to do at various points in the future and how many tasks you want to do in the present day for a set of different wages. For example, you may be asked on April 1, April 8, and April 15 how many tasks you would like to do on April 15. The decisions made on April 1 and April 8 are called your **decisions about future work** and the decisions made on April 15 are called your **decisions about current**

work. Note that your decisions about future work may be the same or may be different from your decisions about current work.

Your screen will have a list of various wages as well as a date of work. You will be asked how many tasks you would like to do at each wage on the date of work. Next to each wage is a slider bar with values between 0 and 100; you will use this slider bar to indicate how many tasks you would like to complete at each wage. Below is an example of how the decision screen might look like.


Decision about work

Current date: April 1. Date of work: April 15.

Wage per task	How many tasks do you want to complete?	Total # of tasks	Total earnings
$ 0.15	0 — 100	48	$ 7.20
$ 0.18	0 — 100	54	$ 9.72
$ 0.23	0 — 100	37	$ 8.51
$ 0.26	0 — 100	50	$ 13.00
$ 0.31	0 — 100	65	$ 20.15

Important: each of choices you make above has an equal chance to be realized.

Date of payment for total earnings: May 15.

Figure A.1: Example of a work decision screen.

Your decisions cannot affect which future wages might be realized. **One of these decisions will be randomly selected to become the decision-that-counts, so you will actually have to do the tasks specified in one of these decisions. In other words, each decision that you report has a positive chance of becoming the one you will need to complete. Therefore, it is in your interest to treat each choice you make as if it will be the one that will determine your tasks and answer honestly about your work pref-**

erences.

Determination of Decision-that-counts

In addition to people's decisions over doing work for monetary payments at different points in time, we are also interested in people's preferences over which of those decisions should be implemented. Therefore, in addition to making decisions about how many tasks to do at a given wage, you will also be asked how much you want decisions made at different points in time to be the decision-that-counts. To do this, you will be asked to provide a **decision-that-counts percentage** (henceforth DTC percentage) in each of your decisions about future work. For example, you may be asked for DTC percentage decisions on April 1 and April 8 about work on April 15. Below is an example of how the decision screen might look like.

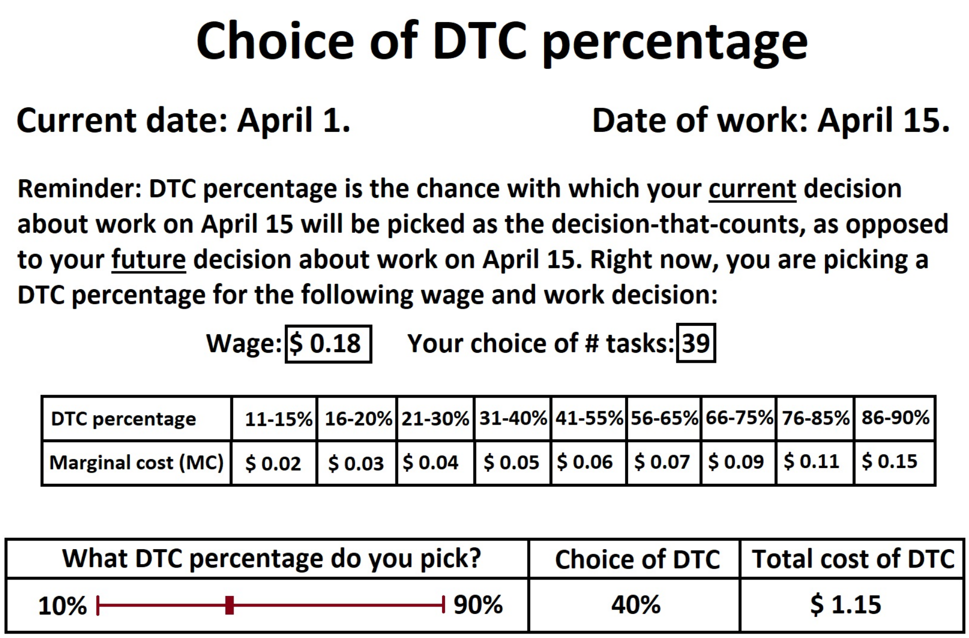


Figure A.2: Example of a DTC percentage decision screen.

You will use your slider bar to select a number **between 10% and 90%. The DTC percent-**

age is the probability that your decision about future work will become the decision-that-counts. Therefore, 1 minus the DTC percentage is the probability that your decision about current work will become the decision-that-counts. For example, if on April 1 you selected an DTC percentage of 30% for your April 15 decision, then there is a 30% chance that your April 1 decision will be the decision-that-counts and a 70% chance that your April 15 decision will be the decision-that-counts. Also, each possible DTC percentage choice may be associated with a monetary cost or monetary reward, denoted as **cost** or **reward**, respectively. **These costs or rewards are only paid for the decision-that-counts, and each of the choices you make has a positive chance of being realized.**

To illustrate this, suppose that on April 1 and 8 you are making DTC percentage decisions about work on April 15. Suppose that on April 1, at a wage of $0.18 you selected a DTC percentage of 30% which costs $0.10, and at a wage of $0.25 you selected a DTC percentage of 40% which costs $0.20. Suppose that on April 8, at a wage of $0.18 you selected an DTC percentage of 35% which costs $0.15, and at a wage of $0.25 you selected an DTC percentage of 40% which costs $0.20. If the decision-that-counts is your April 1 decision at a wage of $0.18, you would have to pay the cost of $0.10. If the decision-that-counts is your April 15 decision at a wage of $0.18, you would have to pay the cost of $0.10. In neither case, however, will you have to pay the cost of $0.20 from your decisions at a wage of $0.25, nor would you have to pay the cost of $0.15 from your April 8 decision at the wage of $0.18. All monetary costs or rewards will be deducted from or added to your earnings at the end of the experiment. **Your DTC percentages will be used to to determine the decision-that-counts, so they will affect the tasks you will actually have to do. Therefore, it is in your interest to treat each choice you make as if it will be the one that will determine your tasks and answer honestly about your preferences.**

Selection of the Decision-that-counts

The selection of the decision-that-counts will occur in three stages.

In the first stage, we will determine which wage determines your decision-that-counts. Since you made decisions across different weeks for five different wages, we will randomly select a number between 1 and 5 to determine which wage's decisions will be used to determine the decision-that-counts. Note that since you made decisions for five different wages, all wages are equally likely to determine the decision-that-counts.

The second stage determines which of your decisions about future work will be used to determine the decision that counts. If you made two decisions about future work for the date in question, we will randomly select a number between 1 and 2 to determine which decision about future work will be used to determine the decision that counts. For example, if you were making decisions about April 15 work on April 1 and April 8, this stage will determine which of the latter two dates will be used in the next stage. Note that since you made decisions on two previous dates, all decisions about future work are equally likely to determine the decision-that-counts.

The third stage determines whether your decision-that-counts will come from an decision about future work or a decision about current work. **We will randomly select a number between 1 and 100. If that number is less than or equal to your chosen DTC percentage for the wage selected in stage 1 on the date selected in step 2, the decision-that-counts will be the decision about future work (from the date selected in step 2). If that number is greater than your chosen DTC percentage for the wage selected in stage 1 on the date selected in step 2, the decision-that-counts will be the decision about current work. Therefore, the DTC percentage that you select precisely determines the likelihood with which your decision about future work or you decision about current work becomes the decision-that-counts.** For example, if at the wage selected in stage 1 and the decision about future work selected in stage 2 you selected an DTC percentage of 30%, then there is a 30% chance that your decision about future work will be the decision-that-counts and a 70% chance that your decision about current work will be the decision-that-counts.

At every stage, you will see a screen with a random number generator, so that you can explicitly observe the process of selecting the decision-that-counts.

Bibliography

Augenblick, N., M. Niederle, and C. Sprenger (2015). Working over time: Dynamic inconsistency in real effort tasks. *The Quarterly Journal of Economics 130*(3), 1067–1115.

Augenblick, N. and M. Rabin (2019). An experiment on time preference and misprediction in unpleasant tasks. *The Review of Economic Studies 86,* 941–975.

Aziz, H., Y. Bachrach, E. Elkind, and M. Paterson (2011). False-name manipulations in weighted voting games. *Journal of Artificial Intelligence Research 40,* 57–93.

Bachrach, Y. and E. Elkind (2008). Divide and conquer: False-name manipulations in weighted voting games. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 2,* pp. 975–982. Citeseer.

Bergemann, D. and U. Hege (2005). The Financing of Innovation: Learning and Stopping. *The RAND Journal of Economics 36*(4), 719–752.

Bergemann, D. and J. Välimäki (2001). Stationary Multi-Choice Bandit Problems. *Journal of Economic Dynamics and Control 25*(10), 1585–1594.

Bhargava, H. K. (2022). The Creator Economy: Managing Ecosystem Supply, Revenue Sharing, and Platform Design. *Management Science 68*(7), 5233–5251.

Bolton, P. and M. Dewatripont (2005). *Contract Theory,* Volume 1 of *MIT Press Books*. The MIT Press.

Bonatti, A. and J. Hörner (2011, April). Collaborating. *American Economic Review 101*(2), 632–63.

Bonatti, A. and J. Hörner (2017). Career Concerns with Exponential Learning. *Theoretical Economics 12*(1), 425–475.

Brahma, S., L. Njilla, and S. Nan (2022). Optimal auction design with malicious sellers. In *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops),* pp. 661–666.

British Broadcasting Corporation (2016, Mar). Four times when having an online poll was a bad idea. https://www.bbc.com/news/uk-35860830.

Carrera, M., H. Royer, M. Stehr, J. Sydnor, and D. Taubinsky (2019). How are preferences for commitment revealed? *NBER Working Paper 26161*.

Chen, D. L., M. Shonger, and C. Wickens (2016). otree–an open source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance 9(C)*, 88–97.

Chorppath, A. K. and T. Alpcan (2011). Adversarial behavior in network mechanism design. In *Proceedings of the 5th International ICST Conference on Performance Evaluation Methodologies and Tools*, pp. 506–514.

Chowdhury, S. M. and O. Gürtler (2015). Sabotage in contests: a survey. *Public Choice 164*(1-2), 135–155.

Collins, B. and B. Popken (2019, Jun). Trolls target online polls following first democratic presidential debate. *National Broadcasting Company News*. https://www.nbcnews.com/tech/tech-news/trolls-target-online-polls-following-first-democratic-presidential-debate-n1023406.

Conitzer, V. (2008). Anonymity-proof voting rules. In C. Papadimitriou and S. Zhang (Eds.), *Internet and Network Economics*, Berlin, Heidelberg, pp. 295–306. Springer Berlin Heidelberg.

Elkind, E., P. Faliszewski, and A. Slinko (2011). Cloning in elections: Finding the possible winners. *Journal of Artificial Intelligence Research 42*, 529–573. cloning on the other side of the ballot box.

Fioravanti, F. and J. Massó (2022). False-name-proof and strategy-proof voting rules under separable preferences. *Available at SSRN 4175113*.

Frenkel, S., K. Browning, and T. Lorenz (2020, Jun). Tiktok teens and k-pop stans say they sank trump rally. *New York Times*.

Gary-Bobo, R. J. and T. Jaaidane (2000). Polling mechanisms and the demand revelation problem. *Journal of Public Economics 76*(2), 203 – 238.

Gomes, R., D. Gottlieb, and L. Maestri (2016). Experimentation and project selection: Screening and learning. *Games and Economic Behavior 96*, 145–169.

Gomes, R. and A. Pavan (2016). Many-to-many matching and price discrimination. *Theoretical Economics 11*(3).

Gul, F. and W. Pesendorfer (2003). Temptation and self-control. *Econometrica 69*(6), 1403–1435.

Guo, Y. (2016, August). Dynamic delegation of experimentation. *American Economic Review 106*(8), 1969–2008.

Halac, M., N. Kartik, and Q. Liu (2016, 03). Optimal Contracts for Experimentation. *The Review of Economic Studies 83*(3), 1040–1091.

Hirsch, A. V. and J. P. Kastellec (2022). A theory of policy sabotage. *Journal of Theoretical Politics 34*(2), 191–218.

Hörner, J. and A. Skrzypacz (2016). Learning, Experimentation and Information Design. Working paper.

Invernizzi, G. M. (2020). Electoral competition and factional sabotage. *Unpublished Manuscript*.

Ishida, J. (2012). Dynamically sabotage-proof tournaments. *Journal of Labor Economics 30*(3), 627–655.

Jain, S. and K. Qian (2021). Compensating Online Content Producers: A Theoretical Analysis. *Management Science 67*(11), 7075–7090.

Jiang, Y., J. Kang, D. Niyato, X. Ge, Z. Xiong, C. Miao, Xuemin, and Shen (2022). Reliable distributed computing for metaverse: A hierarchical game-theoretic approach.

John, A. (2020). When commitment fails – evidence from a field experiment. *Management Science 66*(2), 503–529.

Jones, C. I. (2002). Sources of U.S. Economic Growth in a World of Ideas. *American economic review 92*(1), 220–239.

Jones, C. I. (2022). Taxing Top Incomes in a World of Ideas. *Journal of Political Economy 130*(9), 2227–2274.

Jullien, B. and A. Pavan (2018, 07). Information Management and Pricing in Platform Markets. *The Review of Economic Studies 86*(4), 1666–1703.

Kaspi, H. and A. Mandelbaum (1998). Multi-Armed Bandits in Discrete and Continuous Time. *The Annals of Applied Probability 8*(4), 1270–1290.

Keller, R., S. Rady, and M. Cripps (2005). Strategic Experimentation with Exponential Bandits. *Econometrica 73*(1), 39–68.

Lambert, N. and Y. Shoham (2008). Truthful surveys. In C. Papadimitriou and S. Zhang (Eds.), *Internet and Network Economics,* Berlin, Heidelberg, pp. 154–165. Springer Berlin Heidelberg.

Liu, C., S. Wang, L. Ma, X. Cheng, R. Bie, and J. Yu (2017). Mechanism design games for thwarting malicious behavior in crowdsourcing applications. In *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications,* pp. 1–9.

Manso, G. (2011). Motivating Innovation. *The Journal of Finance 66*(5), 1823–1860.

O'Donoghue, T. and M. Rabin (1999). Doing it now or later. *American Economic Review 89*(1), 103–124.

O'Donoghue, T. and M. Rabin (2001). Choice and procrastination. *The Quarterly Journal of Economics 116*(1), 121–160.

Peitz, M., S. Rady, and P. Trepper (2017). Experimentation in two-sided markets. *Journal of the European Economic Association 15*(1), 128–172.

Rastegari, B., A. Condon, and K. Leyton-Brown (2007). Revenue monotonicity in combinatorial auctions. *ACM SIGecom Exchanges 7*(1), 45–47.

Sadler, E. (2021). Dead ends. *Journal of Economic Theory 191*, 105167.

Sadoff, S., A. Samek, and C. Sprenger (2020). Dynamic inconsistency in food choice: Experimental evidence from two food deserts. *Review of Economic Studies 18*(4), 1954–1988.

The Economist (2020, May). No one knows how many people live in north macedonia. https://www.economist.com/europe/2020/05/14/no-one-knows-how-many-people-live-in-north-macedonia.

Thomas, C. D. (2021, February). Strategic Experimentation with Congestion. *American Economic Journal: Microeconomics 13*(1), 1–82.

Toussaert, S. (2018). Eliciting temptation and self control through menu choices: A lab experiment. *Econometrica 86*(3), 859–889.

Weber, A. (1929). *Theory of the location of industries*. University of Chicago Press. (Translated by Carl Joachim Friedrich).

Yang, G., S. He, and Z. Shi (2017). Leveraging crowdsourcing for efficient malicious users detection in large-scale social networks. *IEEE Internet of Things Journal 4*(2), 330–339.

Yokoo, M. (2003). Characterization of strategy/false-name proof combinatorial auction protocols: Price-oriented, rationing-free protocol. In *IJCAI*, pp. 733–742.

Yokoo, M. (2008). *False-Name-Proof Auction*, pp. 308–310. Boston, MA: Springer US.

Yokoo, M., T. Matsutani, and A. Iwasaki (2006). False-name-proof combinatorial auction protocol: Groves mechanism with submodular approximation. In *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, pp. 1135–1142.

Yokoo, M., Y. Sakurai, and S. Matsubara (2001). Robust combinatorial auction protocol against false-name bids. *Artificial Intelligence 130*(2), 167–181.

Yokoo, M., Y. Sakurai, and S. Matsubara (2004). The effect of false-name bids in combinatorial auctions: new fraud in internet auctions. *Games and Economic Behavior 46*(1), 174–188.