Selection Neglect in Policing Decisions *

Santiago De Martini[1] Juan B. González[1]
Santiago M. Perez-Vincent[2]

[1]University of Southern California
[2]Inter-American Development Bank

March 22, 2025

(Most Recent Version Here)

Abstract

Police officers exhibit wide racial disparities when predicting which individuals are more likely to commit a crime. We show this pattern of statistical discrimination can be partially explained by common cognitive biases that distort how officers analyze crime data. This paper focuses on two of these biases: not accounting for the fact that crime data is not a representative sample (selection neglect) and exaggerating differences in criminality across groups (representativeness). We model police inference as a decision maker who infers the type of an individual who belongs to a specific group, when the decision maker only has information about a group-specific prior and an individual signal. In particular, we focus on how representativeness can distort such priors and how selection neglect distorts the perception of the signal. We design a novel framed field experiment to evaluate the predictions of the model and estimate these biases at the individual level, having police officers from across Latin America as participants. Finally, we study whether there is an association between these biases and discrimination of minorities by studying the correlation between the estimated biases and an Implicit Association Test.

*We thank Colin Camerer, Giorgio Coricelli, Jonathan Libgober, Alejandro Martínez-Marquina, Paulina Oliva, Antonio Rangel, and the participants of the Graduate Behavioral and Experimental Economics Conference at Caltech for thoughtful feedback on the study.

1 Introduction

Evidence shows systematic racial disparities in policing decisions based on crime predictions. Compared to White people, Black and Hispanic individuals are stopped and searched more often (Pierson et al., 2020), are exposed to more policing in their neighborhoods (Chen et al., 2023; Vomfell & Stewart, 2021), face more interactions with the police in general (Ba et al., 2021), and receive worse police service when victimized (Rizzotti, 2024). Crime predictions by the police are not only often biased, they are also inaccurate. Conditional on being stopped, Black and Hispanic individuals are less likely to have committed a crime (Pierson et al., 2020) or to carry a weapon (Fryer, 2019). The systematic statistical discrimination in prediction-led interactions are crucial to understand the racial disparities in arrests, police brutality, and murder by the police (Fryer, 2019; USSC, 2023), disparities so wide that murder by the police has become the leading cause of death for young Black males in the US (Edwards et al., 2019). Indeed, Chen et al. (2023) show that prediction-led policing decisions — over-patrolling Black and Hispanic neighborhoods — explain around 60% of the racial gap in arrests. Therefore, to understand and mitigate racial disparities in policing decisions, it's crucial to understand why systematic statistical discrimination arises in the first place.

What role do crime predictions play in policing decisions? Since these decisions follow the objective of maximizing arrests (Stashko, 2023), departments and officers need to predict the location of future crimes as well as who will commit them (Brayne, 2020). In most cases, these predictions rely on historical data collected by the police and, increasingly, on crime prediction algorithms (Mohler et al., 2015). Thus, previous crime data is used as inputs for statistically inferring the likelihood of crime across areas and individuals and guide further policing decisions. Even in the absence of taste-based discrimination, systematic cognitive distortions in the inference process have the potential of generating statistical discrimination. For instance, Dube et al. (2024) show that increasing cognitive reflection before making decisions reduce discrimination among officers of the Chicago Police Department. This paper contributes to this literature by modeling and experimentally testing how a common cognitive bias, selection neglect, distorts distort crime prediction among police officers and generate inaccurate statistical discrimination.

Selection neglect arises when a decision maker considers a selected — non-random — sample as representative of a population (Enke, 2020). The sampling or selection process then shapes the results of inference based on that sample. For instance, we might think most people in the country hold the same political beliefs as us if we base our population inference on the sample of our friends, with whom we became friends in the first place in part because we share political beliefs. Indeed, selection neglect is a key mechanism to why echo chambers and polarization emerge (Brundage et al., 2024). In economics, selection neglect explains why people infer the national income distribution from the sample of people they interact with (Cruces et al., 2013), or why investors are overoptimistic about returns when their expectation is based on previous successful investments (Barron et al., 2024; Jehiel, 2018).

Selection problems compound in dynamic settings: when current predictions select future data collection, selection neglect can generate a cycle of biased sampling leading to biased predictions, and vice versa. In the context of crime, enforcement decisions are especially vulnerable to these cycles, as Hübert and Little (2023) and Che et al. (2024) show theoretically. The argument is simple. Policing decisions follow the predicted criminality of areas and individuals, predictions based on previous crime data (number of arrests in an area, for instance). Notice that, conditional on a level of crime, the number of arrests is increasing in the level of policing in an area (Chen et al., 2023). Therefore, as minority neighborhoods are more intensely policed than their counterparts, the share of non-White criminals arrested could be higher even if crime rates were the same across groups. If decision makers in police departments neglect selection and don't account for policing levels, they might infer that crime rates among minority populations are higher than what they actually are and in turn increase policing in those areas even more, creating a vicious "cycle of bias" (Dittmann et al., n.d.). Neglecting selection leads to inferring crime likelihoods through a misspecified model and a misallocation of policing, amplifying over time policing disparities across racial groups.

While taste-based discrimination has been shown to drive policing decisions to some extent (Fryer, 2019), we show selection neglect can drive disparities in policing outcomes in the absence of taste differences. Take the case of widespread predictive policing algorithms. These algorithms predict the location of future crime, a purely statistical inference exercise with no role for taste. However, predictive policing algorithms don't account for data selection and so generate spurious statistical discrimination when trained with crime data selected by policing activity (Brayne, 2020; Lum & Isaac, 2016). As cognitive distortions often show more potential for interventions than taste-based bias, in this paper we focus on how selection neglect affect statistical inference and what interventions can reduce biases in policing decisions.

This paper models how selection neglect generates inaccurate statistical discrimination, building on previous work by Esponda et al. (2023) and Hübert and Little (2023). In particular, our decision maker forms their beliefs of a type t of an individual who belongs to group g, based on a group-specific prior distribution and an observed signal. As a result, the estimated type $\hat{t}$ is a function of the structural parameters and the data observed by the individual. Even though the model is motivated by how police force form their beliefs about the crime rates of different groups, the application of it is more general an extends beyond this specific context.

The main contribution of this paper is to provide novel experimental evidence from the field to estimate the structural parameters in an ecologically valid environment. We recruit N participants among a pool of workers in the justice and safety sectors across Latin America, including M police officers, for a framed field experiment [1]. In the experiment, there are two cities consisting of a continuum of neighborhoods, and participants are asked to estimate the crime rate (*type*) of some of these neighborhoods, based on observed information. We fix participants' crime-rate priors by showing the distribution of crime rates across cities. Then, in each round[2], a neighborhood from each city is randomly drawn from its city-specific distribution and participants observe a noisy signal of its crime rate: previous police reports on the neighborhood. However, as the number of crimes reported depend on the level of patrolling on the neighborhood, this signal is selected and not necessarily representative of the crime rate of the neighborhood. Participants are informed about this selection process as well as about the level of patrolling in each neighborhood. Therefore, if the participant accounts for data selection, she will have the information to back up the undistorted signal (indicative of the true type). After observing this signal and the number of patrols, participants estimate the crime rate of the neighborhood and are paid accordingly to the accuracy of their guess.

[1]The participant pool is a panel of workers invited to participate at IADB training courses.

[2]The experiment consists of 25 rounds.

We design several variations of the experiment to answer possible concerns. First, we exogenously vary the level of selection across neighborhoods, making neighborhoods in both cities equally patrolled in one treatment, and letting one city to be more patrolled in another. This design allows us to estimate representative prior distortion with and without selection neglect, to rule out that the former is affected by the latter. Second, we vary whether participants can observe to which city each neighborhood belongs to. Similarly, when city belonging is not observed, representativeness should not bias belief updating, such that we would be able to estimate selection neglect *without* representative distortion. Third, we test whether the combination of both biases lead to more inaccurate statistical discrimination by including both the information of the city of each neighborhood and by varying the level of patrolling across neighborhoods. Finally, we vary the setting of the experiment to test whether inference problems by the police are specific to the crime context. Specifically, we conduct the same task but instead of asking about the crimes in a neighborhood, participants will have to guess the number of trees.

The last part of our experiment consists in eliciting individual measures discrimination against minority groups to relate these estimates to the estimated individual parameters. This would allow us to test whether individuals who are more prone to fall into cognitive biases can also more likely to discriminate outside the lab. We argue that the cognitive distortions discussed in this paper not only have efficiency implications[3] but can also form incorrect beliefs that result in inaccurate statistical discrimination.

2 Theoretical Framework

2.1 Selection Neglect and Statistical Inference

In this section we model how selection neglect can lead to inaccurate statistical discrimination. The model draws from Esponda et al. (2023) to develop the inference framework and from Hübert and Little (2023) for the particular functional form of selection neglect. While we focus on the application to the context of policing, the model applies to any inference process

[3] By creating inefficient patrolling allocations across neighborhoods, already explored theoretically by Che et al. (2024)

where data might be selected.

A decision maker (e.g., a police officer) must estimate the type t of some individual. For instance, the type could reflect the crime rate of a neighborhood or whether an individual committed a crime. The individual belongs to a group g (e.g., gender, race, city) that is observed by the decision maker. Types are drawn from a distribution $f(t|g)$, which varies across groups. The decision maker observes a noisy signal s, which is informative of the individual's type. In the policing context, s correspond to previous crime data from a neighborhood or a report about a suspect. Signals are distributed according to some distribution $h(s|t, g)$.

We assume that the types are normally distributed around a group-specific mean (μ_g) with common variance σ^2. Similarly, signals are normally distributed, centered at the true type t and with group-specific variance υ_g^2. Therefore, $t \sim \mathcal{N}(\mu_g, \sigma^2)$ and $s \sim \mathcal{N}(t, \upsilon_g^2)$. A Bayesian decision maker would update their posterior beliefs of t given the signal s and the group information g following $f^p(t|s, g) = \frac{h(s|t,g)f(t|g)}{\int h(s|t,g)f(t|g)\,dt}$. The optimal decision given the signal and the observed group would be the expected posterior belief: $E[t|g, s] = \int f^p(t|s, g)\, t\, dt$. In this setting, the optimal estimate of $\hat{t}$ for a Bayesian decision is a convex combination of the group prior and realized signal[4]:

$$\hat{t} = \omega_g \mu_g + (1 - \omega_g)s \tag{1}$$

where ω_g is the weight put on the prior mean. A Bayesian agent would choose this weight according to the precision of the prior relative to the signal, this is, $\omega_g^* = \frac{\upsilon_g^2}{\sigma^2 + \upsilon_g^2}$.

We consider the case when the signal might be distorted by *selection neglect*. Selection neglect refers to situations where the decision maker takes a selected (non-random) sample as representative of the population. For instance, Barron et al. (2024) and Jehiel (2018) study this bias in the context of equity investors. Specifically, they demonstrate that investors base their decisions on the outcomes of past investments while neglecting the fact that this data excludes investments that were never made, ultimately resulting in over-investment. Enke (2020) takes a set of 6 draws and asks participants to estimate the mean of the set. To guide their estimate, participants observe a subset of up to 5 signals, but are only shown those

[4] See Appendix A1 for a detailed derivation.

signals that are consistent with their priors (in line with their prior beliefs). Even though participants know the signal generating process, they fail to consider the signals are selected, leading to systematic under or over estimation. This recent literature consistently shows that decision makers often fail to account for data selection when making inference.

Policing decisions are highly vulnerable to data selection issues, as crime data depends on the level of policing and reporting[5]. If a neighborhood is more intensely policed, the police is more likely to observe the crimes happening there. The same happens if patrolling differs across racial or income groups: if individuals of some group are stopped and searched more often, a higher share of criminals[6] will be found in this group due to oversampling. We denote the type t as the unobserved true criminality of subjects or neighborhoods of group g. Each group is subject to a level of policing p_g, which is known to the decision maker. By policing, the decision maker draws a signal s that is informative of the true type t. In the case of individuals, signals can be evidence coming from stopping and searching suspects, and in the case of neighborhoods, the observed number of crimes in a given neighborhood. In both cases, the signal is increasing in the level of policing: more policing implies a higher likelihood of observing crime (Chen et al., 2023). We denote this mapping as $c(.)$ and let the observed crime signal be $s = c(p_g, t; \epsilon_g)$. We don't impose a functional form on $c(.)$, but we assume it to have an additively separable deterministic and a stochastic component: $c(p_g, t; \epsilon_g) = c(p_g, t) + \epsilon_g$. The deterministic component is an increasing function of the true type t and the level of policing p_g, while the stochastic component accounts for noise in the data collection noise, with $\epsilon_g \sim \mathcal{N}(0, \upsilon_g)$. Note that $\mathbb{E}\left[\frac{\partial c(p_g,t;\epsilon_g)}{\partial p_g}\right] > 0$, which means that the number of arrests are, in expectation, increasing with the policing level.

When a decision maker observes the signal s, she might or might not realize it's selected when inferring the type t. Take the two extreme cases. If the decision maker perfectly accounts for data selection, she realizes that observed crime is a function of the level of policing and uses the inverse of this function to map the observed data to the crime level generating it.

[5] While the level of crime reporting by victims also affects what crime signals are observed, this paper focus on the role of policing because its endogeneity. Policing decisions are made based on previously selected crime data, whereas reporting decisions are not.

[6] In this context, a higher share of criminals means a higher number of criminals in a given group relative to the group's population.

She then uses this backed-up crime level as the signal. This is, a decision maker that perfectly accounts for selection knows that, for a level $\bar{p}_g$ of policing, $t = E[c(\bar{p}_g, t; \epsilon_g)^{-1}]$, and she updates her beliefs using the backed-up signal $s^* = c(\bar{p}_g, t; \epsilon_g)^{-1}$. On the other hand, when the decision maker neglects that the observed crime data is selected, she will interpret it as representative of the true type. That is, she incorrectly believes that $t = c(\bar{p}_g, t; \epsilon_g)$, so uses the signal $s = c(\bar{p}_g, t; \epsilon_g)$ to update. In between the two extremes, we follow Hübert and Little (2023) and assume the perceived signal $\tilde{s}$ is a convex combination of the observed data s and the backed-up signal s^* that accounts for selection: $\tilde{s} = \lambda s + (1-\lambda)s^*$, where λ captures the level of selection neglect.

Selection neglect affects the optimal estimate $\hat{t}$ by taking the potentially distorted $\tilde{s}$ as signal. The optimal estimate then takes the following form:

$$\hat{t} = \omega_g \mu_g + (1-\omega_g)\tilde{s} = \omega_g \mu_g + (1-\omega_g)\left[\lambda s + (1-\lambda)s^*\right] + \epsilon \tag{2}$$

From the optimal estimate, we can calculate the prediction error as:

$$\Delta = \hat{t} - t = \omega_g \underbrace{(\mu_g - t)}_{\Delta_t} + (1-\omega_g)\underbrace{(s^* - t)}_{\Delta_{s^*}} + (1-\omega_g)\underbrace{\lambda(s - s^*)}_{\Delta_s} + \epsilon \tag{3}$$

Equation 3 is sheds light on the sources of prediction error that the model incorporates. First, Δ_t captures how different is the true type t from the prior μ_g, adjusted by the inference weight on the prior (ω_g). Δ_{s^*} reflects the difference between the correctly backed up signal and the true type due to sampling noise, and adjusted by the weight put on the signal $(1-\omega_g)$. Finally, Δ_s is the gap between the selected data (i.e., perceived signal when $\lambda = 1$) and the corrected signal (i.e., the perceived signal when $\lambda = 0$). The loading of this term into the prediction error depends on the level of selection neglect λ: when selection is perfectly accounted for ($\lambda = 0$), this gap has no effect. Finally, the model includes a standard computational error ϵ which we assume to have mean zero.

Statistical Discrimination

We define statistical discrimination SD as the ratio of the expected crime predictions across two groups A and B:

$$SD_{AB} = \frac{E[\hat{t}_A]}{E[\hat{t}_B]} \tag{4}$$

Intuitively, there is statistical discrimination when $SD \neq 1$, this is, one group is predicted to have a higher expected criminality than the other. However, this can be accurate if the crime distributions are different across groups. Thus, we decompose statistical discrimination into an accurate (aSD) and inaccurate component statistical discrimination (iSD): $SD = aSD \cdot iSD$. We consider as accurate statistical discrimination to the component of the difference between the predicted crime rates across groups that aligns to the actual crime distributions: $aSD_{AB} = \frac{E[t_A]}{E[t_B]} = \frac{\mu_A}{\mu_B}$. On the other hand, the inaccurate component is the part of the prediction that doesn't follow the actual crime distribution, and can be expressed, using Equation 2, as:

$$iSD_{AB} = \frac{\omega_A + (1-\omega_A)(1 + \lambda\left(\frac{E[c(p_A, t_A)]}{\mu_A} - 1\right))}{\omega_B + (1-\omega_B)(1 + \lambda\left(\frac{E[c(p_B, t_B)]}{\mu_B} - 1\right))} \tag{5}$$

Intuitively, crime predictions for group A will be relatively overestimated when the expected crime signal, given the level of policing, is higher than the expected crime level of the group. This holds if and only if the decision maker neglects selection to some extent ($\lambda > 0$) and doesn't disregard signals when updating ($\omega_A < 1$). Conversely, there will be no inaccuracies when all the weight is put on the prior or when the subject perfectly accounts that data is selected ($\lambda = 0$) and therefore is able to back up the true crime rate, or when .

Mechanisms

Understanding the mechanisms driving selection neglect in police decision-making is as crucial as evaluating its presence. By identifying the causes of this bias, we can better design policy interventions to mitigate its effects. To this end, we propose several mechanisms that we will test experimentally, as detailed in Section 3.

Evidence suggests that in environments with large information sets, agents allocate scarce cognitive resources by focusing only on a subset of available information while disregarding the rest (Bordalo et al., 2012; Gabaix & Laibson, 2005; Simon, 1955; Stigler, 1961). Moreover, and the tendency to overlook information increases with the complexity of the problem. In other words, excessive information can be counterproductive—when information becomes too complex, individuals struggle to discern which features are relevant. This phenomenon has been observed in areas such as technology adoption in farming (Hanna et al., 2012) and taxation (Chetty et al., 2009). Importantly, these studies highlight that when relevant information is made salient, agents can better select information to improve their decisions.

The saliency and complexity of information are directly relevant to crime prediction. Crime forecasting is a complex problem involving numerous pieces of information.[7] In such a setting, decision-makers may struggle to determine which pieces of information are most useful. Specifically, since law enforcement officers rely on a broad set of information to predict crime levels, they may overlook the role of police presence in a given area and focus solely on reported crime. This, in turn, increases the likelihood of selection neglect, as officers may disregard the impact of local police presence on observed crime rates.

We are also interested in studying the role of social and observational learning. This idea has been formalized in Banerjee (1992). Intuitively, this model captures the idea that for a new decision-maker (such as a rookie police officer), it is rational to imitate senior officers because they may possess information that the new officer lacks. However, this can have long-term consequences: if new decision-makers rely on imitation rather than independent learning, it can lead to the perpetuation of incorrect decision-making (Bikhchandani et al., 2024). A growing body of applied literature examines this phenomenon in various contexts, demonstrating that individuals (Cai et al., 2009; Falk & Ichino, 2006) as well as firms (DellaVigna & Gentzkow, 2019) are both susceptible to such behavior.

In the context of policing decisions, social learning may play a substantial role. Crime prediction is a complex task that requires integrating various types of information. Given this complexity, it may be strategic for an officer to imitate what other officers do. Consequently, if experienced officers neglect data selection when predicting crime[8] it becomes more likely

[7]For example, decision-makers may be presented with local demographic or income characteristics, further increasing the complexity of the task.

[8]A direct example from our experimental design would be officers using arrest data to infer crime levels

that newer officers will also neglect data selection. This dynamic suggests that small shifts in the behavior of a subset of officers could scale up, leading to widespread effects within the police force.

Finally, selection neglect can be interpreted as a model misspecification problem (Bohren & Hauser, 2024; Esponda & Pouzo, 2016). Specifically, in our context, decision-makers may have access to all relevant information but fail to infer which information is truly relevant for updating their beliefs. This selective attention process can lead to persistent errors in prediction if decision-makers cannot learn which information is useful (Schwartzstein, 2014). For example, an officer deciding how to allocate patrols based on local crime rates may fail to recognize that observed arrests depend on police presence and are not directly mapped into local crime levels. This misinterpretation is likely to persist if, as in our context, the true state (i.e., actual crime rates) is never revealed. Since patrol allocation officers have limited opportunities to update their beliefs, they may infer incorrect differences in crime rates across neighborhoods—not because they neglect the information per se, but because they fail to correctly process how information should be used to extract meaningful insights from observed data.

2.2 Selection Neglect in Exploration Dynamics

Selection neglect can distort the inference process of a decision maker that observes selected signals, as shown in the previous section. These distortions can compound in dynamic settings where future data collection depends on current (potentially distorted) predictions. In this section, we expand the previous model to these dynamic settings using a multi-arm bandits framework. A decision maker faces several arms (slot machines), and each arm has a reward probability unknown to the decision maker. At each decision, the decision maker chooses which arms to pull arm, receiving from each arm pulled a reward which is a Bernoulli random variable. At each trial, the agent has to trade-off two goals: exploitation and exploration. Since the agent holds some beliefs over the reward probabilities, it might seem optimal to *exploit* this knowledge by pulling the arm with higher expected reward. However, there is informational value in *exploring* by pulling different arms, as they might find out another arm

while neglecting differences in police presence across neighborhoods.

is actually more profitable. Selection neglect can affect both goals. Consider the extreme case where an agent doesn't realize that current beliefs depend on previous pulling decisions and interprets the lack of reward from non-pulled arms as no-reward signals. In this case, neglecting the selection of previous data leads to distorted beliefs that will affect which arm is chosen for *exploitation*. On the other hand, the agent can neglect the informational value of selecting data in a diversified way, thus underestimating the value of *exploration*. It's intuitive to see how these two effects can compound. If one arm has been pulled more often than others, resulting in more rewards in absolute terms, an agent that neglects selection will keep exploiting this arm (as they underestimate the expected value of other arms) and not exploring others (as they undervalue exploration).

It's straightforward to apply this framework to the context of crime prediction and policing decisions. Different neighborhoods have different crime probabilities, and the police decision maker is rewarded for catching these crimes. In a context of scarce resources (not all neighborhoods can be exhaustively patrolled), where to police becomes a exploration-exploitation problem. Patrolling those neighborhoods with perceived high crime rates will result in a higher expected reward in the short term, whereas policing uniformly across areas will result in more informative crime data. If selection is neglected on the side of exploitation, officers will not take into account that the perceived higher crime rate in an overpoliced neighborhood might be the result of excessively sampling data there, leading to a higher value of exploitation in that area. If selection is neglected on the side of exploration, officers will discount the informational value of policing traditionally underpoliced areas, reducing the incentive to divest resources from overpoliced ones. Through both effects, it's easy to see how selection neglect might lead to the empirical patterns of statistical discrimination and overpolicing we observe in policing decisions (Chen et al., 2023).

What remains of this section formalizes this intuition. Multi-arm bandit problems have been extensively studied, and that is because their closed form solution are famously difficult to find. In economics, previous has used bandits to model problems of strategic experimentation (Keller et al., 2005), research decisions (Zhuo, 2023), psychiatric treatment prescriptions (Currie & MacLeod, 2020), or product selection by retailers (Jin et al., 2021). While the theory literature derives equilibrium predictions, most empirical applications acknowledge the intractability of bandit problems and opt to model decision makers that use some sort of heuristics to ease computations. Following the literature and anticipating our experimental design, we define the objective function of our agent, explain why the curse of dimensionality makes the problem intractable, and propose two different heuristics that approximate optimality at much reduced computational costs.

2.2.1 Objective Function

A decision maker faces N arms they can pull. We denote the choice a_{it} to pull arm i at time t as $a_{it} = \{0, 1\}$. As at time t they can pull $n_t \leq N$ of them, the agent faces a resource constraint $\sum_i^N a_{it} = n_t \forall\, t$. We denote by $\mathbf{a}_t$ the vector of actions at time t. The outcome of a pull can be successful and give a reward $R_{it} = 1$, or unsuccessful and give no reward $R_{it} = 0$. If an arm is not pulled, it gives no reward. At time t, the decision maker has an information set I_t that comprises all past actions $\mathbf{a}_1, \mathbf{a}_2, ...\mathbf{a}_{t-1}$ and outcomes $\mathbf{R}_1, \mathbf{R}_2, ...\mathbf{R}_{t-1}$.

When pulled, the probability of reward is p_{it}, which the agent has incomplete information about. Let p_{it} have a prior distribution $F_t(p_{it})$. Then, after observing the history of actions and outcomes, the decision maker has a posterior belief of the probability $\tilde{F}_t(p_{it}|I_t)$. In the simplest case where the reward probabilities are constant and arms are independent from each other, a Bayesian agent would hold a belief $\tilde{F}$ that follows a Beta distribution [9]. In case of receiving a reward, we allow these rewards to vary across arms. We denote the utility payoff from an arm as $\pi_{it}(a_{it}, p_{it}; \omega)$, which depends on whether the arm has been pulled ($\pi_{it}(a_{it} = 0, p_{it}; \omega_i = 0)$ and an arm-specific weight ω_i. Thus, our agent maximizes the posterior expected payoff:

$$\max_{\mathbf{a_1}(I_1),...\mathbf{a_T}(I_T)} \sum_t^T \sum_i^N \int \pi_{it}(a_{it}, p_{it}; \omega)\; d\tilde{F}_t(p_{it}|I_t), \text{ subject to} \sum_i^N a_{it} = n_t \forall\, t \tag{6}$$

In the theoretical literature, it's often assumed that agents find the optimal action sequence in bandits with finite temporal horizons by backwards induction. The first step of this approach is setting up a Bellman equation, which takes the information set as a state variable and the action vector as a control. In our case, we can define the dynamic programming

[9] Where the posterior belief is $P(p_i|I_t) = f(p_i;\ \alpha_0 + \sum_t R_{it},\ \beta_0 + \sum_t (a_{it} - R_{it}))$ of a Beta distribution.

problem throught the value function:

$$V_t(I_t) = \max_{\mathbf{a}_t} \underbrace{\sum_i^N \int \pi_{it}(a_{it}, p_{it}; \omega)\ d\tilde{F}_t(p_{it}|I_t)}_{\text{Exploitation component}} + \underbrace{E_{I'_{t+1}}\left[V_{t+1}(I'_{t+1})|I_t, \mathbf{a}_t\right]}_{\text{Exploration component}} \tag{7}$$

We can see the Bellman equation has two additively separable components. The first captures the expected value of *exploiting* the information currently available, while the second captures the continuation value: the expected change in the value function due to *exploring* and acquiring extra information. The value of exploration, this is, the continuation value, is obtained by integrating over all possible future states conditional on the current state I_t. Note that I_t comprises all past actions and outcomes. As the number of possible states I_T at the end of the time horizon becomes excessively large, the curse of dimensionality makes solving the Bellman equation by backward induction intractable. As the curse of dimensionality is a common problem in bandits, previous research has characterized several workarounds to approximate the optimal policy at a feasible computational cost. We focus on two of these approximations: index approximations and reinforcement learning.

2.2.2 Index Approximation

The curse of dimensionality arises from the large number of possible future states one has to integrate over to find the continuation value in the Bellman equation (7). To solve this issue, index solutions approximate the continuation value without integrating over all possible states, and thus optimize using an approximated value function $\hat{V}_{it}$, that takes the functional form:

$$\hat{V}_{it}(I_t, \mathbf{a}_t; \omega, \theta) = \underbrace{\int \pi_{it}(a_{it}, p_{it}; \omega)\ d\tilde{F}_t(p_{it}|I_t)}_{\text{Same exploitation component}} + \underbrace{B_{it}(I_t, \mathbf{a}_t; \theta)}_{\text{Exploration bonus}} \tag{8}$$

Similar to the Bellman equation, the approximated value function has two additively-separable components. The first is identical to the Bellman equation: the agent still calculates an expected payoff from pulling each of the arms. The second part is the function $B_{it}(I_t, \mathbf{a}_t; \theta)$, which is the approximation to the continuation value, as is known in the theoretical literature as "exploration bonus". This approximation doesn't integrate over all future states, and it only takes as inputs the information directly available to the agent. Thus, the optimal solution is found by calculating the index $\hat{V}_{it}$ for every arm, and then pulling the n_t arms with a highest index. As no backward induction is needed, this approximation avoids the dimensionality curse. While several index solutions are available and many functional forms have been proposed for $B_{it}(\cdot)$, we choose an Upper-Confidence Bounds (UCB) approach [10], where the exploration bonus takes the following form:

$$B_{it}(I_t, \mathbf{a}_t; \theta) = a_{it}\sqrt{\frac{\theta}{\sum_1^t a_{it}}} \tag{9}$$

The exploration bonus depends on whether an arm is pulled: if it isn't pulled, no new information will be revealed. The parameter θ captures the agent's preference for exploration: when $\theta = 0$, the agent places no value in the new information acquired from exploration. Finally, $B_{it}(\cdot)$ is a convex decreasing function in the number of pulls for arm i, $\sum_1^t a_{it}$. Intuitively, $B_{it}(\cdot)$ approximates the upper confidence bound of the expected posterior reward: if an arm has not been much explored, there distribution of the expected reward has a high variance. As an arm is pulled many times and more information is acquired, this variance decreases, and so does the upper confidence bound of the expected payoff. In other words, the value of exploration decreases with the amount of exploration already done in that arm.

Why using an UCB index and not other indices common in the theoretical literature? Index solutions begin when Gittins (1979) and Gittins and Jones (1979) proved that, in the standard bandit setting, the optimal policy could be reduced to calculating a dynamic allocation index for every arm and then pulling the arm with the highest index. However, this solution — commonly known as the Gittins Index — is only optimal for the standard setting, which varies from the objective function described in Equation (6) in several ways. First, the standard bandit models a setting where reward probabilities don't change over time. In our case, $\tilde{F}_t(p_{it}|I_t)$ can change over time, reflecting changes in criminality over time. For these non-stationary bandits (called "restless bandits"), the Gittins index is suboptimal, but other index approximations like the Whittle index or UCB are more robust (Whittle,

[10]We modify the standard UCB index proposed by Auer et al. (2002) by having a constant and individual-specific preference for exploration θ. Using a constant parameter for this preference is more common in current applications of UCB indices (Lattimore & Szepesvári, 2020).

1983). Second, in the standard bandit only one arm is pulled at each time period, while our setting allows for several arms to be pulled — several neighborhoods can be patrolled at the same time. UCB approximation is robust in this case, known as "combinatiorial" bandits. Third, the standard setting has uncorrelated arms, while in our case we allow for correlation between arms by making the reward probability depend on the entire information set, this is, all previous actions and outcomes for all arms. For these "contextual" bandits, UCB-like indices also achieve optimality at lower costs, while Gittins indices are suboptimal [11]. In addition, Gittins indices are famously difficult to compute (Chakravorty & Mahajan, 2014), while UCB indices are more intuitive. For this reason, they are the most common in empirical applications (Jin et al., 2021; Zhuo, 2023). For all these reasons, an UCB index like the one described in Equation (9) is a better choice than other indices commonly used in the theoretical literature.

2.2.3 Reinforcement Learning Approximation

Also, Reinforcement Learning literature (Lefebvre et al., 2017; Palminteri & Lebreton, 2022), easier to compute and

, VALUE OF INFO (Ambuehl & Li, 2018; Anderson, 2012; Charness et al., 2021; Ichihashi, 2025)

3 Experimental Design

3.1 The Inference Task

Abstract Version

We design a simple ball-and-urn experiment to test the predictions of our theoretical framework. Participants learn there are two groups of boxes with 20 balls inside, each of these balls is either white or black. Boxes have different proportions of white and black balls, with one of the groups of boxes having a higher number of black balls than the other

[11] For an exhaustive and detailed discussion of different index approximations, see Lattimore and Szepesvári (2020).

on average. Participants learn the prior distribution of black balls across the two groups of boxes.

The data generating process is simple. A box from each group is selected at random, and some number of balls are independently sampled from each of them with replacement. We let the number of draws to vary between 10 and 30, following a XXXXX distribution. For each box, participants observe its group, the number of black balls sampled, and the number of draws, and then guess the number of black balls in each of the boxes. Accurate guesses are incentivized.

Framed Version

The decision maker in our model infers the type of individuals who belong to different groups, after observing a noisy signal of their type. As our focus is on statistical discrimination, separate from taste-based discrimination, we present participants with individuals belonging to *abstract* groups. However, even with abstract groups, there's a risk of participants interpreting these groups as genders or ethnicities and projecting their unobservable beliefs onto them (Mullainathan, 2002). To avoid this risk while keeping a similar environment to the crime prediction and policing decisions, we take individuals as two abstract groups as *neighborhoods* in two different *cities*: the Brown city and the Green city. Each neighborhood, as individuals in the model, has a different type, t, that in our context corresponds, for simplicity, to a measure of *crime level* that goes from 0 to 100. In particular, each neighborhood belongs to one of the two cities described above. Moreover, each city has its own continuum of neighborhoods with different crime levels. In particular, the distribution of crime rates of by neighborhood is identical in shape and dispersion across cities. However, we let mean crime level of the neighborhood of one city is to be larger than the other (see Figure A1 for an illustration). Thus, Brown neighborhoods (i.e., neighborhoods from Brown city) have an average type of μ_B, while Green neighborhoods have an average type of μ_G, with $\mu_B > \mu_G$. Both crime level distributions have a standard deviation of σ.

Before starting the experiment, participants learn about the distribution of crime levels across the two cities. While crime distributions are not actually observed in real settings, we fix prior beliefs to leave no place for representative prior distortion or stereotypes and to isolate the effect of selection neglect as a misperception of the signal [12].

Participants then learn about the data-generating process. In our experiment, we inform them that these neighborhoods are monitored by surveying cameras covering all streets. The recorded footage is sent to both the national and local police forces, where analysts review it to report observed crimes. However, each police force lacks sufficient analysts to review 100% of the footage.

We explain that each analyst is assigned a random subset of locations or cameras to observe, ensuring a representative understanding of crime distribution across the neighborhood. This clarification is crucial to prevent participants from assuming that footage from high-crime areas is more likely to be analyzed. Additionally, we emphasize that the two police forces operate independently, with no coordination between them. Consequently, the signal participants receive is the sum of arrests made by both forces. Mathematically, this corresponds to the sum of two independent random samples of crimes drawn without replacement.

Let p_{ng}^L and p_{ng}^N be the number of analysts that review the footage of neighborhood n of city g at the local and national level, respectively. We tell the participants that each of the analysts is able to observe 5% of the footage. We let the reported crime to have an idiosyncratic mean-zero reporting error. This can be rationalized by the fact that there are some crimes that the analysts do not see, and there are other situations that are reported as crime that actually are not. In this case, the signal observed by police force F with A_F analysts will be equal to c^F. The decision maker can then back up the correct (i.e., the neighborhood-level) signal as

$$s^{*F} = \frac{100}{p_{ng}^F \cdot 5} c^F$$

In our experiment, participants do not see these force-specific signals. Instead, we present them with the total reported crimes by the analysts by both police forces. Namely, the presented signal corresponds to:

$$s = s^L + s^N$$

Now, let $p_{ng} = p_{ng}^L + p_{ng}^N$, then the correct backed up signal from the perspective of the

[12] Appendix A1 includes how the model would change if representative prior distortion was added. We decide not to include it so as to focus on the effect of selection neglect.

individual should be

$$s^* = \frac{100}{p_{ng} \cdot 5} \cdot s$$

In each round of the task, the computer randomly draws one neighborhood from each of the two cities, each with their own crime level t_{ng}. For each neighborhood, the computer selects a level of total analysts p_{ng} and draws a signal s_{ng}[13]. Participants then observe: the prior distributions, the number of analysts and the signals. After being presented with this information, they make an incentivized guess[14] of the true crime level of each neighborhood.

Importantly, the experimental interface provides participants with a calculator to help them make their guess. Participants decide whether they want to use the calculator or not; we record their calculations. We include a calculator for two reasons. First, we want to isolate the effect of selection neglect from numeracy skills and calculation errors. Second, recording participants' calculations offers valuable insights about the decision making process or model that participants use for inference — the choice process, beyond the outcome of that inference — the choice. For instance, if a participant makes calculations using the signals but without adjusting them to the selection level, the recorded calculation is direct evidence of neglecting selection. This use of the calculator adds to the recent literature using novel choice process tools, like notes in Fréchette et al. (2024) or custom price calculators in Martınez-Marquina and Rasocha (pre-published).

To situate our task within the experimental literature, our participants observe a signal whose noise reflects sampling uncertainty, rather than perception noise as in Bordalo et al. (2016) and Esponda et al. (2023) (see Figure **??** for an illustration). Our experiment is informationally equivalent to a ball-and-urn standard setting. We can consider the two cities to be two different groups of urns, where each urn is a neighborhood. In turn, each urn (neighborhood) has the same number of balls where each ball contains a number (crime observed by an analyst), and these numbers correspond to the type of the urn (crime level). A different number of balls is sampled from each urn (total number of analysts), which determines the

[13]The signal is constructed using $s_{ng} = \frac{p_{ng}}{100} \cdot t_{ng} + \epsilon_{ng}$. Therefore, the computer only has to get a random draw for ϵ_{ng} and then the signal is constructed directly.

[14]We pay participants \$ $\frac{2}{|\hat{t}-t|}$ in each round.

observed signal. Instead of intentionally hiding some ball color — as in López-Pérez et al. (2022), or the high/low draws — as in Enke (2020), selection in our experiment acts through the level of sampling in each neighborhood. This selection mechanism reflects the real setting, where some neighborhoods are over or under surveyed, and thus more or less crime data is sampled from them. Finally, by grouping neighborhoods in two cities, we can explore whether selection neglect varies when signals confirm or contradict the prior, relating selection neglect to the literature on confirmation boas (Benjamin, 2019).

Treatment Design

Inference Task is repeated in 2 blocks of 10 rounds each. Across blocks, participants complete the Inference Task followed by the Patrol Allocation Task in each round. In Block 1, the level of selection in the Inference Task is exogenously selected. In particular, the number of analysts in each neighborhood is uniformly drawn among the set of integers between 5 and 15 (recall that each analyst observes 5% of the footage) under some constraints to be able to test fundamental predictions of the model: (1) a minimum number of draws of $p_{ng} = 100$ per city, to test no inaccurate statistical discrimination arises in the absence of selection, (2) there are some rounds where both neighborhoods face the same level of policing, and (3) there are enough signals that contradict the priors.

In Block 2, participants are randomly assigned to 4 treatment groups for the last 10 rounds. The treatments are designed to test some potential mechanisms causing selection neglect (see Section 2.1) as well as testing some interventions to ameliorate the bias it generates. The treatments considered are the following:

1. **Control**. Participants complete 10 additional rounds with exogenous selection.

2. **Salience**. To test whether this mechanism is driving selection neglect, we experimentally reduce the salience of the selection information (number of analysts) by including decoy attributes of each neighborhood that are irrelevant for inference. We expect participants in the treatment group to reveal a higher selection neglect (i.e., a higher λ).

3. **Peer Effects**. We show participants how other participants made crime predictions and allocated patrols. In particular, we show half of the participants in this treatment

group a case where selection is neglected, and the other half a case where selection is accounted for. If participants are vulnerable to peer effects, then the participants treated in the former group will neglect the data more than those in the latter.

4. **Top-down misspecification**. We want to evaluate whether model misspecification is one of the causes of why selection neglect arises. To test this hypothesis, we provide participants with a short training on selection neglect and the cycle of bias it can generate, specifying what model should be used to account for data selection.

Subset Analysis

Without any additional treatment arms or variations in the experiments, we aim to study other hypotheses by partitioning our data in several ways. In particular:

1. **No selection, less distortion**. When the selected signal is equal to the backed-up one ($s = s^*$), even participants that neglect data selection should show less prediction error. Equation 3 shows this prediction clearly: when $s = s^*$, $\Delta_s = 0$, reducing the overall prediction error Δ across the board. We validate this prediction by restricting the sample to those draws where the level of policing leads to undistorted signals, this is, $p_{ng} = 20$.[15]

2. **Unbiased selection, no spurious discrimination**. A key prediction of the model is that in the absence of differential selection across groups, selection neglect will not generate inaccurate statistical discrimination. In other words, if the number of analysts is the same across neighborhoods, there is no reason for a participant to form inaccurate discrimination towards one group. To test this prediction, we restrict the sample to the subset of observations where both neighborhoods have an equal level of policing. While selection neglect can still bias each prediction, it should bias in the same way the predictions for neighborhoods in both cities, not amplifying discrimination across cities.

3. **Confirmation Bias**. Previous literature has shown people are more likely to trust information that confirms their priors, and update more following confirmatory signals

[15] Recall that you need 20 analysts in a police force to be able to survey the footage of the whole neighborhood and rule out selection bias.

(Benjamin, 2019; Charness & Dave, 2017; Lord et al., 1979). In our experiment, participants have prior beliefs about the crime levels of each city given by the known crime distributions across neighborhoods. Suppose that, $\mu_G > \mu_B$, then, before observing any signals, participants should predict that $t_{nG} > t_{nB}$. Similar to confirmation bias, selection neglect might be a motivated cognition tool to protect the ego from wrong priors (Möbius et al., 2022). If this is the case, participants might be more likely to recur to selection to explain why the signals contradict their priors ($s_{nG} < s_{nB}$), but neglect selection when the signals confirm those priors ($s_{nG} > s_{nB}$). To test this hypothesis, we divide the sample between those observations confirming the prior and those contradicting the priors. We expect to estimate a lower λ among those rounds where the priors are confirmed, relative to when the signal contradicts them.

3.2 The Exploration-Exploitation Task

Framed Version

In the Inference Task, we separate crime predictions from policing decisions by making the level of policing, and thus the crime signals, exogenous to participants' predictions. While this is necessary to isolate the consequences of neglecting the past selection of data in current predictions, the potential consequences of selection are greater in dynamic settings where signals, predictions, and decisions are all linked. The Exploration-Exploitation Task models this setting. Participants observe a grid of neighborhoods, and in each neighborhood a crime can happen or not happen — crime follows a Bernouilli process. Each neighborhood has an unobserved criminality level that reflects the probability of crime happening there. In addition, each neighborhood can be patrolled or not. If a neighborhood is patrolled and crime happens there, crime is caught with some probability. If a neighborhood is not patrolled, crime is not caught. Participants observe what neighborhoods are patrolled, and whether crime is caught or not in each of them.

We divide this part of the experiment in 3 blocks, with each block being a variation of the Exploration-Exploitation Task. In Block 1, all neighborhoods are patrolled. After observing the different crime signals (whether crime has been caught or not), participants rank the neighborhoods according to their perceived level of criminality, and are paid according to their accuracy. This prediction is orthogonal to the patrol allocation, as in every trial of this block all neighborhoods are patrolled no matter the crime prediction. In Block 2, only one neighborhood is randomly chosen to be patrolled. After observing which neighborhood is patrolled and whether crime has been caught or not in each neighborhood, participants rank them according to their predicted crime, and correct guesses are incentivized.

Why having a first block with all neighborhoods patrolled and then a block with only one? This design allows us to test whether participants neglect previous data selection by comparing both blocks. In this context, we can characterize selection neglect as considering the lack of signals from unpatrolled neighborhoods (no crime can be caught there, by design) as signals of no crime. In the first block, signals are unselected, so we can characterize how each individual learns from positive (crime) and negative (no crime) signals. In the second block, each neighborhood can have three possible outcomes: positive signal, negative signal or none. If an individual makes a similar downward update after receiving a negative signal than after receiving no signal, this individual is exhibiting selection neglect.

In Block 3, the task is closest to the real setting. In each trial, participants now choose which neighborhood to patrol. As a result, they will observe whether crime was caught there or not, while no crime is caught in the unpatrolled neighborhoods. After observing this information, participants decide again where to patrol. Participants are paid by the number of times they manage to catch crime due to their allocation. The experimental design of Block 3 is akin to a multi-arm bandit experiment, where choosing each arm leads to receiving a reward with some unknown probability. This is the standard setting to capture the trade-off between exploration and exploitation.

In addition to the multi-arm bandit, we elicit the willingness to pay for unbiased information during Block 3. In the context of policing, victimization surveys offer a population-based alternative to crime data generated by the criminal justice system that is more robust to endogenous selection issues (Lum & Isaac, 2016). However, these surveys are underutilized by crime analysts when predicting crime and making policing decisions (Brayne, 2020). To understand how participants value unbiased information, during Block 3 we offer them the opportunity to buy access to a victimization survey that reveals, with some noise, the crime probabilities of each neighborhood. While previous literature has shown that people systematically undervalue the informativeness of information structures (Ambuehl & Li, 2018; Charness et al., 2021), in the context of policing several mechanisms could be driving this undervaluation. First, there could be a data endowment effect, where police officers develop an attachment to the data they have generated and devalue data they have not produced themselves. Second, an inaccurate cost-benefit analysis of the expected value of using unbiased versus selected data. To disentangle these mechanisms, we randomize when participants receive the offer —either during trial 15 or trial 25, and whether they receive it alongside a cost-benefit analysis of the expected gains from using the survey relative to the data they have.

Abstract Version

The abstract version follows the exact same design, but instead of a spatial policing framing we use a classic multi-arm bandit. Instead of neighborhoods, participants are presented with abstract symbols — the procedure most validated in the cognitive science and neuroeconomics literature (see Palminteri and Lebreton (2022) for examples). Each symbol is associated with some probability of reward, and participants explore and learn these probabilities by choosing one of the symbols in each trial.

The block design in the abstract version is identical to the framed version. In Block 1, in each trial participants receive a signal from each symbol, and rank their beliefs about which symbol is more likely to give a reward. In Block 2, participants only observe the signal of one randomly chosen symbol, and then rank their beliefs. In Block 3, participants choose one symbol in each trial, and earn the realized reward from the chosen symbol. During Block 3, participants report their willingness to pay for unbiased information about the reward probability of each symbol, following the same timing and cost-benefit treatments as in the framed version.

3.3 Implicit Discrimination

One of the main theoretical implications of selection neglect is that it can lead to dynamics of statistical discrimination against overpoliced groups (see Section 2). To test whether this prediction extends beyond the statistical the abstract setting of the Inference and the Patrol Allocation Tasks, we conduct a modified Implicit Association Test (IAT) to measure discrimination. The idea behind the IAT is that humans are able to process belief-congruent information faster than incongruent information. The standard IAT (Greenwald et al., 1998) makes participants classify names or pictures into two target groups (e.g., White and Black people), and adjectives into two attribute groups (e.g., positive and negative sentiment). The classification is done by pressing either the E or I key of a keyboard, with each target and attribute group being classified on one side. The task of classification is repeated many times, for several blocks. In some blocks, one target group has to be classified on the same side as the positive attributes, and in another blocks that target group is classified on the same side as the negative attributes. If there is an implicit association between a target group and positive sentiments (e.g., White people and positive sentiments), response times and error rates will be lower when both are classified on the same side (congruent trials) than when that target group is classified on the same side as negative sentiments (incongruent trials). This test has shown ample external validity (Greenwald et al., 2009), and predicts several relevant economic outcomes (Alesina et al., 2024; Glover et al., 2017). We modify the standard IAT to capture implicit associations of criminality to different racial groups that are relevant to the context where our experiment will take place (Latin America). In this modified IAT, participants classify acts into criminal or not, and pictures of Indigenous and White individuals.

3.4 Numeracy

We explore whether numeracy skills and cognitive reflection are associated with selection neglect and statistical discrimination. We include two questions to test numeracy skills (Kahan et al., 2012), one that test cross-multiplication, and one to measure cognitive reflection (Frederick, 2005). The questions are available in the Appendix A2. A measure of numeracy will allow to test whether the results are due to poor understanding of probabilities rather than a neglect of selection.

4 Empirical Analysis

The main goal of our experiment is to identify the individual parameter of selection neglect, λ. It is important to keep in mind that the econometritian observes the city prior distributions (from where she can infer μ_g), the signals and the correctly-backed up signal s and s^*, and the true and estimated crime level for a given neighborhood t_{ng} and $\hat{t}_{ng}$. It is important to keep in mind that, contrarily to the econometritian, participants only observe the prior distributions, the signal of crime observed s, and the total analysts $p_{ng} = p^L_{ng} + p^N_{ng}$, and they are asked to guess $\hat{t}_{ng}$. With this information, we can also back up $s*$.

In each round, a different neighborhood n is drawn from the distribution for each city g. Given that participants play different rounds of our experiment, we are able to get within subject variation, which allows us to estimate the selection neglect parameter at the individual levels. Note that we can substract s^* from each side of Equation 2 to get the following expression

$$\hat{t} - s^* = \omega_g(\mu_g - s^*) + \underbrace{(1-\omega_g)\lambda}_{\gamma_1}(s - s^*) + \epsilon \tag{10}$$

We are inducing random variation in both p_{ng} and t and therefore on $(\mu_g - s^*)$ and $(s - s^*)$[16], such that we are able get consistent estimates of ω_g and γ_1. Then, it is easy to show that $\hat{\lambda} = \frac{\hat{\gamma_1}}{1-\omega_g}$. It is natural then to use the Delta method to do inference about the structural parameters.

We acknowledge that our empirical strategy cannot separately identify round-specific ω_g. Therefore, we follow the approach adopted by Esponda et al. (2024) in a similar context and estimate only group-specific weights. Recall that this parameter captures the weight a participant assigns to the signal when forming their beliefs. This presents a challenge in cases where, for example, the number of analysts in each police force is sufficiently large to observe almost all of the footage from each neighborhood. In such cases, the variance of the signal would be close to zero Consequently, the decision-maker should assign a significantly higher weight to the signal in rounds where the number of analysts is large compared to rounds where it is small.

[16] Which are at the same time linearly independent (i.e., rank condition holds).

To address this limitation, we take several precautions. First, we do not allow extreme patrolling scenarios, such as the one described above, ensuring that the decision-maker cannot easily infer significant differences in signal variance across rounds. Second, we do not disclose the exact breakdown of analysts between the two forces. This prevents participants from accurately calculating the variance of the signal. For example, if the total number of analysts is 20, it could correspond to 10 analysts in each force (high variance) or 20 analysts in one force and 0 in the other (low variance). By withholding this information, we limit the extent to which participants can adjust the weight they assign to the signal based on differences in the number of analysts. Third, we experimentally test this assumption with a subsample to evaluate its validity.

To obtain these estimates, we use data from 15 rounds with endogenous and exogenous rounds alternate. The first 5 rounds are practice rounds for participants, while the last 5 rounds are used to estimate treatment effects.

For the treatment design, we adopt the same specification as in Equation 11, but we implement a *between-participants* rather than a *within-participants* estimation. This approach mitigates issues that could arise when comparing results from the last 5 rounds of the experiment to those from earlier stages, such as learning effects or fatigue. Consequently, following a similar logic as in the within-participant estimation, we recover group-specific parameters, where groups are defined based on the treatment groups outlined in Section 3.1.

Let i denote each participant in the pool I, let r denote the round of the experiment, and let T represent the set of treatment groups, such that T_i corresponds to the treatment group assigned to participant i. We then estimate the following regression:

$$\hat{t} - s^* = \sum_{t \in T} 1\{T_i = t\} \left(\omega_{t,g}(\mu_g - s^*)_{r,i} + (1 - \omega_{t,g})\lambda_t (s - s^*)_{r,i}\right) + \alpha_i + \epsilon \qquad (11)$$

Our parameters of interest are $\omega_{t,g}$ and λ_t. In this specification, we include participant fixed effects. We expect to find a lower $\lambda_{t,g}$ among the treatment groups relative to the control. Differences in the magnitude of this treatment-specific selection neglect parameter will help us determine which treatment is more effective in reducing selection neglect.

References

Alesina, A., Carlana, M., La Ferrara, E., & Pinotti, P. (2024). Revealing Stereotypes: Evidence from Immigrants in Schools. *American Economic Review*, *114*(7), 1916–1948.

Ambuehl, S., & Li, S. (2018). Belief updating and the demand for information. *Games and Economic Behavior*, *109*, 21–39.

Anderson, C. M. (2012). Ambiguity aversion in multi-armed bandit problems. *Theory and Decision*, *72*(1), 15–33.

Auer, P., Cesa-Bianchi, N., & Fischer, P. (2002). Finite-time Analysis of the Multiarmed Bandit Problem. *Machine Learning*, *47*(2), 235–256.

Ba, B., Knox, D., Mummolo, J., & Rivera, R. (2021). The role of officer race and gender in police-civilian interactions in Chicago. *Science*, *371*(6530), 696–702.

Banerjee, A. V. (1992). A simple model of herd behavior. *The quarterly journal of economics*, *107*(3), 797–817.

Barron, K., Huck, S., & Jehiel, P. (2024). Everyday Econometricians: Selection Neglect and Overoptimism When Learning from Others. *American Economic Journal: Microeconomics*, *16*(3), 162–198.

Benjamin, D. J. (2019, January 1). Chapter 2 - Errors in probabilistic reasoning and judgment biases. In B. D. Bernheim, S. DellaVigna, & D. Laibson (Eds.), *Handbook of Behavioral Economics: Applications and Foundations 1* (pp. 69–186, Vol. 2). North-Holland.

Bikhchandani, S., Hirshleifer, D., Tamuz, O., & Welch, I. (2024). Information cascades and social learning. *Journal of Economic Literature*, *62*(3), 1040–1093.

Bohren, J. A., & Hauser, D. N. (2024). Misspecified models in learning and games. *Working Paper*.

Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics*, *131*(4), 1753–1794.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly journal of economics*, *127*(3), 1243–1285.

Brayne, S. (2020, November 2). *Predict and Surveil: Data, Discretion, and the Future of Policing*. Oxford University Press.

Brundage, M., Little, A. T., & You, S. (2024). Selection Neglect and Political Beliefs. *Annual Review of Political Science*, *27*, 63–85.

Cai, H., Chen, Y., & Fang, H. (2009). Observational learning: Evidence from a randomized natural field experiment. *American Economic Review*, *99*(3), 864–882.

Chakravorty, J., & Mahajan, A. (2014). Multi-armed bandits, gittins index, and its calculation. *Methods and applications of statistics in clinical trials: Planning, analysis, and inferential methods*, *2*, 416–435.

Charness, G., & Dave, C. (2017). Confirmation bias with motivated beliefs. *Games and Economic Behavior*, *104*, 1–23.

Charness, G., Oprea, R., & Yuksel, S. (2021). How do People Choose Between Biased Information Sources? Evidence from a Laboratory Experiment. *Journal of the European Economic Association*, *19*(3), 1656–1691.

Che, Y.-K., Kim, J., & Mierendorff, K. (2024, September 12). *Predictive Enforcement*. arXiv: 2405.04764.

Chen, M. K., Christensen, K. L., John, E., Owens, E., & Zhuo, Y. (2023). Smartphone Data Reveal Neighborhood-Level Racial Disparities in Police Presence. *The Review of Economics and Statistics*, 1–29.

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *American economic review*, *99*(4), 1145–1177.

Cruces, G., Perez-Truglia, R., & Tetaz, M. (2013). Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics*, *98*, 100–112.

Currie, J. M., & MacLeod, W. B. (2020). Understanding Doctor Decision Making: The Case of Depression Treatment. *Econometrica*, *88*(3), 847–878.

DellaVigna, S., & Gentzkow, M. (2019). Uniform pricing in us retail chains. *The Quarterly Journal of Economics*, *134*(4), 2011–2084.

Dittmann, A. G., Dobson, K. S. H., & Schweitzer, S. (n.d.). How behavioral science interventions can disrupt the cycle of bias in AI-assisted police work. *Behavioral Science & Policy*.

Dube, O., MacArthur, S. J., & Shah, A. K. (2024). A Cognitive View of Policing. *The Quarterly Journal of Economics*, qjae039.

Edwards, F., Lee, H., & Esposito, M. (2019). Risk of being killed by police use of force in the United States by age, race–ethnicity, and sex. *Proceedings of the National Academy of Sciences*, *116*(34), 16793–16798.

Enke, B. (2020). What You See Is All There Is. *The Quarterly Journal of Economics*, *135*(3), 1363–1398.

Esponda, I., Oprea, R., & Yuksel, S. (2023). Seeing What is Representative. *The Quarterly Journal of Economics*, *138*(4), 2607–2657.

Esponda, I., & Pouzo, D. (2016). Berk–nash equilibrium: A framework for modeling agents with misspecified models. *Econometrica*, *84*(3), 1093–1130.

Esponda, I., Vespa, E., & Yuksel, S. (2024). Mental Models and Learning: The Case of Base-Rate Neglect. *American Economic Review*, *114*(3), 752–782.

Falk, A., & Ichino, A. (2006). Clean evidence on peer effects. *Journal of labor economics*, *24*(1), 39–57.

Fréchette, G. R., Vespa, E., & Yuksel, S. (2024, November 19). *Extracting Models From Data Sets: An Experiment*. 5026751.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25–42.

Fryer, R. G. (2019). An Empirical Analysis of Racial Differences in Police Use of Force. *Journal of Political Economy*, *127*(3), 1210–1261.

Gabaix, X., & Laibson, D. (2005). Bounded rationality and directed cognition. *Harvard University*, *20*.

Gittins, J. C. (1979). Bandit Processes and Dynamic Allocation Indices. *Journal of the Royal Statistical Society: Series B (Methodological)*, *41*(2), 148–164.

Gittins, J. C., & Jones, D. M. (1979). A dynamic allocation index for the discounted multi-armed bandit problem. *Biometrika*, *66*(3), 561–565.

Glover, D., Pallais, A., & Pariente, W. (2017). Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores. *The Quarterly Journal of Economics*, *132*(3), 1219–1260.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17–41.

Hanna, R., Mullainathan, S., & Schwartzstein, J. (2012). Learning through noticing: Theory and experimental evidence in farming. *National Bureau of Economic Research.*

Hübert, R., & Little, A. T. (2023). A Behavioural Theory of Discrimination in Policing. *The Economic Journal*, *133*(655), 2828–2843.

Ichihashi, S. (2025). Information and policing. *Journal of Economic Theory*, *225*, 105992.

Jehiel, P. (2018). Investment Strategy and Selection Bias: An Equilibrium Perspective on Overoptimism. *American Economic Review*, *108*(6), 1582–1597.

Jin, R., Simchi-Levi, D., Wang, L., Wang, X., & Yang, S. (2021). Shrinking the Upper Confidence Bound: A Dynamic Product Selection Problem for Urban Warehouses. *Management Science*, *67*(8), 4756–4771.

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, *2*(10), 732–735.

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

Lattimore, T., & Szepesvári, C. (2020, September). *Bandit Algorithms* (1st edition). Cambridge University Press.

Lefebvre, G., Lebreton, M., Meyniel, F., Bourgeois-Gironde, S., & Palminteri, S. (2017). Behavioural and neural characterization of optimistic reinforcement learning. *Nature Human Behaviour*, *1*(4), 1–9.

López-Pérez, R., Pintér, Á., & Sánchez-Mangas, R. (2022). Some conditions (not) affecting selection neglect: Evidence from the lab. *Journal of Economic Behavior & Organization*, *195*, 140–157.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence. *Journal of Personality and Social Psychology*, *37*(11), 2098–2109.

Lum, K., & Isaac, W. (2016). To Predict and Serve? *Significance*, *13*(5), 14–19.

Martınez-Marquina, A., & Rasocha, V. (pre-published). *Like There is No Tomorrow.*

Möbius, M. M., Niederle, M., Niehaus, P., & Rosenblat, T. S. (2022). Managing Self-Confidence: Theory and Experimental Evidence. *Management Science*, *68*(11), 7793–7817.

Mohler, G. O., Short, M. B., Malinowski, S., Johnson, M., Tita, G. E., Bertozzi, A. L., & Brantingham, P. J. (2015). Randomized Controlled Field Trials of Predictive Policing. *Journal of the American Statistical Association*, *110*(512), 1399–1411.

Mullainathan, S. (2002). A memory-based model of bounded rationality. *The Quarterly Journal of Economics*, *117*(3), 735–774.

Palminteri, S., & Lebreton, M. (2022). The computational roots of positivity and confirmation biases in reinforcement learning. *Trends in Cognitive Sciences*, *26*(7), 607–621.

Pierson, E., Simoiu, C., Overgoor, J., Corbett-Davies, S., Jenson, D., Shoemaker, A., Ramachandran, V., Barghouty, P., Phillips, C., Shroff, R., & Goel, S. (2020). A large-scale analysis of racial disparities in police stops across the United States. *Nature Human Behaviour*, *4*(7), 736–745.

Rizzotti, L. (2024). *Racial Discrimination against Crime Victims: Evidence from Violent Crime Investigations*. https://lucarizzotti.github.io/lucarizzotti.com/JMP_Luca_Rizzotti.pdf

Schwartzstein, J. (2014). Selective attention and learning. *Journal of the European Economic Association*, *12*(6), 1423–1452.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, 99–118.

Stashko, A. (2023). Do Police Maximize Arrests or Minimize Crime? Evidence from Racial Profiling in U.S. Cities. *Journal of the European Economic Association*, *21*(1), 167–214.

Stigler, G. J. (1961). The economics of information. *Journal of political economy*, *69*(3), 213–225.

USSC. (2023, November 9). *2023 Demographic Differences in Federal Sentencing*. United States Sentencing Commission. Retrieved November 12, 2024, from https://www.ussc.gov/research/research-reports/2023-demographic-differences-federal-sentencing

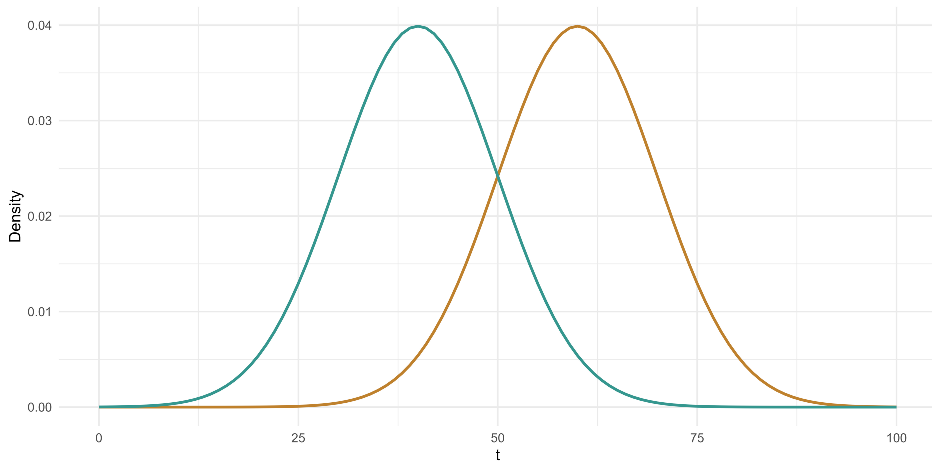
Vomfell, L., & Stewart, N. (2021). Officer bias, over-patrolling and ethnic disparities in stop and search. *Nature Human Behaviour*, *5*(5), 566–575.

Whittle, P. (1983, June). *Optimization Over Time, Dynamic Programming and Stochastic Control* (Volume 2 edition). Wiley.

Zhuo, R. (2023). Exploit or explore? An empirical study of resource allocation in research labs.

Tables and Figures

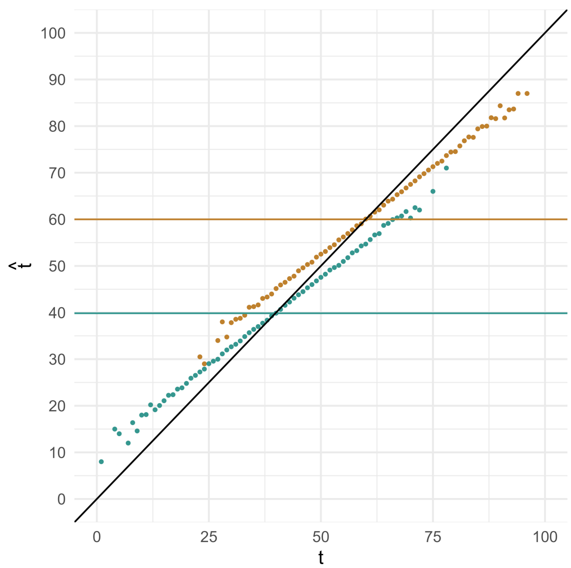
Figure A1: Example of prior distributions presented in the experiment



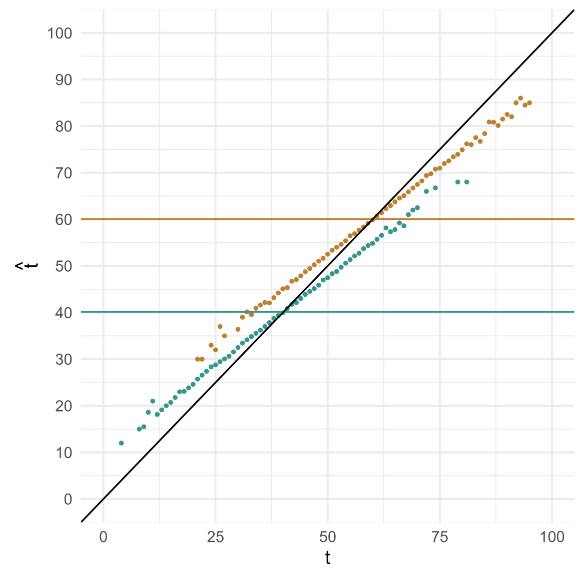
Notes: Each density corresponds to the distribution of types (crime rates) of neighborhoods across Brown city and Green city.

Figure A2: Simulation of Inference Task, effect of SN

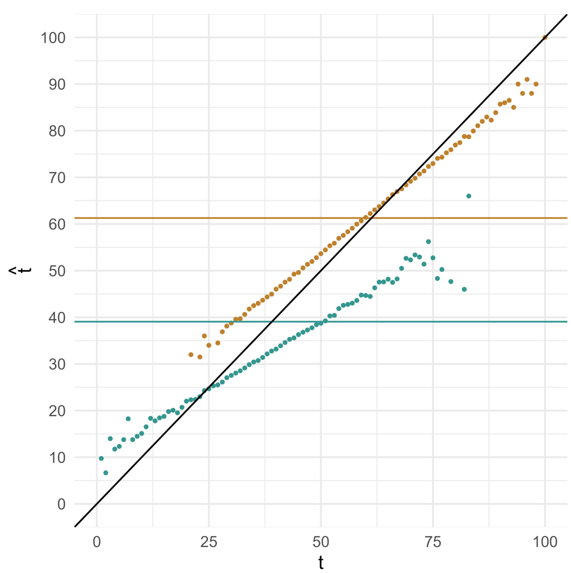
(a) Balanced Selection with low SN



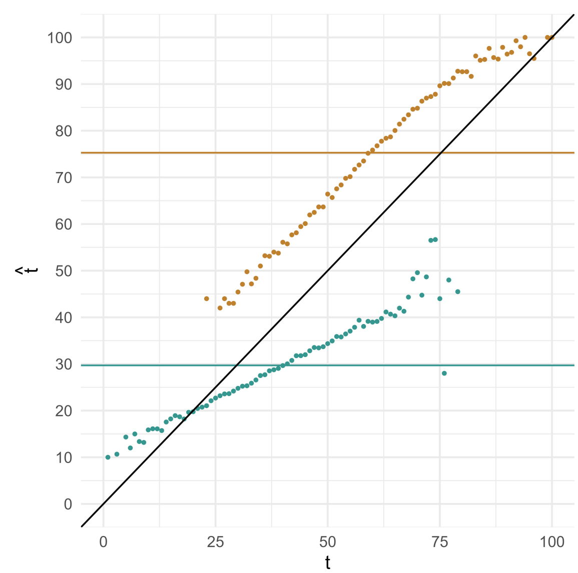
(b) Balanced Selection with high SN



(c) Differential Selection with low SN

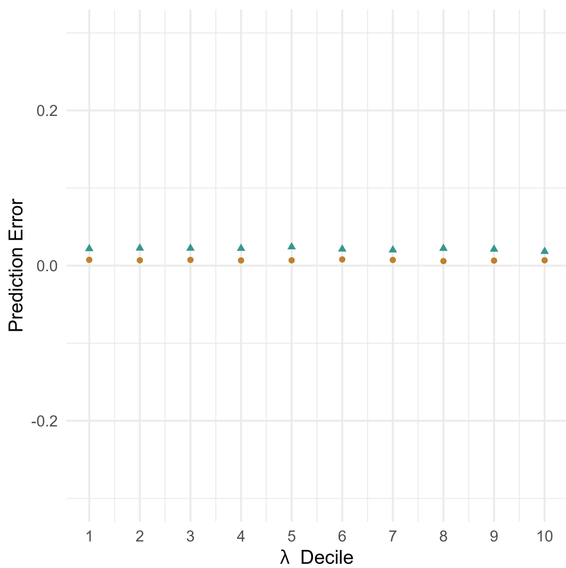


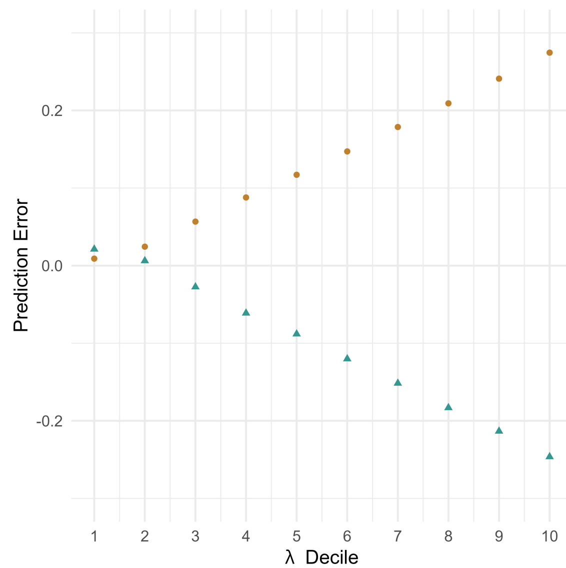
(d) Differential Selection with high SN



Notes: N = 250,000 (10,000 individuals, 25 rounds). Panels (a) and (b) show Balanced Selection ($p_B = p_G = 100$), while Panels (c) and (d) show the Differential Selection subset ($p_B > 100, p_G < 100$). Panels (a) and (c) show the subset of individuals with low selection neglect ($\lambda < 0.2$), while Panels (b) and (d) show those individuals with high selection neglect ($\lambda > 0.8$),

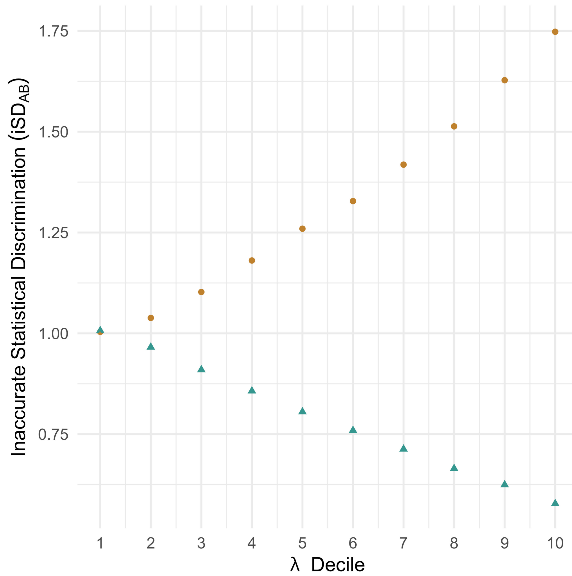
Figure A3: Simulated Prediction Error





Notes: Total N = 500,000 (10,000 individuals x 25 rounds x 2 estimations). We simulate 1,000 individuals with $\lambda = 0$ and 9,000 with $\lambda \sim \mathcal{N}(0.5, 1)$ and truncated to $\lambda \in [0, 1]$. Panels (a) shows the average prediction error by decile of selection neglect in the Balanced Selection subset ($p_B = 100, p_G = 100$, N=99,925), and Panel (b) shows the Differential Selection subset. Brown circles show the case when Brown group is overselected ($p_B > 100, p_G < 100$, N = 199,771), and Green triangles the subset where the Green group is overselected ($p_B < 100, p_G > 100$, N=199,870).

Figure A4: Simulated Inaccurate Statistical Discrimination



Notes: The figure displays the average inaccurate statistical discrimination (ratio of predictions across groups, net of the actual distribution differences). Brown circles show the case when Brown group is overselected ($p_B > 100, p_G < 100$, N = 199,771), and Green triangles the subset where the Green group is overselected ($p_B < 100, p_G > 100$, N=199,870).

A1 Appendix Derivations

Lemma: *Let $t \sim \mathcal{N}(\mu_g, \sigma^2)$ and $s \sim \mathcal{N}(t, V_g^2)$ then the optimal Bayesian estimate $\hat{t}$ is given by $\hat{t} = \mu_g \omega_g + s(1 - \omega_g)$ with $\omega_g = \frac{V_g^2}{V_g^2 + \sigma^2}$*

Proof. Start by noting that

$$\begin{aligned}
f^p(t|s,g) &= f(t|g)h(s|t,g) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(t-\mu_g)^2}{2\sigma^2}\right) \cdot \frac{1}{\sqrt{2\pi V_g^2}} exp\left(-\frac{(s-t)^2}{2V_g^2}\right) \\
&= K exp\left(-\frac{1}{2}\left(\frac{t^2 - 2t\mu_g + \mu_g^2}{\sigma^2} - \frac{s^2 - 2st + t^2}{2V_g^2}\right)\right)
\end{aligned}$$

Now note that every term that does not include t can be seen as proportionally constant, so we can drop it and include it in the constant K.

$$\begin{aligned}
f^p(t|s,g) &= \tilde{K} exp\left(-\frac{1}{2}\left(\frac{t^2 - 2t\mu_g + \mu_g^2}{\sigma^2} - \frac{s^2 - 2st + t^2}{2V_g^2}\right)\right) \\
&= \tilde{K} exp\left(-\frac{1}{2}\left(\frac{t^2(V_g^2 + \sigma^2) - 2t(\mu_g V_g^2 + s\sigma^2}{V_g^2\sigma^2}\right)\right) \\
&= \tilde{K} exp\left(-\frac{1}{2}\left(\frac{t^2 - 2t\left(\frac{\mu_g V_g^2 + s\sigma^2}{V_g^2 + \sigma^2}\right)}{\frac{V_g^2\sigma^2}{V + g^2 + \sigma^2}}\right)\right) \\
&= \hat{K} exp\left(-\frac{1}{2}\frac{\left(t - \frac{\mu_g V_g^2 + s\sigma^2}{V_g^2 + \sigma^2}\right)^2}{\frac{V_g^2\sigma^2}{V + g^2 + \sigma^2}}\right)
\end{aligned}$$

Note the role of the normalization factor in the derivation of this expression.

This approximates to a normal distribution with mean $\frac{\mu_g V_g^2 + s\sigma^2}{V_g^2 + \sigma^2}$ and variance $\frac{V_g^2\sigma^2}{V + g^2 + \sigma^2}$. Hence,

$$\mathbb{E}^p[t|s,g] = \mu_g \omega_g + s(1 - \omega_g)$$

where $\omega_g = \frac{V_g^2}{V_g^2 + \sigma^2}$

Finally, its easy to show that the optimal guess for $\hat{t} = \mathbb{E}^p[t|s,g]$ since $\mathbb{E}^p[t|s,g]$ is the value with the highest probability mass, given that its *pdf* is given by $f^p(t|s,g)$, which is a Normal distribution. □

Lemma: *Let* $t|g \sim \mathcal{N}(\mu_g, \sigma^2)$ *and* $t|-g \sim \mathcal{N}(\mu_{-g}, \sigma^2)$, *then* $\tilde{f}(t|g) = \kappa f(t|g) R(t, g, -g)^{\gamma^p}$ *is a normal distribution with mean* $\mu_g + \gamma(\mu_g - \mu_{-g}$ *and variance* σ^2

Proof. Let's start with the ratio

$$
\begin{aligned}
R(t,g,-g)^{\gamma^p} = \left(\frac{f(t|g)}{f(t|-g)}\right) &= \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{\gamma}{2\sigma^2}\left((t-\mu_g)^2 - (t-\mu_{-g})^2\right)\right) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{\gamma}{2\sigma^2}\left(t^2 - 2t\mu_g + \mu_g^2 - t^2 + 2t\mu_{-g} - \mu_{-g}^2\right)\right) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2\sigma^2}\left(\gamma(2t(\mu_g - \mu_{-g}) + \mu_g^2 - \mu_{-g}^2)\right)\right)
\end{aligned}
$$

Now premultiply this by $f(t|g)$

$$
\begin{aligned}
\tilde{f}(t|g) &= \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2\sigma^2}\left(t^2 - 2t\mu_g + \mu_g^2 + \gamma(2t(\mu_g - \mu_{-g}) - \mu_g^2 + \mu_{-g}^2)\right)\right) \\
&= \kappa \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(t - (\mu_g + \gamma(\mu_g - \mu_{-g}))^2}{2\sigma^2}\right)
\end{aligned}
$$

Where in the last step we absorved a set of constants in the normalziation factor κ. Thus, the distorted distribution is normally distributed with the mean $\mu_g + \gamma(\mu_g - \mu_{-g}$ and variance σ^2.

□

A2 Numeracy Questions

1. **Numeracy 1**: The chance of getting a viral infection is .0005. Out of 10,000 people, about how many of them are expected to get infected?

2. **Numeracy 2**:Imagine that we roll a fair, six-sided die 1,000 times. Out of 1,000 rolls, how many times do you think the die would come up as an even number?

3. **Cross-multiplication 1**: In a jar there are some balls. If 75% of the jar has 60 balls, how many balls has the jar?

4. **Cross-multiplication 2**: 2 out of 3 students in a class are right handed. If there are 18 right handed students in the class, how many students are there in the class?

5. **CRT**: In a lake, there is a patch of lilypads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?