Persuade or Price?
Bayesian Persuasion vs. Pigouvian Taxation Under Congestion[*][†]

Wen Lu
Georgia State University

Job Market Paper

December 5, 2025

Please click here for the latest version.

Shared resources suffer from overuse when users ignore costs imposed on others. I experimentally compare taxes versus strategic information provision for managing congestion. Groups of five choose between a shared resource with congestion and an uncertain alternative. Without intervention, two of five participants use the shared resource, matching theoretical predictions. A tax reduces usage to 1.4 participants, improving welfare. Information provision achieves 75 percent compliance with tailored recommendations yet increases congestion to 2.5 participants and reduces welfare. This failure occurs because partial information creates strategic gambling: participants told to avoid the resource sometimes enter, betting others' compliance leaves it underutilized. Results show information policies can backfire when individual sophistication undermines coordination, while simple taxes prove more reliable. This provides first experimental evidence comparing price and information mechanisms for managing shared resources.

Keywords: Common-pool resources, Bayesian persuasion, Pigouvian taxation, information design, experimental economics
JEL Codes: C92, D82, H23, Q20

---

[*] Previous Title: Incentive or Information? Experiments on Interventions to Reduce Negative Externalities

[†] I gratefully acknowledge the financial support provided by the AYS Dissertation Fellowship and the Experimental Economics Center (ExCEN). I thank Garth Heutel, Sarah Jacobson, Vjollca Sadiraj, James Cox, and participants at the Economic Science Association North American Meeting 2025, the Southern Economic Association Annual Meeting 2025, 2024, Western Economic Association International 2025, and seminars at Georgia State University. Contact: wenlu0222@gmail.com.

Contents

1 Introduction

Environmental and natural resource problems, particularly congestion problems, are often plagued by negative externalities. These problems arise when decision-makers fail to account for the social costs their actions impose on others, leading to overuse and inefficiency. Examples include road traffic congestion (such as New York City's upcoming congestion pricing plan), fisheries depletion, shared computing infrastructure overload, and attention-based digital platform crowding. Economists typically model these as common-pool resource (CPR) problems, where resources are rival but non-excludable. Left unregulated, these systems tend toward overutilization, creating a gap between individual incentives and collective welfare.

Two broad classes of policy interventions are available to address such problems. Market-based policies such as Pigouvian taxes directly alter incentives by making socially costly actions more expensive. Information-based policies such as strategic disclosure influence behavior indirectly by shaping beliefs about the state of the world. While the theoretical properties of these tools are well understood in isolation, little is known about their comparative effectiveness in strategic environments where multiple agents face both risk (about the state of nature) and strategic uncertainty (about others' actions).

This paper compares taxation with Bayesian persuasion (Kamenica and Gentzkow, 2011), an information design tool that influences behavior by providing strategically designed signals rather than altering payoffs. In this approach, a planner who initially does not know the state learns it and then sends tailored recommendations to different users, designed to coordinate them toward efficient outcomes. By structuring signals strategically, the planner can influence the equilibrium distribution of behavior. This method is especially relevant when state uncertainty plays a central role in decision-making.

This tradeoff is not merely theoretical. New York City approved a congestion pricing plan set to take effect in 2025, charging drivers entering Manhattan's central business district during peak hours. While this Pigouvian approach is grounded in economic theory, it has sparked strong political opposition and equity concerns. A contrasting example comes from digital platforms like Google Maps, which influence behavior through information rather than price. By selectively displaying route options, these platforms shape traffic patterns without direct monetary costs. However, Google Maps sends the same information to all users, unlike the tailored signals in Bayesian persuasion that could potentially achieve more efficient coordination.

I use a laboratory experiment to cleanly compare these approaches. I study a simple environment where groups of five participants repeatedly choose between two options: a risky outside option whose value depends on an uncertain state of nature, and a shared resource whose value declines with congestion. Participants do not observe the state before choosing and face uncertainty about others' actions, creating scope for both price and information to improve coordination. I implement two between-subjects treatments in a within-subjects design. In the tax treatment (BT), I impose a fixed cost on choosing the shared resource, following the classic Pigouvian prescription. In the information treatment (BI), I use Bayesian persuasion: a central planner observes the state and sends private, non-binding recommendations to participants that are designed to coordinate them toward the social optimum.

The experiment yields three main findings.

First, in the absence of any intervention, participants coordinate on an ex-ante rather than a state-contingent equilibrium. In the baseline rounds, the average number of entrants into the shared resource is essentially the same in the Good and Bad states. Group-level entry is about 1.87 out of 5 participants in both cases, close to the ex-ante Nash benchmark of two entrants when states are equally likely. Participants largely ignore state differences in their decisions. This state-invariant behavior suggests that agents use coarse, ex-ante decision rules when facing both risk and strategic uncertainty.

Second, both interventions successfully reduce inefficiency in the state where the outside option strictly dominates. In the Good state, where the outside option is clearly superior, entry into the shared resource falls from approximately 38 percent in the baseline to about 26 percent in Stage 2. This indicates that both price and information help participants recognize and act on state information they previously ignored.

Third, the interventions diverge dramatically in the state where coordination is critical. In the Bad state (where the social optimum is 2 out of 5 entrants), the information treatment increases entry from 35 percent to 50 percent, while the tax treatment decreases entry from 40 percent to 28 percent. In terms of group-level entry, the information treatment moves from about 1.73 to 2.49 entrants, overshooting both the baseline and the social optimum of two - moving opposite to theoretical predictions. The tax treatment moves from exactly two entrants to about 1.38 in Stage 2, which moves in the theoretically predicted direction. Relative to the ex-ante equilibrium benchmark of two entrants and the ex-ante optimal level of one entrant under the tax, this shift represents about 62 percent of the desired movement.

I find evidence that the information treatment likely backfires due to strategic gambling on partial information. The optimal information structure in this setting reveals the state perfectly to a subset of participants while leaving others with residual uncertainty. Those who learn with certainty that the state is Bad rationally choose the shared resource. Those who receive partial information face a tempting gamble: if the state turns out to be Bad and if most others follow recommendations, then the shared resource can deliver high payoffs when congestion is low. Approximately 27 percent of those with partial information deviate by entering the shared resource, and this strategic gambling pushes aggregate entry to about 2.49 participants, above both the baseline and the social optimum.

The tax treatment, by contrast, reliably reduces entry across both states in a way that lines up with ex-ante reasoning. In the Bad state, the state-contingent social optimum calls for two entrants, but once we recognize that participants behave ex-ante rather than state by state, the relevant target under the tax becomes the ex-ante optimal level of one entrant. The tax reduces group entry from two to about 1.38, which is incomplete but substantial progress toward this target. This reveals a trade-off between behavioral robustness and precision. Taxes change behavior uniformly and predictably, even when agents rely on coarse ex-ante decision rules, while information is theoretically capable of implementing optimal allocations but is behaviorally fragile when partial revelation creates opportunities for strategic gambling.

This paper makes four contributions.

First, it provides the first experimental comparison of Pigouvian taxation and Bayesian persuasion in a common-pool resource setting with multiple strategic agents and state-dependent payoffs. While taxation has been studied extensively, experimental work on Bayesian persuasion remains limited, particularly in multi-receiver settings with externalities.

Second, the experiment documents a novel baseline regularity: participants coordinate on ex-ante equilibria as if the state were unknown, rather than responding to state-specific information. Group-level entry is essentially identical across states and centers around the ex-ante Nash prediction. This pattern suggests that models incorporating coarse reasoning or robust decision-making may better capture behavior in congested environments than models that assume fully state-contingent best responses.

Third, I identify a new failure mechanism for information-based interventions. High individual compliance does not guarantee correct aggregate outcomes when signal structures create heterogeneous beliefs and opportunities for strategic gambling. In my information treatment, a 75 percent compliance rate coincides with aggregate behavior that moves opposite to theoretical predictions in the Bad state. The deviators are not confused; they exploit the informational asymmetries created by optimal persuasion in a way that amplifies congestion. This highlights that even theoretically optimal information design can fail when strategic responses undermine coordination.

Fourth, the experimental design contributes substantively by implementing a multi-receiver Bayesian persuasion environment where receivers' actions create externalities for each other. This extends the experimental literature on information design, which has focused primarily on single-receiver or small-group settings without externalities (Nguyen, 2017; Fréchette et al., 2022; Aristidou et al., 2023). The multi-agent CPR structure introduces new behavioral complexities that are policy-relevant but experimentally underexplored, and it allows a direct comparison between information and price instruments in the same strategic environment.

The remainder of the paper proceeds as follows. Section 2 reviews related literature. Section 3 presents the theoretical framework and derives predictions. Section 4 describes the experimental design and procedures. Section 5 presents the empirical results. **??** discusses mechanisms and policy implications. Section 6 concludes.

2 Related Literature

This paper connects four strands of literature: experimental studies of common-pool resource management, the emerging literature on Bayesian persuasion and information design, the longstanding literature on Pigouvian taxation, and comparisons of informational versus incentive-based policy interventions.

2.1 Common-Pool Resources and Congestion Games

The tragedy of the commons, formally analyzed by Hardin (1968) and modeled by Ostrom (1990), predicts that rational individuals will overuse shared resources in the absence of property rights or regulation. Laboratory experiments have extensively documented this pattern. Walker et al. (1990) show that participants in common-pool resource games extract more than the social optimum but less than the Nash equilibrium prediction, with behavior converging toward Nash over time. Subsequent work has explored various institutional solutions including communication (Ostrom et al., 1992), graduated sanctions (**?**), and voting on regulations (Walker et al., 2000).

Most relevant to this paper is the experimental literature comparing policy interventions in CPR settings. Delaney and Jacobson (2016) compare monetary incentives and information provision, finding that both reduce overuse but that the combination is most effective. Critically, their information treatment provides feedback about past outcomes rather than strategic information about the current state, and they do not explore settings with environmental uncertainty about the value of non-extraction. My contribution is to compare incentives with strategically designed information revelation (Bayesian persuasion) in a setting where state uncertainty about the outside option creates scope for information to coordinate behavior across states.

A parallel literature studies congestion games, which share the structure of CPR problems. Rapoport et al. (2009) document systematic deviations from equilibrium in coordination games with negative externalities. Anderson et al. (2008) study a laboratory environment where subjects choose between a safe outside option and a risky, congestible route, finding reasonable alignment with equilibrium predictions but substantial variance that reduces welfare. However, their setup does not include exogenous state uncertainty about payoffs.

Recent work has begun introducing uncertainty into congestion experiments, though the sources and roles of uncertainty vary. Qiao and Noussair (2025) examine how traffic apps influence congestion when there is uncertainty about travel times, varying the share of drivers with access to real-time information. Lu et al. (2011) and Mak et al. (2015) explore how feedback about unchosen routes affects learning. However, these studies focus primarily on ex-post information provided after decisions to facilitate learning across rounds. By contrast, my design emphasizes ex-ante belief formation, where signals are delivered before choices and strategically designed to influence behavior through Bayesian persuasion.

One dimension of uncertainty that remains underexplored is uncertainty about the return of the outside option or non-congested route, a common feature of real-world decisions. Should I take the highway that might be under construction, or join the congested local road? This added uncertainty makes the game more realistic and opens questions about how information affects decision-making when strategic interaction depends not only on what others will do, but also on what agents believe about the environment.

This paper builds on the theoretical model of Das et al. (2017), who analyze a congestion game where agents choose between a congestible route and an outside option whose cost depends on an exogenous state. They show how a planner can use signals to guide agents toward efficient allocations. While their work is fully theoretical and assumes a continuum of agents, it highlights how uncertainty about the outside option can be leveraged through strategic information design. My study adapts their framework to a finite-agent experimental environment and investigates how individuals respond to belief-based interventions under strategic interaction.

2.2 Bayesian Persuasion and Information Design

The theoretical framework for information design was established by Kamenica and Gentzkow (2011), who characterize optimal information structures when a sender with commitment power can design signals to influence a receiver's posterior beliefs. The key insight is that by committing to a signal structure before observing the state, a sender can implement outcomes impossible with simple cheap talk or full revelation. Bergemann and Morris (2016, 2019) extend this to study information design in games with multiple receivers, introducing the concept of Bayes correlated equilibrium (BCE). While persuasion with a single receiver is fundamentally about optimization, the multi-receiver case becomes one of equilibrium selection: receivers engage in a Bayesian game, and persuasion becomes selecting an optimal BCE from the feasible set.

Experimental evidence on Bayesian persuasion is growing but remains limited to simple settings. Nguyen (2017) provides the first laboratory test using an intuitive interface with envelopes and cards, finding that subjects' behavior as both senders and receivers is qualitatively consistent with theory but deviates quantitatively, with senders providing too little information and receivers over-responding. Fréchette et al. (2022) use a slider interface where senders choose the probability of sending truthful signals, confirming that both senders and receivers exhibit systematic biases but behavior converges toward theory with experience. Au et al. (2023) allow senders to directly select posteriors rather than information structures, assisting belief updating. Aristidou et al. (2023) compare Bayesian persuasion with mechanism design, finding both can influence behavior but mechanisms are more robust to bounded rationality.

Importantly, all existing experimental studies of Bayesian persuasion examine settings with either one sender and one receiver, or one sender and a small number of receivers without strategic interaction among receivers. The only exception is Ziegler (2023), who explores persuasion with two strategic receivers and examines whether public or private messages are more effective depending on whether receivers' actions are strategic complements or substitutes. However, his receivers do not face congestion externalities, and his sender is self-interested rather than a benevolent planner implementing a welfare-maximizing signal structure. His focus is on the communication channel (public vs. private), while mine is on whether belief manipulation can substitute for monetary incentives.

My contribution to this literature is implementing Bayesian persuasion in a multi-agent strategic environment with congestion externalities. This introduces a new complexity: even if each individual correctly interprets their signal, strategic interaction can cause aggregate outcomes to diverge from the planner's intention. The theoretical literature on persuasion with multiple receivers and externalities (**??**) has identified conditions under which optimal persuasion can coordinate behavior, but experimental evidence has been lacking. My results show these conditions may be violated in practice when signal structures create opportunities for strategic gambling.

Implementation approach. Since this paper employs a one-sender–multiple-receivers structure, and unlike prior studies where the sender strategically chooses the signal structure to maximize personal payoffs, it features a benevolent planner who implements a fixed information policy derived from the theoretical social optimum. Neither of the standard experimental approaches, a physical signal construction (Nguyen, 2017) or slider-based probability selection (Fréchette et al., 2022) is utilized. The sender is not an actual player receiving monetary rewards based on receivers' choices. Their role is solely to implement an ex ante optimal signal structure. This design simplifies the experimental interface while preserving the logic of Bayesian persuasion in a centralized policy environment, and the posterior is implicitly embedded in the signal pattern, where subjects need only recognize that receiving recommendation A means "more than half likely" it's the better choice, enabling coordination without full Bayesian updating.

2.3 Pigouvian Taxation and Market-Based Policies

The idea that externalities can be corrected by taxes or subsidies aligning private and social costs dates to Pigou (1920). Laboratory experiments have generally confirmed that such taxes reduce externalities. Plott (1983) shows taxes can move behavior toward competitive equilibrium in induced-value markets. Anderson et al. (2008) find that congestion pricing increases welfare by discouraging excessive entry. Cason and Gangadharan (2016) compare taxes and subsidies in public goods contexts, finding both effective but with different distributional consequences.

However, these experiments typically implement taxes in settings where the optimal rate is straightforward to calculate and apply uniformly across agents and states. In my setting, the efficient allocation is state-contingent: in one state everyone should avoid the shared resource (optimal entry X=0), while in another state exactly two out of five should enter (optimal entry X=2). A uniform tax cannot achieve both objectives simultaneously—it can only discourage entry on average. This highlights a fundamental limitation of Pigouvian taxation when externalities are state-dependent: market-based policies trade precision for robustness.

Furthermore, implementing taxes in practice raises familiar challenges. Users often perceive them as unfair or coercive, and political resistance is common. Kallbekken et al. (2011)

show that even when a tax is economically efficient, people may reject it due to tax aversion, especially when they don't understand its purpose or revenues aren't transparently recycled. Andreassen et al. (2024) find that combining taxes with subsidies can increase public support by addressing fairness concerns.

More fundamentally, in CPR settings where resources like water, fisheries, or air quality are shared among many users, implementing price-based solutions is difficult because monitoring usage and enforcing compliance are often impractical. Additionally, determining the correct tax level requires accurate externality measurement, which may be infeasible in decentralized environments.

Field evidence comes primarily from congestion pricing. Studies of London, Stockholm, and Singapore generally find that congestion pricing reduces traffic volumes and travel times, though effectiveness varies with implementation details and public acceptance remains challenging. Notably, these schemes use time-varying prices rather than state-contingent information, consistent with the insight that adjusting prices is administratively simpler than designing information revelation mechanisms.

2.4 Comparison of Policy Instruments

A small literature directly compares different regulatory instruments. Cason and Gangadharan (2016) compare taxes, subsidies, and cap-and-trade in pollution control, finding all reduce emissions but with different efficiency and distributional properties. Ferraro and Price (2013) review field experiments comparing economic incentives and behavioral interventions in environmental conservation, concluding that context matters greatly. Allcott and Rogers (2014) show that information about social norms can persistently reduce energy consumption, suggesting behavioral interventions can sometimes compete with price-based policies.

This paper is most closely related to two recent experimental studies. Aristidou et al. (2023) compare information design with direct mechanisms in a laboratory setting but focus on a two-receiver environment without externalities or congestion, and their main question concerns relative robustness to bounded rationality rather than comparative welfare effects in CPR problems. Delaney and Jacobson (2016) compare taxes with information nudges and social norm messaging in a CPR extraction experiment but use simple feedback about past outcomes rather than strategic information design, and their setting lacks state uncertainty about the outside option.

My contribution is providing a clean experimental comparison of Pigouvian taxation and Bayesian persuasion, two policy tools grounded in distinct theoretical frameworks, within the same strategic environment featuring both state uncertainty and congestion externalities. The laboratory's key advantage is that I can implement theoretically optimal versions of both policies and isolate the behavioral factors causing them to succeed or fail. By embedding environmental uncertainty and multi-agent strategic behavior into the same experimental structure, I shed light on when and how each tool works, and what tradeoffs they entail for effective CPR management.

3 Theoretical Framework

This section develops the theoretical foundation for the experiment. Section 5.1 introduces a baseline binary-choice congestion game with an uncertain outside option, adapted from Das et al. (2017). Section 3.2 characterizes the socially optimal allocation, providing a benchmark for evaluating interventions. Section 3.3 formalizes the information treatment as a Bayesian persuasion problem where the planner commits to a fixed signal structure to induce optimal entry. Section 3.4 derives the incentive-compatible tax aligning private and social incentives. Section 3.5 provides the experimental calibration. The resulting conditions guide experimental design and generate testable predictions.

3.1 The Baseline Model

Similar to Das et al. (2017), consider a group of agents (drivers) choosing between two roads, a and b. Road a is affected by an exogenous state—either under construction (Bad State) or not (Good State)—introducing uncertainty into its payoff. Road b is subject to congestion, where each additional entrant reduces the return for everyone. This setup resembles a standard CPR game as individual decisions impose negative externalities. However, it differs in two key ways: agents make a binary choice between two discrete options, and the payoff from the outside option (Road a) depends on an uncertain state.

Suppose there are $n+1$ risk-neutral players total, consisting of one Sender (S) and n Receivers (R). S acts as a benevolent social planner aiming to induce an efficient outcome, while each R is a self-interested agent maximizing individual payoff. All Receivers move simultaneously. This section focuses on analyzing Receivers' payoffs and equilibrium behavior, while the Sender's incentives and socially optimal outcome are discussed subsequently.

Each R chooses action $x_i \in \{0,1\}$, where $x_i = 1$ means entering road b and $x_i = 0$ means entering road a. The payoff function is:

$$\pi_i = A - \omega(1 - x_i) - \gamma X x_i, \quad i \in \{1, 2, 3, \ldots, n\} \tag{1}$$

where:

- Each player's decision is binary, endowment normalized to 1
- Reaching destination yields value A
- ω is the cost of choosing road a ($\omega = 0$ if Good state, $\omega = 1$ if Bad state)
- State $S \in \{G, B\}$ drawn by Nature with $\Pr(G) = p$, $\Pr(B) = 1 - p$ (common knowledge)
- γ: individual congestion cost from choosing road b

- $X = \sum_{j=1}^{n} x_j$: total number choosing road b

Under Full Information

Suppose all agents are self-interested own-payoff maximizers with knowledge of the state. The payoffs for R_i, given all other agents' actions fixed, are shown in Table 1.

Table 1: Payoff Matrix with State Probabilities

Agent's choice	$\omega = 0$ (Pr $= p$)	$\omega = 1$ (Pr $= 1 - p$)
$x_i = 0$	A	$A - 1$
$x_i = 1$	$A - \gamma X$	$A - \gamma X$

When $\omega = 0$ (Good state), choosing $x_i = 0$ is a dominant strategy since the cost of road a is zero. All agents choose a, leading to $X = 0$.

When $\omega = 1$ (Bad state), agents compare costs of a and b. Cost of a is fixed at 1, while cost of b is congestion-dependent at γX. Agents continue choosing b until costs equalize:

$$1 = \gamma X \implies X^{NE} = \frac{1}{\gamma}, \text{ when } \omega = 1 \tag{2}$$

This determines the equilibrium number choosing road b under full information when $\omega = 1$. When $\omega = 0$, equilibrium is $X = 0$ as choosing $x_i = 0$ is dominant.

Under Uncertainty

When the state is unknown, equilibrium requires identifying the marginal R indifferent between $x_i = 0$ and $x_i = 1$. The marginal R represents the threshold where entering b's cost exactly offsets the expected benefit.

The marginal R's payoff when choosing $x_i = 1$ is $A - \gamma X$ regardless of state. When choosing $x_i = 0$, she faces environmental uncertainty: if $\omega = 0$ (probability p), payoff is A; if $\omega = 1$ (probability $1 - p$), payoff is $A - 1$.

The equilibrium condition from the marginal R's indifference:

$$A - \gamma X = pA + (1-p)(A-1) \tag{3}$$

Solving for X:

$$X^{NE} = \frac{1-p}{\gamma} \tag{4}$$

This determines equilibrium entry when the state is unknown. The number decreases as probability of $\omega = 0$ increases since agents anticipate lower expected cost when avoiding congestion. The parameter γ moderates entry: as congestion costs rise, fewer agents choose b.

3.2 Social Optimum

The Sender (S) acts as a benevolent social planner aiming to maximize joint payoff of all Receivers. The socially optimal outcome maximizes total welfare accounting for externalities each entry decision imposes.

Rewrite R's utility: $\pi_i = A - \omega + (\omega - \gamma X)x_i$

The social planner determines optimal number choosing road b (variable X) to maximize aggregate utility:

$$\max_X (n - X)(A - \omega) + X(A - \gamma X) \tag{5}$$

Under full information where S knows the state, decision-making follows straightforward structure. When $\omega = 0$, choosing a incurs no cost, meaning S and R incentives align. Optimal decision: all R choose a, implying $x_i = 0$ for all i and $X = 0$.

When $\omega = 1$, S maximizes:

$$\max_X (n - X)(A - 1) + X(A - \gamma X) \tag{6}$$

Taking derivative:

$$\frac{d}{dX}[(n-X)(A-1)+X(A-\gamma X)] = -(A-1)+(A-\gamma X)-\gamma X = 1-2\gamma X = 0 \tag{7}$$

The socially optimal number yields:

$$X^{SO} = \frac{1}{2\gamma}, \text{ when } \omega = 1 \tag{8}$$

Since when $\omega = 0$, optimal entry is $X^{SO} = 0$, the expected socially optimal entry is:

$$E[X^{SO}] = p \times 0 + (1-p)\frac{1}{2\gamma} = \frac{1-p}{2\gamma} \tag{9}$$

The socially optimal entry differs from self-interested equilibrium $\frac{1-p}{\gamma}$. In individual equilibrium, each agent decides whether to enter based solely on own payoff, without accounting for the congestion externality imposed on others. More agents enter than is socially optimal since each only considers private cost and doesn't internalize additional congestion created. The social planner accounts for this externality and limits entry to improve welfare. This misalignment between private and social incentives underscores the need for intervention—information design or incentive mechanisms—to reduce inefficiencies and bring equilibrium closer to social optimum.

3.3 Information Treatment

In the Information Treatment, S can send signals to R before R decides, potentially influencing choices. Before making an entry decision, each R observes private signal $s_i \in \{\alpha, \beta\}$ providing information about state $S \in \{G, B\}$.

This study interprets signals as action recommendations: α represents S recommending R choose road a ($x_i = 0$), and β represents S recommending R choose road b ($x_i = 1$). This follows from the revelation principle ensuring that without loss of generality, signals can be relabeled as direct action recommendations without affecting equilibrium (Kamenica and Gentzkow, 2011). The sender's strategy consists of choosing a probability distribution over signals under each state. The set of feasible signal probabilities available to S:

$$P_S = \{(p_\alpha, p_\beta) | p_\alpha, p_\beta \in [0,1]\} \tag{10}$$

where $p_\alpha = \Pr(\alpha|G)$ (probability of "recommend road a" given Good state) and $p_\beta = \Pr(\beta|B)$ (probability of "recommend road b" given Bad state).

Optimal Choice: Partial Private Signals

Because public partial information fails to shift behavior efficiently, the optimal approach sends private partial signals. Unlike public signals, private signals allow different agents to receive different information, leading to heterogeneous beliefs. Since each R updates posterior belief independently based on received signal, private signals create strategic uncertainty among receivers. This belief variation can influence equilibrium entrants X more effectively, steering toward social optimum.

Specifically, the goal is inducing posterior belief p' such that each individual maximizes own utility while ensuring the number of R entering b aligns with socially optimal entry level.

Implementation Strategy

To implement socially optimal outcome, S uses the following strategy: when state is Good (no construction, $\omega = 0$), since S and all R prefer road a, signal is fully revealing. Sender always recommends road a, i.e., $p_\alpha = 1$. When state is Bad (under construction, $\omega = 1$), road a incurs cost while overuse of road b leads to congestion. Socially optimal number choosing road b is $X^{SO} = \frac{1}{2\gamma}$. Given n receivers, this corresponds to fraction $\frac{1}{2\gamma n}$ of population.

To implement this allocation through signals, S directly selects a subset of R to receive each recommendation. Specifically, when state is Bad, Sender assigns exactly $\frac{1}{2\gamma n}$ of receivers message "recommend road b," and remaining $1 - \frac{1}{2\gamma n}$ receive "recommend road a." This ensures precisely the socially optimal number are encouraged to enter congestible route when state is unfavorable.

Posterior Beliefs

Following Fréchette et al. (2022), optimal strategy for S is fully revealing state when $\omega = 0$, leading to equilibrium result $p_\alpha = 1$. This ensures no agent in state G receives β, meaning $\Pr(\beta|G) = 0$.

Since all agents receiving β must be in state B:

$$\Pr(B|\beta) = 1, \quad \Pr(G|\beta) = 0 \tag{11}$$

Given $p_\beta = \frac{1}{1+p}$ and $p_\alpha = 1$, posteriors are:

$$\Pr(G|\alpha) = \frac{p_\alpha p}{p_\alpha p + (1 - p_\beta)(1 - p)} = \frac{1+p}{2} \tag{12}$$

$$\Pr(B|\alpha) = \frac{(1 - p_\beta)(1 - p)}{p_\alpha p + (1 - p_\beta)(1 - p)} = \frac{1-p}{2} \tag{13}$$

3.4 Incentive Treatment: Pigouvian Tax

In this treatment, S can impose Pigouvian tax to disincentivize R from choosing congested road if state is Bad. To introduce tax, fixed amount τ is subtracted from payoff of choosing road b:

$$\pi_i = A - (\omega)(1 - x_i) - (\gamma + \tau)Xx_i \tag{14}$$

Marginal R's equilibrium condition:

$$A - (\gamma + \tau)X = pA + (1 - p)(A - 1) \tag{15}$$

Solving for X:

$$X = \frac{1 - p}{\gamma + \tau} \tag{16}$$

Optimal Pigouvian tax τ aligning private incentive with social optimum is found by setting these equal:

$$\frac{1 - p}{\gamma + \tau} = \frac{1 - p}{2\gamma} \tag{17}$$

Solving: $\tau^* = \gamma$

Key limitation: This tax achieves *expected* optimum $E[X^{SO}]$ but cannot implement state-contingent optima. The tax is **blunt but robust**: reliably reduces entry on average but over-corrects in Bad state (where some entry optimal) and under-corrects in Good state (where zero entry optimal).

3.5 Experimental Calibration

In the experimental environment, each participant chooses between two options: Box A or Box B. Box A yields a fixed payoff that depends on the realized state of the world. In the Good state, the payoff from Box A is 30. In the Bad state, the payoff from Box A is 10. Box B is a shared, congestion-sensitive resource with a payoff that decreases linearly in the number of entrants. Specifically, if X participants choose Box B, the payoff is $30 - 5X$.

Groups consist of five participants. The realized state is displayed on the screen prior to each choice, but payoffs depend not only on the state but also on the number of participants entering Box B.

Information Treatment Calibration

State = 30 (Good). When the state is Good, Box A yields a certain payoff of 30. Box B yields at most 25 when $X = 1$ and declines further with additional entrants. Since Box A strictly dominates Box B for any $X \geq 1$, both the Nash equilibrium and the social optimum are:

$$X^{SO} = X^{NE} = 0$$

corresponding to an allocation of 5 choosing Box A and 0 choosing Box B.

State = 10 (Bad). When the state is Bad, Box A yields only 10. Box B yields $30 - 5X$. A social planner maximizes total group payoff:

$$W(X) = (5 - X) \cdot 10 + X \cdot (30 - 5X)$$

The socially optimal level of entry into Box B is:

$$X^{SO} = 2$$

corresponding to an allocation of 3 choosing Box A and 2 choosing Box B.

In contrast, the symmetric, state-contingent Nash equilibrium in the Bad state equates the individual payoffs:

$$10 = 30 - 5X \Rightarrow X^{NE} = 4$$

which corresponds to 1 choosing Box A and 4 choosing Box B.

Signal Structure in the Information Treatment

To implement the social optimum in the Bad state, the planner sends private, non-binding recommendations according to the following structure:

- If the state is Good (30): all five participants receive “Choose A”

- If the state is Bad (10): two participants receive "Choose B" and three participants receive "Choose A"

This creates two types of information in the Bad state:

- Two participants are fully informed that the state is Bad, since "Choose B" is never sent in the Good state.
- Three participants receive a partially informative signal, since "Choose A" is sent in both states.

With a prior of $P(\text{Good}) = 0.5$, Bayes' rule implies:

$$P(\text{Good} \mid \text{"Choose A"}) = \frac{0.5}{0.5 + 0.5 \cdot \frac{3}{5}} = \frac{0.5}{0.8} = 62.5\% \tag{18}$$

$$P(\text{Good} \mid \text{"Choose B"}) = 0 \tag{19}$$

Thus, two participants receive full information that the state is Bad, while three participants update to an interior posterior that leaves strategic uncertainty.

Incentive Treatment Calibration

In the tax treatment, a constant Pigouvian tax of $\tau = 5$ is imposed on choosing Box B. The new payoff from Box B becomes:

$$\pi_B = 30 - 5X - 5 = 25 - 5X$$

The ex ante equilibrium condition becomes:

$$\mathbb{E}[\pi_A] = \mathbb{E}[\pi_B]$$

$$20 = 25 - 5X \Rightarrow X^{NE} = 1$$

Thus, the tax treatment predicts one entrant in Box B, corresponding to a 20 percent individual entry rate in both states.

Unlike the information treatment, the tax does not generate state-contingent allocations. Instead, it implements a single ex ante prediction that trades off efficiency in the Good and Bad states.

3.6 Theoretical Predictions

Baseline (Stage 1). Participants observe the realized state and choose between Box A and Box B without intervention. The state-contingent Nash equilibria are:

- Good state: $X^{NE} = 0$
- Bad state: $X^{NE} = 4$

However, under symmetric mixed strategies and with $P(\text{Good}) = 0.5$, the ex ante equilibrium satisfies:

$$X^{NE} = \frac{1-p}{\gamma} = 2$$

which corresponds to an individual entry rate of 40%.

Information Treatment (Stage 2, BI). If all participants follow recommendations:

- Good state: $X = 0$
- Bad state: $X = 2$

This treatment is calibrated to achieve the state-contingent social optimum.

Tax Treatment (Stage 2, BT). With tax $\tau = 5$:

- Predicted entry: $X = 1$ in both states
- Corresponding to a 20% individual entry rate

Comparative Predictions. Relative to Stage 1 baseline outcomes:

1. In the Good state, both treatments should reduce entry into Box B, moving behavior toward $X = 0$.
2. In the Bad state, the information treatment should maintain entry at $X = 2$, while the tax treatment should reduce entry toward $X = 1$.
3. In expectation, both treatments should weakly increase efficiency relative to baseline. Under full compliance and correct belief updating, the information treatment should weakly dominate the tax treatment by implementing the state-contingent optimum.

3.7 Hypotheses

The following hypotheses are derived from the theoretical structure of the game and the calibration of the two interventions, prior to observing any data.

H1 Baseline equilibrium behavior. In Stage 1, group-level entry into Box B will be consistent with equilibrium behavior implied by the payoff structure. In expectation, entry will be concentrated around the ex ante benchmark of $X = 2$ out of 5 participants.

H2 Good-state response to intervention. In the Good state, where Box A strictly dominates Box B for any congestion level, both the information treatment and the tax treatment will reduce entry into Box B relative to the baseline stage.

H3 Bad-state response to intervention. In the Bad state, where coordination on intermediate entry is required for efficiency, both the information treatment and the tax treatment will move group-level entry closer to the efficient target of $X = 2$ compared to the baseline stage.

H4 Welfare improvement. Relative to baseline, both interventions are expected to weakly increase group-level welfare by reducing inefficient congestion in the shared resource.

4 Experimental Design

4.1 Framing

The experiment was designed to investigate how incentive-based (Pigouvian tax) and information-based (Bayesian persuasion) interventions influence individual behavior in a stylized CPR environment with congestion. To avoid framing effects commonly associated with real-world scenarios such as driving or traffic, the experiment adopts a neutral, abstract representation using balls and boxes. In this setup, participants choose between two abstract options (Box A and Box B), which correspond conceptually to a risky outside option and a congestion-sensitive shared resource, respectively.

This ball-and-box representation minimizes potential bias from real-world associations like fairness, responsibility, or prior experience, and has precedent in experimental economics as a tool for presenting probabilistic environments in a cognitively accessible but emotionally neutral way (Holt and Laury, 2002; Charness and Gneezy, 2010). It allows the experiment to isolate strategic behavior and belief formation while avoiding demand effects that might arise from more emotionally or politically charged language like “roads,” “traffic,” or “tolls.” The neutral framing was consistent across all treatments and was particularly important for isolating the effects of information and incentive design.

Each participant plays the role of a decision-maker who must repeatedly choose between two boxes. The color of the ball drawn from Box A is determined by nature and yields a payoff that depends on the state of the world. The payoff from choosing Box B, by contrast, depends on the number of other participants who also choose Box B, reflecting a congestion externality. The experimental treatments vary how much information or incentive participants receive about the state of the world, with the goal of testing whether persuasion or pricing is more effective at aligning behavior with the social optimum.

Furthermore, although the theoretical model is formulated in terms of value minus cost, the experimental interface presented all outcomes as positive payoffs. This choice was made for clarity and ease of understanding, and to minimize potential framing effects such as loss aversion or punitive interpretations (Kahneman and Tversky, 1979). The payoff structure used in the lab is mathematically equivalent to the theoretical model and preserves its equilibrium logic, while offering a cleaner and more intuitive experience for participants.

4.2 Baseline Game Structure

Participants were randomly matched into groups of five in each round, and all decisions were made simultaneously within these groups. The random matching protocol was implemented to minimize reputation effects and ensure that each round represented an independent strategic interaction. Groups were re-matched after every round, preventing participants from developing cooperative norms or engaging in repeated-game strategies. This design choice aligns with the theoretical model's assumption of one-shot interactions and allows for cleaner identification of treatment effects.

Within each group, participants chose between two options:

Box A (Outside Option): Contains a ball with unknown color, either yellow or blue. The color is drawn by nature and determines the payoff. Yellow corresponds to the Good state and yields a high payoff of 30. Blue corresponds to the Bad state and yields a lower payoff of 10. The probability distribution over states (denoted by $p = 0.5$) is common knowledge and explicitly stated in the instructions, but the actual color of the ball is unknown to participants at the time of decision. The true color (i.e., the realized state) is revealed after each round, providing feedback that allows for learning across rounds.

Box B (Shared Option): Always provides a positive payoff that decreases with the number of participants choosing it. It starts from a base of 30 and declines linearly with congestion. Specifically, each participant who chooses Box B receives a payoff of $30 - 5X$, where X is the number of participants who selected Box B in that round. This payoff structure was clearly displayed in a table format on participants' screens, showing the payoff for each possible value of X from 1 to 5. This transparency ensures that participants fully understand the congestion externality and can make informed decisions.The payoff structure was clearly displayed:

Table 2: Box B Payoff Structure

Number Choosing B (X)	Payoff from B
1	25
2	20
3	15
4	10
5	5

This binary-choice, CPR-style game captures the tension between risk (uncertain state of Box A) and congestion (strategically uncertain but diminishing returns in Box B). It reflects real-world tradeoffs between exploring uncertain but potentially high-return options versus joining others in a predictable but crowded alternative. The parameters were calibrated to create meaningful tension: the Good state payoff (30) is high enough to dominate Box B under any congestion level, while the Bad state payoff (10) creates genuine strategic uncertainty about which option to choose.

4.3 Treatment Design

The experiment features two between-subject treatments, each consisting of two stages with 10 rounds per stage. This two-stage design allows us to observe baseline behavior before intervention and then measure the impact of policy tools within-subject, while maintaining between-subject treatment assignment to avoid contamination effects. The 10-round structure per stage provides sufficient observations for learning to stabilize while keeping the session length manageable.

Baseline–Tax (BT): Participants first play 10 rounds under the baseline condition. This allows us to establish a within-subject baseline for comparison. In the second stage (rounds 11-20), a Pigouvian tax is introduced with clear announcement and explanation.

Participants who choose Box B incur a fixed cost of 5 points, reducing its net payoff to $25 - 5X$. The tax is subtracted from the payoff and clearly displayed on the decision screen. Importantly, the tax is implemented as a simple, transparent deduction rather than being framed as a penalty or punishment, to minimize potential reactance effects.

The tax is calibrated so that, under rational expectations, the Nash equilibrium number of entrants into Box B in the Bad state matches the social optimum of $X = 2$. Without the tax, the equilibrium prediction is $X = 4$ (setting $10 = 30 - 5X$). With the tax, the equilibrium becomes $10 = 25-5X$, yielding $X = 3$. However, accounting for risk aversion and coordination uncertainty, the empirical equilibrium is expected to be closer to $X = 2$, which is the socially optimal level. This calibration allows us to test whether pricing mechanisms can effectively align private incentives with social welfare in a congested environment.

Baseline–Information (BI): Participants also begin with 10 rounds of the baseline condition, ensuring comparability with the BT treatment. In the second stage (rounds 11-20), a Bayesian persuasion intervention is introduced. At the beginning of each round in Stage 2, after the state is realized but before participants make their choice, a central planner sends signals to each participant based on the true state of nature.

The signal structure implements a deterministic persuasion mechanism designed to achieve the social optimum:

- **If the state is Good (Yellow ball, payoff = 30):** All five participants receive a recommendation to "Choose Box A." This recommendation is truthful and aligns individual and social incentives, as Box A strictly dominates Box B in this state.
- **If the state is Bad (Blue ball, payoff = 10):** Exactly 2 participants (chosen randomly by the computer) receive a recommendation to "Choose Box B," while the remaining 3 participants receive a recommendation to "Choose Box A." This allocation corresponds to the social optimum in the Bad state, where $X^{SO} = 2$ maximizes total group welfare.

Each player is assigned one of 5 uniquely labeled “Recommendation Balls” by the computer after the state is realized but not showen to them. The assignment is done without replacement, meaning each player receives a different label within their group. This ensures that the signal distribution is deterministic at the group level while appearing individualized to each participant. Participants are informed that recommendations are based on the true state and are designed to help coordinate behavior, but they are not explicitly told the exact conditional probabilities. Instead, they are told that:

- If they receive “Choose Box A,” the state is more likely to be Good, but there is still some chance it could be Bad.
- If they receive “Choose Box B,” the state is definitely Bad (since no one receives this recommendation in the Good state).

This framing preserves the logic of Bayesian persuasion while simplifying interpretation for participants who may struggle with conditional probability calculations.

The use of deterministic signal assignment is intentional and departs from standard stochastic implementations of Bayesian persuasion for several reasons. First, stochastic signals at the individual level can introduce noise in laboratory settings where subjects may struggle to interpret conditional probabilities correctly (Fréchette et al., 2022). By contrast, deterministic assignments at the group level ensure that the social planner’s strategy is transparent and interpretable, reducing cognitive load and promoting belief updating consistent with theoretical predictions. Second, since the sender in this experiment is a benevolent planner rather than a strategic agent, there is no incentive to manipulate or disguise the signal structure. The fixed mapping from state to recommendation simplifies comprehension without sacrificing the key mechanism of belief manipulation. Third, this design ensures that if all participants follow their recommendations, congestion in the Bad state is exactly at the social optimum, providing a clean benchmark for evaluating compliance and welfare gains.

Importantly, recommendations are non-binding, participants are free to choose either box regardless of the recommendation they receive. This design tests whether information alone, without changing payoffs, can guide behavior toward efficient outcomes through belief manipulation and coordination.

4.4 Experimental Protocol

All sessions were conducted at the Experimental Economics Center (ExCEN) at Georgia State University, using the oTree platform (version 5.11) for experiment delivery (Chen et al., 2016). Student subjects were recruited through ExCEN's existing subject pool, which consists primarily of undergraduate and graduate students from diverse academic backgrounds. Recruitment was conducted via email invitations and subjects could sign up for available time slots on a first-come, first-served basis.

Each subject participated in only one session and one treatment, ensuring a clean between-subject design with no contamination from prior exposure to alternative mechanisms. A total of 170 subjects participated across the two treatments, distributed as follows: 85 subjects in BI (3 sessions: 30, 30, and 25 participants respectively), and 85 subjects in BT (3 sessions: 30, 30, and 25 participants respectively). Sessions lasted approximately 90 minutes, including time for instruction reading, comprehension quizzes, the main experiment, and payment.

At the beginning of each session, subjects received printed instructions specific to their treatment (see Appendix A for full text). The experimenter read the instructions aloud to ensure common knowledge and to allow subjects to ask clarifying questions. Subjects were encouraged to follow along with their printed copy and to raise their hand privately if they had questions. After the instruction period, subjects completed a computerized comprehension quiz designed to verify understanding of the game structure, payoff calculations, and treatment-specific features (such as the tax or recommendation system). The quiz included questions such as:

- "If 3 people choose Box B, what is the payoff for each person who chose Box B?"
- "If the state is Good (Yellow ball) and you choose Box A, what is your payoff?"
- BT:"If you choose Box B and 2 other people also choose Box B, what is your payoff?"
- BI: "If your Recommendation Ball says "B", what must be true??"

Subjects could not proceed to the main experiment until all quiz questions were answered correctly. If a subject answered incorrectly, they received immediate feedback and could attempt the question again. The experimenter was available to answer questions privately during this phase, but no strategic advice was given. This procedure ensured that all participants had a clear understanding of the game mechanics before making consequential decisions.

All decisions were made privately on individual computer terminals in isolated carrels to prevent communication or observation of others' choices. The experimental interface was designed to be intuitive and consistent across treatments, with clear visual displays of payoff tables, round outcomes, and (in Stage 2 of BI) signal recommendations. Participants made their choice by clicking on either "Box A" or "Box B" on their screen, and they could review payoff information before confirming their decision.

The experiment was implemented as a repeated random matching game. Within each stage, participants were randomly re-matched into new groups of five at the beginning of every round. This matching protocol was explained clearly in the instructions to ensure participants understood that they would not interact with the same group members repeatedly. The random matching serves two important purposes: it minimizes reputation effects and repeated-game strategies that could confound our analysis of one-shot interactions, and it generates a larger number of independent observations for statistical analysis.

After all participants in a session completed their decisions in a given round, the computer revealed the outcome information: the realized state (color of the ball in Box A), the number of participants in their group who chose each box, and the resulting payoff. This feedback screen remained visible for 15 seconds before the next round began, allowing participants time to process the information and update their beliefs. In the BI treatment during Stage 2, participants also saw a reminder of the recommendation they received in that round alongside the outcome, facilitating learning about the informativeness of recommendations.

Between Stage 1 and Stage 2, the experimenter provided a brief verbal announcement explaining the introduction of the new policy (tax or recommendations). For BT sessions, the announcement explained: "Starting in round 11, there will be a cost of 5 points if you choose Box B. This cost will be subtracted from your payoff." For BI sessions, the announcement explained: "Starting in round 11, before you make your choice, you will receive a recommendation from the computer. The computer's recommendation is based on information about which color ball is in Box A. You are free to follow or ignore this recommendation, it is not binding. However, the recommendations are designed to help your group achieve higher total payoffs." These announcements were scripted and read verbatim in all sessions to ensure consistency.

This paper followed a single-blind protocol: subjects were identified only by participant ID numbers, and all interactions were anonymous. Subjects were explicitly informed in the instructions that their choices would not be revealed to other participants by name and that all data would be kept confidential. This anonymity reduces social pressure and allows for more honest revelation of preferences.

Final earnings were calculated in experimental currency units (ECU) at the end of the session. One round was randomly selected for payment from each stage (two rounds total), and the sum of payoffs from these two rounds, plus a $3 show-up fee, constituted the subject's total earnings. The exchange rate was set at $0.50 per ECU, and the random payment mechanism was used to ensure incentive compatibility while limiting stake size. This approach follows standard practice in experimental economics (Cubitt et al., 1998) and ensures that participants treat each round seriously without being overwhelmed by the cumulative stakes.

Subjects earned an average of \$13.07 (SD = \$4.41), with earnings ranging from a minimum of \$5.5 to a maximum of \$18. This range reflects both the stochasticity of the random payment selection and variation in performance across rounds. Payments were made in cash immediately after the session ended, and subjects signed a payment receipt confirming the amount received. No subject declined payment or withdrew from the experiment after starting.

Complete experimental instructions for both treatments, including screenshots of the decision interface and comprehension quiz questions, are provided in Appendix.

5 Results and Discussion

5.1 Baseline Behavior and Ex-Ante Equilibrium Play

I begin by examining behavior in the first ten rounds, which form the baseline condition before any intervention is introduced. Since the baseline game is identical across treatments, I pool observations from the Tax (BT) and Information (BI) sessions in Stage 1.

Table 3 reports the average number of entrants into Box B by state. In the Good state (YELLOW), where Box A yields a sure payoff of 30 points and strictly dominates Box B for any congestion level, the average group-level entry is $X = 1.878$ (SD = 1.053). A one-sample t-test strongly rejects the Nash prediction of $X = 0$ ($t = 75.69$, $p < 0.001$). In the Bad state (BLUE), where Box A yields only 10 points and the state-contingent Nash equilibrium predicts $X = 4$, the observed average is $X = 1.869$ (SD = 1.068), again far from the benchmark ($t = -79.85$, $p < 0.001$).

Average entry is almost identical across states: $X = 1.88$ in Good and $X = 1.87$ in Bad. This pattern suggests that participants do not condition their choices on the realized state, even though state information is available on the decision screen. Instead, behavior appears organized around the ex-ante environment. When states are equally likely, the ex-ante Nash equilibrium predicts $X = 2$, where the expected payoff from Box A ($\mathbb{E}[\pi_A] =$

Table 3: Average entry into Box B by state in baseline rounds (Stage 1, pooled across treatments)

State	Mean X	SD	N	vs. NE (t)	vs. Ex-ante (t)
Good (YELLOW)	1.878	1.053	1,800	75.69^{***}	–
Bad (BLUE)	1.869	1.068	1,600	-79.85^{***}	–
Overall	1.874	1.060	3,400	–	-6.96^{***}

Notes: t-statistics test against state-contingent Nash equilibria (NE): $X = 0$ for Good state, $X = 4$ for Bad state. Ex-ante test compares overall mean to $X = 2$. $^{***}p < 0.001$.

$0.5 \times 30 + 0.5 \times 10 = 20$) equals the expected payoff from Box B at that congestion level. The pooled baseline average is close to this benchmark at $X = 1.874$ (SD = 1.060), though a one-sample test rejects equality with $X = 2$ ($t = -6.96$, $p < 0.001$).

This discovery of ex-ante equilibrium play has important theoretical implications. Standard game theory assumes that rational agents condition their strategies on all available information. Yet here, participants systematically ignore state realizations in favor of a simpler decision rule based on expected values. This behavior may reflect cognitive economizing—the mental cost of maintaining and executing state-contingent strategies may exceed the expected benefit. Alternatively, participants may doubt whether others will process state information correctly, leading to strategic uncertainty that favors robust, state-invariant choices. The simultaneous presentation of state information and choice options in my experimental interface, rather than sequential revelation, may contribute to this pattern by reducing the salience of state information relative to standard designs.

The practical implication is that policy interventions may face populations using simpler decision rules than theory assumes. If agents naturally gravitate toward ex-ante rather than state-contingent behavior, calibrating interventions based on full-information benchmarks may lead to systematic errors. This finding suggests that robust policy design should account for the possibility that agents use coarse decision rules even when finer information is available.

5.2 Treatment Effects: Divergent Responses to Tax and Information

I now compare how the Tax (BT) and Information (BI) interventions affect entry into Box B in Stage 2 (rounds 11–20). Because the baseline game is identical across treatments, Stage 1 outcomes serve as a natural comparison for each treatment.

Table 4 summarizes group-level entry X by treatment, stage, and state. In the Bad state (BLUE), where the outside option is low and the congestion-sensitive resource is relatively attractive in theory, the two interventions move behavior in opposite directions.

Table 4: Group-level entry into Box B by treatment, stage, and state

Treatment	State	Stage	Mean X	SD	N
BI	Bad (BLUE)	1	1.734	1.123	790
		2	2.494	0.967	425
	Good (YELLOW)	1	1.901	0.996	910
		2	1.282	0.967	425
BT	Bad (BLUE)	1	2.000	0.994	810
		2	1.379	0.998	435
	Good (YELLOW)	1	1.854	1.108	890
		2	1.325	0.747	415

In the BT sessions, average entry in Bad states falls from $X = 2.000$ (SD = 0.994) in Stage 1 to $X = 1.379$ (SD = 0.998) in Stage 2. The reduction of 0.621 entrants per group is statistically significant ($t = 10.49$, $p < 0.001$). The Pigouvian tax reliably reduces congestion in the state where congestion is most costly.

Table 5: Difference-in-differences analysis for Bad state

Treatment	Stage 1	Stage 2	Within-Δ	t-stat	p-value
BT (Tax)	2.000	1.379	-0.621	-10.49	< 0.001
BI (Information)	1.734	2.494	$+0.760$	11.80	< 0.001
Between-treatment Δ (Stage 2)	1.115		–	16.63	< 0.001

In contrast, the BI sessions display an increase in congestion in Bad states. Average entry rises from $X = 1.734$ (SD = 1.123) in Stage 1 to $X = 2.494$ (SD = 0.967) in Stage 2, an increase of 0.760 entrants per group ($t = 11.80$, $p < 0.001$). In Stage 2, congestion in the Bad state is therefore markedly higher in BI than in BT: $X = 2.494$ versus $X = 1.379$. The difference of 1.115 entrants per group is highly significant ($t = 16.63$, $p < 0.001$).

This divergence reveals a fundamental difference in how price and information mechanisms operate. The tax treatment's success reflects the power of simple, salient incentives. By directly entering participants' payoff calculations, the tax requires no complex inference or belief formation. Participants immediately understand that Box B now costs 5 points more, and they adjust accordingly. The reduction from $X = 2.000$ to $X = 1.379$ represents 62% progress toward the ex-ante optimal level of $X = 1$ under the tax—not perfect calibration, but clear movement in the intended direction.

The information treatment's failure, by contrast, stems from the complexity it introduces. While the planner's signal structure is theoretically optimal, it creates asymmetric information that participants exploit rather than follow. Those receiving "Choose B" recommendations learn with certainty that the state is Bad, making Box B attractive despite congestion. Those receiving "Choose A" recommendations face genuine uncertainty, and some gamble that others' compliance will leave Box B under-congested. This strategic sophistication, rather than confusion or irrationality, drives the perverse increase in congestion.

The pattern in Good states provides additional insight. Both treatments reduce entry (BI: from 1.901 to 1.282; BT: from 1.854 to 1.325), suggesting that interventions help participants recognize state information they previously ignored. However, this success in Good states cannot compensate for the information treatment's failure in Bad states, where coordination matters most.

5.3 Information Treatment Breakdown

To understand why the information treatment backfires, I examine how participants respond to different recommendations. The design sends "Choose B" recommendations to two participants and “Choose A” recommendations to three participants in Bad states, aiming to implement the social optimum of $X = 2$.

Table 6: Compliance with recommendations (BI treatment, Stage 2)

Recommendation	Compliance Rate	N obs.	N complied
“Choose A”	73.1%	680	497
“Choose B”	81.2%	170	138
Overall	74.7%	850	635

Notes: Difference in compliance rates: 8.1 pp ($\chi^2 = 4.38$, $p = 0.036$).

Overall compliance reaches 74.7%, which would typically indicate a successful information intervention. However, the asymmetric compliance rates—81.2% for “Choose B” versus 73.1% for “Choose A”—reveal the mechanism of failure. Those receiving “Choose B” know with certainty the state is Bad (since this signal is never sent in Good states), making Box B’s payoff of 20 (with two entrants) clearly superior to Box A’s payoff of 10. High compliance is rational.

Table 7: Choice patterns with and without recommendations (BI treatment, Bad state)

Condition	P(Choose A)	SD	N
Baseline (Stage 1, Bad state)	0.653	0.476	790
Stage 2, received “Choose A”	0.710	0.454	255
Stage 2, received “Choose B”	0.188	0.392	170

Those receiving “Choose A” face a more complex decision. The signal is partially informative, updating their posterior to $P(\text{Good}|\text{“Choose A”}) \approx 0.625$. While this should theoretically induce choosing A (expected value of 22.5 versus 20 from Box B at the social optimum), approximately 27% deviate. These deviators are not confused—they are making a sophisticated strategic calculation. If others follow recommendations, Box B will be under-congested, offering high payoffs if the state turns out to be Bad (which occurs with 37.5% probability conditional on their signal).

The aggregate effect is perverse: high compliance by “Choose B” recipients ($2 \times 0.812 \approx$

1.62 entrants) combined with strategic deviation by "Choose A" recipients ($3 \times 0.269 \approx 0.81$ entrants) yields total entry of approximately 2.43, close to the observed 2.494. The information intervention thus fails not from misunderstanding but from strategic sophistication—the very rationality that makes participants responsive to information also enables them to exploit the asymmetries it creates.

This finding challenges the optimistic view of information provision as a light-touch alternative to regulation. When information creates heterogeneous beliefs in strategic settings, high individual compliance can coexist with aggregate failure. Policy makers should be particularly cautious about information interventions in environments where strategic complementarities amplify individual deviations.

5.4 Tax Treatment Breakdown

The tax treatment's reduction in entry from $X = 2.000$ to $X = 1.379$ might initially appear to overshoot the social optimum of $X = 2$ in Bad states. However, this interpretation misunderstands the appropriate benchmark. Given that participants play according to ex-ante rather than state-contingent equilibria, the relevant target is the ex-ante optimum under the tax.

With $\tau = 5$, the new equilibrium condition becomes:

$$\begin{aligned} \mathbb{E}[\pi_A] &= \mathbb{E}[\pi_B - \tau] && (20) \\ 20 &= 25 - 5X && (21) \\ X &= 1 && (22) \end{aligned}$$

The observed value of $X = 1.379$ represents 62% movement toward this target from the baseline of 2.000. The tax successfully shifts behavior in the intended direction, though calibration remains imperfect. This likely due to risk aversion or residual behavioral factors not captured in the basic model.

The tax's effectiveness stems from its simplicity. Unlike information that must be processed, interpreted, and incorporated into beliefs about others' behavior, a price change directly enters the decision calculus. Participants need not wonder how others will respond to the tax or whether others understand it correctly, and the dominant strategy shifts uniformly for all players. This transparency may explain why the tax effect appears immediately in round 11 with no adjustment period, while the information treatment shows gradual deterioration over time.

Moreover, the tax achieves spillover benefits in Good states, where entry falls from 1.854 to 1.325. Since any entry in Good states is inefficient, this reduction improves welfare even though the tax was calibrated for the Bad state. This illustrates how blunt instruments can sometimes outperform targeted interventions through their very simplicity.

5.5 Welfare Implications: When Good Intentions Produce Bad Outcomes

The welfare consequences of the two treatments provide the ultimate test of their effectiveness. Table 8 reports group welfare by treatment and stage.

Table 8: Group welfare by treatment and stage

Treatment	Stage	Mean welfare	SD	% of Max	N
BI	1	92.8	12.3	84.4%	1,700
BI	2	87.3	11.8	79.4%	850
BT	1	93.0	12.1	84.5%	1,700
BT	2	95.7	11.5	87.0%	850

Notes: Maximum theoretical welfare = 110. Stage 2 difference: $t = 2.34$, $p = 0.019$.

The tax treatment modestly improves welfare from 93.0 to 95.7 points (+2.9%), achieving 87% of the theoretical maximum. The information treatment, however, reduces welfare from 92.8 to 87.3 points (-5.9%), falling to just 79.4% of maximum. The 8.4-point gap between treatments in Stage 2 is economically substantial and statistically significant ($t = 2.34$, $p = 0.019$).

The welfare deterioration under information is particularly striking given the 75% compliance rate. This disconnect between micro-level compliance and macro-level outcomes illustrates why intermediate metrics can be misleading. A policy maker observing only compliance rates might conclude the information intervention succeeded, missing the strategic dynamics that produce aggregate failure.

The welfare effects also reveal distributional consequences. Under the tax, the payoff differential between Box A and Box B choosers narrows but remains ordered correctly. Under information, strategic gambling by some "Choose A" recipients creates arbitrary winners (those who successfully gamble on low congestion) and systematic losers (those who comply with "Choose A" when congestion is high). This ex-post inequality, arising from differential information rather than differential choices, may raise fairness concerns beyond pure efficiency considerations.

5.6 Temporal Dynamics: Immediate Impact versus Gradual Unraveling

How treatment effects evolve over time provides insight into their underlying mechanisms. Table 9 reports time trends in Stage 2 entry.

Table 9: Time trends in Stage 2 entry

Treatment	Coefficient	SE	p-value
BT (all states)	-0.016	0.029	0.580
BI (all states)	0.071	0.053	0.190
BI (Bad state only)	0.096	0.045	0.036

Notes: Dependent variable is group-level entry X. Standard errors clustered at session level.

The tax treatment shows no significant time trend ($\hat{\beta} = -0.016$, $p = 0.58$), indicating immediate and sustained effectiveness. Participants quickly understand the price change and maintain their adjusted behavior throughout Stage 2. This stability suggests that the tax effect represents an equilibrium response rather than a transitory reaction.

The information treatment, by contrast, shows positive drift, especially in Bad states ($\hat{\beta} = 0.096$, $p = 0.036$). Entry increases gradually over rounds 11-20, suggesting that strategic gambling intensifies with experience. As participants observe that some "Choose A" recipients successfully enter an under-congested Box B, more may be tempted to deviate in subsequent rounds. This unraveling dynamic further undermines the theoretical appeal of information-based coordination.

The contrasting temporal patterns highlight a crucial difference between price and information mechanisms. Prices create common knowledge that facilitates immediate coordination on a new equilibrium. Information creates heterogeneous beliefs that may initially seem to coordinate behavior but gradually unravel as players learn to exploit informational advantages. For policy makers, this suggests that information interventions may show promising initial results that deteriorate over time—a pattern that short-term pilots might miss.

6 Conclusion

This paper provides experimental evidence comparing Pigouvian taxation and Bayesian persuasion as interventions for managing congestion externalities with state-dependent payoffs. The results challenge conventional wisdom about the relative merits of price versus information mechanisms while revealing fundamental insights about strategic behavior under uncertainty.

Three main findings emerge from the experiment. First, participants naturally coordinate on ex-ante rather than state-contingent equilibria, playing $X \approx 1.87$ regardless of whether the outside option yields 30 or 10 points. This state-invariant behavior, while departing from standard game-theoretic predictions, may reflect a rational response to strategic uncertainty or cognitive constraints. The discovery suggests that policy calibration based on full-information benchmarks may systematically miss the mark.

Second, the two interventions produce dramatically different effects in the critical Bad state where coordination matters most. The tax reduces entry from 2.00 to 1.38 participants, achieving 62% of the movement toward the ex-ante optimal level of $X = 1$. The information treatment increases entry from 1.73 to 2.49 participants, overshooting even the baseline level. This divergence occurs despite 75% compliance with recommendations, highlighting how individual rationality can produce collective irrationality.

Third, the mechanisms driving these effects differ fundamentally. The tax succeeds through simplicity, it directly alters payoffs in a transparent, universal way that requires no complex inference or beliefs about others' behavior. The information treatment fails through sophistication, and it creates asymmetric beliefs that strategic players exploit rather than follow, with those receiving partial information gambling on others' compliance to access under-congested resources.

6.1 Theoretical Implications

The findings contribute to several theoretical literatures. For the theory of congestion games, the discovery of ex-ante equilibrium play suggests that standard refinements assuming state-contingent best responses may be too strong. Models incorporating coarse reasoning or robust decision-making may better capture observed behavior.

For the theory of information design, the results illustrate the fragility of optimal persuasion in multi-agent settings with externalities. While Bayesian persuasion can theoretically implement any Bayes correlated equilibrium, the behavioral requirements: correct belief updating, common knowledge of rationality, and coordinated best responses. All may be unrealistic when stakes are meaningful and strategic considerations are salient.

For the theory of mechanism design more broadly, the experiment highlights a tradeoff between theoretical sophistication and behavioral robustness. Simple mechanisms that are "obviously strategy-proof" may outperform complex mechanisms that are merely strategy-proof in equilibrium. The tax's transparency and immediacy provide behavioral advantages that dominate the information treatment's theoretical superiority.

6.2 Policy Implications

The results offer practical guidance for policy makers choosing between price and information instruments. In settings with congestion externalities and strategic interaction, simple taxes appear more reliable than sophisticated information provision. Even when taxes cannot achieve state-contingent optima, their robust directional effects may dominate the theoretical precision of information design.

The 75% compliance rate with recommendations, combined with aggregate failure, warns against using intermediate metrics to evaluate information interventions. Policy pilots should examine equilibrium outcomes, not just individual responses, and should extend observation periods to detect potential unraveling over time.

The findings are particularly relevant for digital platforms managing congestion through information provision. Navigation apps, for instance, might achieve better outcomes through simple congestion pricing than through sophisticated routing algorithms that create asymmetric information. Similarly, environmental policies relying on information disclosure may be vulnerable to strategic exploitation in ways that simple taxes avoid.

6.3 Limitations and Future Research

Several limitations qualify these conclusions. The experiment uses a stylized environment with binary choices, perfect information about payoff structures, and small group sizes. Real-world settings may involve continuous choices, payoff uncertainty, and large-scale coordination that could alter the relative effectiveness of price versus information mechanisms.

The specific parameters chosen: equal probability of states, linear congestion costs, and particular tax levels. All may influence the results. Future work should explore robustness across parameter spaces, potentially identifying conditions where information interventions perform better.

The experiment also abstracts from important practical considerations such as implementation costs, political feasibility, and distributional concerns. While the tax performed better on efficiency grounds, information interventions might be preferred when price mechanisms face political constraints or raise equity concerns.

Future research could extend this framework in several directions. Combining price and information instruments might achieve better outcomes than either alone. Allowing for heterogeneous agents could reveal whether information interventions work better when targeted to sophisticated players. Examining repeated interactions could test whether reputation effects enhance or undermine information-based coordination.

The experiment ultimately delivers a simple but important message: in strategic settings with externalities, the theoretical appeal of sophisticated mechanisms may be undermined by the very rationality they assume. The humble Pigouvian tax, despite its inability to achieve first-best outcomes, provides behavioral robustness that sophisticated information design lacks. As policy makers increasingly turn to behavioral interventions and nudges, remembering the enduring value of simple price instruments remains important. Sometimes, the blunt tool is the right tool.

References

Allcott, H. and Rogers, T. (2014). The short-run and long-run effects of behavioral interventions. *American Economic Review*, 104(10):3003–3037.

Anderson, L. R., Holt, C. A., and Reiley, D. (2008). Congestion pricing and welfare: An entry experiment. In Cherry, T. L., Kroll, S., and Shogren, J. F., editors, *Environmental Economics, Experimental Methods*, pages 280–291. Routledge.

Andreassen, G. L., Kallbekken, S., and Rosendahl, K. E. (2024). Can policy packaging help overcome pigouvian tax aversion? a lab experiment on combining taxes and subsidies. *Journal of Environmental Economics and Management*, 127:103010.

Aristidou, A., Coricelli, G., and Vostroknutov, A. (2023). Incentives or persuasion? an experimental investigation. Working Paper.

Au, P. H., Kwon, O., and Li, K. K. (2023). A simple experiment on simple bayesian persuasion. *Working Paper.*

Bergemann, D. and Morris, S. (2016). Information design, bayesian persuasion, and bayes correlated equilibrium. *American Economic Review*, 106(5):586–591.

Bergemann, D. and Morris, S. (2019). Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95.

Cason, T. N. and Gangadharan, L. (2016). Taxes and subsidies in voluntary public goods provision. *Working Paper.*

Charness, G. and Gneezy, U. (2010). Laboratory experiments: Implementing independent private values. *Experimental Economics*, 13(3):317–343.

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

Cubitt, R. P., Starmer, C., and Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics*, 1(2):115–131.

Das, S., Kamenica, E., and Mirka, R. (2017). Reducing congestion through information design. In *2017 55th annual allerton conference on communication, control, and computing (allerton)*, pages 1279–1284. IEEE.

Delaney, J. and Jacobson, S. (2016). Payments or persuasion: Common pool resource management with price and non-price measures. *Journal of Behavioral and Experimental Economics*, 62:1–13.

Ferraro, P. J. and Price, M. K. (2013). Money for nothing? a call for empirical evaluation of biodiversity conservation investments. *PLoS Biology*, 11(4):e1001537.

Fréchette, G. R., Lizzeri, A., and Perego, J. (2022). Rules and commitment in communication: An experimental analysis. *Econometrica*, 90(5):2283–2318.

Hardin, G. (1968). The tragedy of the commons. *Science*, 162(3859):1243–1248.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5):1644–1655.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Kallbekken, S., Kroll, S., and Cherry, T. L. (2011). Do you not like pigou, or do you not understand him? tax aversion and revenue recycling in the lab. *Journal of Environmental Economics and Management*, 62(1):53–64.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Lu, X., Gao, S., and Ben-Elia, E. (2011). Information impacts on route choice and learning behavior in a congested network: experimental approach. *Transportation Research Record*, 2243(1):89–98.

Mak, V., Gisches, E. J., and Rapoport, A. (2015). Route vs. segment: an experiment on real-time travel information in congestible networks. *Production and Operations Management*, 24(6):947–960.

Nguyen, Q. (2017). Bayesian persuasion: evidence from the laboratory. *Work. Pap., Utah State Univ., Logan.*

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action.* Cambridge University Press.

Ostrom, E., Walker, J., and Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American political science Review*, 86(2):404–417.

Pigou, A. C. (1920). *The Economics of Welfare.* Macmillan and Co., London.

Plott, C. R. (1983). Externalities and corrective policies in experimental markets. *The Economic Journal*, 93(369):106–127.

Qiao, L. and Noussair, C. N. (2025). Traffic apps and traffic congestion: An experiment. *Conditionally accepted at Management Science.*

Rapoport, A., Seale, D. A., and Ordoñez, L. (2009). Coordination in market entry games with symmetric players. *Organizational Behavior and Human Decision Processes*, 108(2):215–229.

Walker, J. M., Gardner, R., Herr, A., and Ostrom, E. (2000). Collective choice in the commons: Experimental results on proposed allocation rules and votes. *The Economic Journal*, 110(460):212–234.

Walker, J. M., Gardner, R., and Ostrom, E. (1990). Rent dissipation in a limited-access common-pool resource: Experimental evidence. *Journal of Environmental Economics and Management*, 19(3):203–211.

Ziegler, A. (2023). Persuading an audience: Testing information design in the laboratory. Technical report, Tinbergen Institute Discussion Paper.

A Subject Instructions: Treatment BT and BI

Experiment Instructions

Welcome

Welcome to the study!

No Talking Allowed
This is an experiment about decision-making. Your decisions will affect your earnings and may also affect the earnings of others. Please read the instructions carefully. Do not talk with other participants. If you have any questions, raise your hand and an experimenter will come help you privately.

Privacy
This experiment is structured so that no one, including the experimenters or the other participants, will ever know your personal decision. Your privacy is guaranteed because neither your name nor student ID number will be entered into the computer that records your decisions in the experiment. The only identifying information is your seat number, which is used solely to anonymously connect your decisions to your payments.

Multiple Rounds
The experiment consists of two parts. Part 1 has 10 rounds and Part 2 has 10 rounds. **Each round is equally likely to be selected to determine your earnings at the end of the experiment,** so you should take each round seriously.

Random Matching and Anonymity
In each round, you will be randomly and anonymously matched with **four** other participants to form a group of **five**. Group members may change from round to round, and you will not know who they are. Your choice in each round does not affect what happens in other rounds, and past group members will not be identified. You should make your decisions independently each time.

Cash Earnings
You will receive a $3 participation payment. Additional earnings will depend on ONE randomly selected round. Earnings are in U.S. dollars at the end of the experiment.

Experiment Instructions

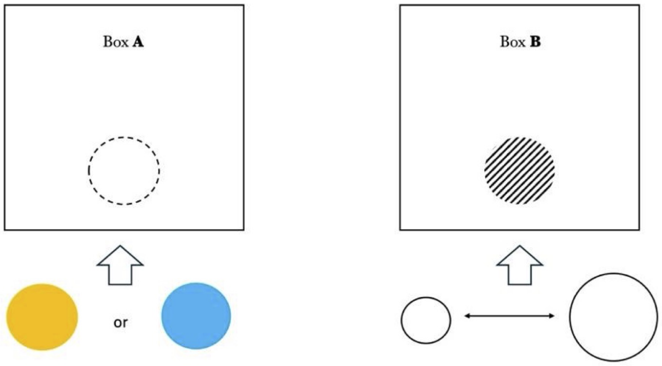
Decision Task Instructions

You will now begin Part 1 of the experiment. All decisions will be made using the computer. Please read the instructions below carefully, as they explain the task and how your earnings will be determined.

Part 1:

On your screen, you are presented with Box A and Box B.

- The computer places one ball in each box.
- The ball in **Box A** has an **unknown color**, either **YELLOW** or **BLUE**.
- The ball in **Box B** has an **unknown size**, which can range from **small to large**.



How the Ball in Box A Is Determined

Before each round, the computer randomly selects the color of the ball in **Box A**:

- It is either **YELLOW** or **BLUE**, with equal chance (1/2)
- The color is the **same for everyone** in your group
- You will **not** see the color before making your choice

How the Ball in Box B Is Determined

The size of the ball in Box B depends on how many people in your group choose Box B in that round. Box B becomes **smaller** as more people choose it

Decision Task

Your task is to choose either Box A or Box B. Your earnings depends on your choice:

Experiment Instructions

Earnings

- If you choose Box A:
  - The earnings is determined by the **color of the ball** in Box A.
  - The ball is:
    - YELLOW → You earn 30.
    - BLUE → You earn 10.
- If you choose Box B:
  - The earning is determined by the **size of the ball** in Box B.
  - The size depends on how many people in your group also choose Box B
  - Specifically:

Number of people choosing Box B (including you)	Earning from Box B
1	25
2	20
3	15
4	10
5	5

You will not know the color or the size of the ball until after you make your decision for the round. At end of each round, you will learn the color and the size of the two balls.

What You See After Each Round

After each round, you will learn:

- The **color of the ball in Box A**
- The **size of the ball in Box B**
- Your **earnings for that round**

Each round is independent:

Each round is a fresh start. The ball in Box A is randomly redrawn, your group changes, and nothing from previous rounds carries over.

Example Decisions

Example	**Your Choice**	**Ball Color / Group Behavior**	**Your Earnings**
1	Box A	**YELLOW**	30
2	Box A	**BLUE**	10
3	Box B	You were the only one	25
4	Box B	2 others chose Box B (3 total)	15

Decision Task Instructions

You will now begin Part 2 of the experiment. All decisions will be made using the computer. Please read the instructions below carefully, as they explain the task and how your earnings will be determined.

Part 2:

Part 2 is similar to Part 1, **but the earnings from Box B are different**.

Earnings

- If you choose Box A:
  - The earnings is determined by the **color of the ball** in Box A.
  - The ball is:
    - YELLOW → You earn 30.
    - BLUE → You earn 10.
- If you choose Box B:
  - The earning is determined by the **size of the ball** in Box B.
  - The size depends on how many people in your group also choose Box B
  - Specifically:

Number of people choosing Box B (including you)	Earning from Box B **In part 2**
1	**20**
2	**15**
3	**10**
4	**5**
5	**0**

Decision Task

Your task is to choose either Box A or Box B. Your earnings depends on your choice:

Example Decisions

Example	**Your Choice**	**Ball Color / Group Behavior**	**Your Earnings**
1	Box A	YELLOW	30
2	Box A	BLUE	10
3	Box B	You were the only one	20
4	Box B	1 others chose Box B (2 total)	15

Experiment Instructions

Welcome

Welcome to the study!

No Talking Allowed
This is an experiment about decision-making. Your decisions will affect your earnings and may also affect the earnings of others. Please read the instructions carefully. Do not talk with other participants. If you have any questions, raise your hand and an experimenter will come help you privately.

Privacy
This experiment is structured so that no one, including the experimenters or the other participants, will ever know your personal decision. Your privacy is guaranteed because neither your name nor student ID number will be entered into the computer that records your decisions in the experiment. The only identifying information is your seat number, which is used solely to anonymously connect your decisions to your payments.

Multiple Rounds
The experiment consists of two parts. Part 1 has 10 rounds and Part 2 has 10 rounds. **Each round is equally likely to be selected to determine your earnings at the end of the experiment,** so you should take each round seriously.

Random Matching and Anonymity
In each round, you will be randomly and anonymously matched with **four** other participants to form a group of **five**. Group members may change from round to round, and you will not know who they are. Your choice in each round does not affect what happens in other rounds, and past group members will not be identified. You should make your decisions independently each time.

Cash Earnings
You will receive a $3 participation payment. Additional earnings will depend on ONE randomly selected round. Earnings are in U.S. dollars at the end of the experiment.

Experiment Instructions

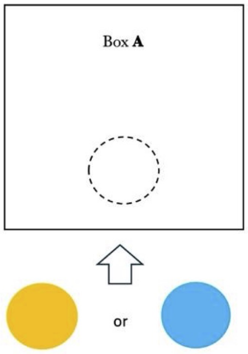
Decision Task Instructions

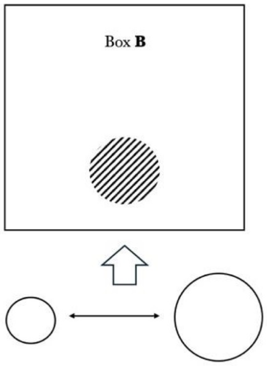
You will now begin Part 1 of the experiment. All decisions will be made using the computer. Please read the instructions below carefully, as they explain the task and how your earnings will be determined.

Part 1:

On your screen, you are presented with Box A and Box B.

- The computer places one ball in each box.
- The ball in **Box A** has an **unknown color**, either **YELLOW** or **BLUE**.
- The ball in **Box B** has an **unknown size**, which can range from **small to large**.





How the Ball in Box A Is Determined

Before each round, the computer randomly selects the color of the ball in **Box A**:

- It is either **YELLOW** or **BLUE**, with equal chance (1/2)
- The color is the **same for everyone** in your group
- You will **not** see the color before making your choice

How the Ball in Box B Is Determined

The size of the ball in Box B depends on how many people in your group choose Box B in that round. Box B becomes **smaller** as more people choose it

Decision Task

Your task is to choose either Box A or Box B. Your earnings depends on your choice:

Experiment Instructions

Earnings

- If you choose Box A:
  - The earnings is determined by the **color of the ball** in Box A.
  - The ball is:
    - **YELLOW** → You earn 30.
    - **BLUE** → You earn 10.
- If you choose Box B:
  - The earning is determined by the **size of the ball** in Box B.
  - The size depends on how many people in your group also choose Box B
  - Specifically:

Number of people choosing Box B (including you)	Earning from Box B
1	25
2	20
3	15
4	10
5	5

You will not know the color or the size of the ball until after you make your decision for the round. At end of each round, you will learn the color and the size of the two balls.

What You See After Each Round

After each round, you will learn:

- The **color of the ball in Box A**
- The **size of the ball in Box B**
- Your **earnings for that round**

Each round is independent:

Each round is a fresh start. The ball in Box A is randomly redrawn, your group changes, and nothing from previous rounds carries over.

Example Decisions

Example	**Your Choice**	**Ball Color / Group Behavior**	**Your Earnings**
1	Box A	**YELLOW**	30
2	Box A	**BLUE**	10
3	Box B	You were the only one	25
4	Box B	2 others chose Box B (3 total)	15

Decision Task Instructions

You will now begin Part 2 of the experiment. All decisions will be made using the computer. Please read the instructions below carefully, as they explain the task and how your earnings will be determined.

Part 2:

Before making your choice, you will receive a **Recommendation Ball** that recommend you to choose Box A or Box B.

Each round, there are **5 players**, and the computer prepares **5 private Recommendation Balls** — one per person.

The Recommendation on each ball depends on the color of the ball in Box A.

How the Recommendation Ball is Generated

If the Ball in Box A is YELLOW:

- All 5 Recommendation Balls are labeled A.

If the Ball in Box A is BLUE:

- 3 Recommendation Balls labeled "A"
- 2 Recommendation Balls labeled "B"

Each participant is randomly given **one** of these Recommendation Balls and sees **only their own**.
You will not see which Recommendations were given to the other players until the round ends.

Decision Task

Each round in Part 2 will proceed in the following order:

1. You Receive Your Recommendation Ball

The computer will draw a Recommendation Ball for you.
You will see only your own Recommendation Ball (say either A or B).

2. Your task is to choose either **Box A** or **Box B**. Your earning depends on your choice.

What You See

- You will see **only your own** Recommendation Ball
- You will **not** see the color of the Ball in Box A
- You will **not** see the Recommendations given to other participants until the round ends

Earnings and What You See After Each Round and Each round is independent: