Attributes: Selective Learning and Influence*

Arjada Bardhi†

January 10, 2024

Abstract

An agent selectively samples attributes of a complex project so as to influence the decision of a principal. The players disagree about the weighting, or relevance, of attributes. The correlation across attributes is modeled through a Gaussian process, the covariance function of which captures pairwise attribute similarity. The key tradeoff in sampling is between the alignment of the players' posterior values for the project and the variability of the principal's decision. Under a natural property of the attribute correlation—the *nearest-attribute property* (NAP)—each optimal attribute is relevant for some player and at most two optimal attributes are relevant for only one player. We derive comparative statics in the strength of attribute correlation and examine the robustness of our findings to violations of NAP for a tractable class of distance-based covariances. The findings carry testable implications for attribute-based product evaluation and strategic selection of pilot sites.

Keywords: attribute covariance, Gaussian sample paths, nearest-attribute property, strategic sampling, powered-exponential covariances

JEL Classification: D83, D81, D72, D04

1 Introduction

Important decisions rely on selective exploration of multiple criteria or attributes: from buyers appraising complex products for purchase to employers assessing the diverse skills of potential employees, policymakers evaluating the heterogeneous spatial impact of large-scale social programs, and firms learning the demand across different consumer segments. Understanding the nature of such selective exploration—which attributes are optimally explored and why—has been of long-standing interest in economics (Huber and McCann (1982), Keeney and Raiffa (1976), Lancaster (1966)).

*I am indebted to Asher Wolinsky, Bruno Strulovici, Yingni Guo, and Georgy Egorov for their formative feedback. I thank the editor Bart Lipman and three anonymous referees, as well as Attila Ambrus, Nemanja Antic, Nina Bobkova, Steven Callander, Alex Frankel, Rachel Kranton, R. Vijay Krishna, Meg Meyer, Xiaosheng Mu, Alessandro Pavan, Andy Postlewaite, Debraj Ray, Philipp Sadowski, Todd Sarver, Takuo Sugaya, Andy Skrzypacz, Ennio Stacchetti, and numerous seminar and conference audiences for valuable feedback. The latter stages of this project were supported by NSF Award SES-1919329.

†Department of Economics, New York University. Email: arjada.bardhi@nyu.edu.

Two features common to these examples but largely overlooked by the literature on attribute learning complicate such understanding. First, in many settings of interest attributes are correlated. Evaluation relies on complex inferences from explored attributes to those left unexplored. Second, agency conflict often undermines attribute exploration; the party selecting which attributes to learn might weight attributes quite differently from the one making the multi-attribute decision.

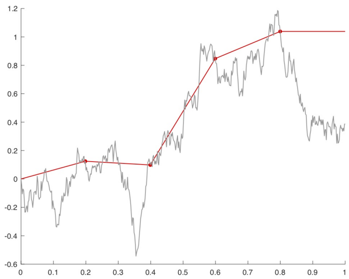
This paper models the strategic interaction between a principal and an agent who jointly evaluate an uncertain project consisting of a multitude of correlated attributes. The players share the same understanding of how any two attributes of the project are correlated but disagree on the relevance of attributes. They also have separate authorities in evaluation: the agent decides which attributes to sample publicly, whereas the principal decides on the project based on the observed sample. We study how properties of the attribute correlation shape strategic sampling.

Such a sampling problem arises, for instance, when a consumer learns his willingness to pay for a multi-attribute product. Ample evidence from marketing suggests that consumers use their beliefs about how the product attributes covary together to draw inferences from attributes that are observed to those that remain unobserved (Walters and Hershfield (2020), Mason and Bequette (1998)). For certain packaged food products, for instance, consumers might be able to observe only the attributes that a health authority requires to be displayed in the nutrition label. How should such a health authority design the content of nutrition labels if concerned that consumers tend to underweight certain attributes, such as health considerations, in favor of other attributes, such as taste? In another context, a policymaker might seek to assess the overall impact of a social program rolled out across several sites from pilot studies run at select sites. The local impact at a given site corresponds to an attribute, and such impact is naturally correlated across sites because of geographic proximity or other shared characteristics. However, pilot studies are often decided upon and run by independent agencies with their own weighting of the importance of different sites. Are the pilot sites among those that the policymaker cares most about?[1]

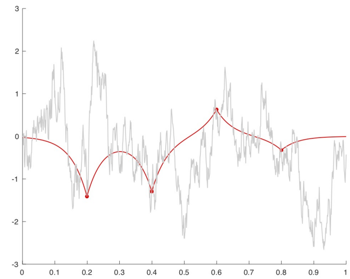
This paper provides a novel and flexible framework for attribute sampling. The project under evaluation is characterized by a mass of attributes $\mathcal{A}$, of which the agent can sample up to k attributes. Attribute realizations follow an unknown mapping, randomly drawn from the space of the sample paths of a general Gaussian process. The process is pinned down by two commonly known parameters: (i) an attribute mean $\mu : \mathcal{A} \to \mathbb{R}$, which encodes the expected realizations of attributes, and (ii) a pairwise covariance $\sigma : \mathcal{A} \times \mathcal{A} \to \mathbb{R}$, which encodes the similarity between any two attributes. This framework covers a vast array of attribute mappings and patterns of extrapolation from a sample of attributes, as shown in Figure 1 for four different Gaussian processes. Each player's value from the project consists of a sum of weighted attribute realizations, where the weight captures the relevance of that attribute for the player.[2] Each player seeks to minimize the quadratic loss from the principal's

[1] A growing literature starting with Allcott (2015) has pointed out the low external validity of pilot findings due to strategic selection of pilot sites. See the discussion in section 5.4.

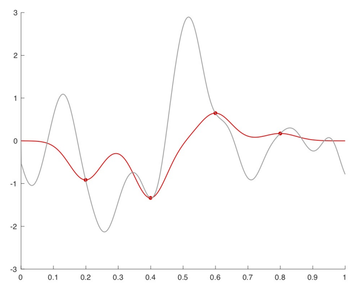
[2] Such linear aggregation of attributes is a special case of the utility specification in Lancaster (1966) and Keeney and Raiffa (1976) and it is standard in many applications of the multi-attribute framework, e.g., Spiegler (2016)



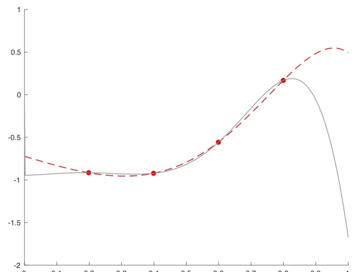
(a) Brownian: $\sigma(a, a') = \min(a, a')$



(b) Ornstein-Uhlenbeck: $\sigma(a, a') = e^{-20|a-a'|}$



(c) Squared-exponential: $\sigma(a, a') = e^{-400(a-a')^2}$



(d) Polynomial: $\sigma(a, a') = (1 + aa')^{10}$

Figure 1: The attribute mapping is in grey and the extrapolated mapping in red. For all plots, $\mathcal{A} = [0, 1]$ *and the attribute mean* $\mu(\cdot)$ *is zero. The dots denote the realizations of the sample* $\mathbf{a} = \{0.2, 0.4, 0.6, 0.8\}$.

prediction of the value of the project.

The contribution of the analysis is threefold. First, we provide a general characterization of the optimal sample for any attribute covariance and any disagreement about attribute weights. In the single-player benchmark—i.e., in the absence of disagreement—the value of a sample is summarized by the *posterior variance* that it induces on the player's estimate of the value of the project. This sufficient statistic encodes the tradeoff between the redundancy within the sample and the sample's generalizability to out-of-sample attributes: the single-player sample consists of attributes that are weakly correlated with each other while strongly correlated with out-of-sample attributes. In contrast, in the presence of disagreement, the agent's payoff from a sample is summarized by two statistics: (i) the covariance that the sample induces between the principal's decision and the agent's value, and (ii) the posterior variance for the principal. These statistics capture the key tradeoff in sampling: the agent prefers the sample to be more informative for the principal only insofar as it contributes to greater alignment between the players. Theorem 2 unpacks this tradeoff in terms of the inferences that players draw about individual attributes in the sample.

To study the precise composition of the optimal sample of attributes, we identify a natural property of the attribute covariance—the *nearest-attribute property* (hereafter, NAP)—that renders this characterization particularly tractable. This property, which is the second key contribution

on boundedly rational consumers, Srinivasan and Shocker (1973) on product evaluation, Nevo (2001) on product differentiation etc.

of the analysis, implies local extrapolation across the attribute space: the players' inference about an out-of-sample attribute is informed only by the neighboring attributes in the sample. In other words, under NAP, the observed attributes segment the project into ex post independent clusters of unobserved attributes, thus rendering the project modular ex post.

We first address whether the optimal sample includes irrelevant attributes, i.e., attributes the realizations of which do not enter directly the value of the project for either player. Intuitively, irrelevant attributes are valuable to sample: for example, in the single-player benchmark they might be informative about many relevant attributes, or in the strategic setting they might be a natural compromise whenever the relevant attributes of the two players are far apart. However, we find that if the attribute covariance satisfies NAP, each attribute in the optimal sample is necessarily relevant for at least one of the players. In the single-player benchmark, irrelevant attributes are too low-powered to be informative about multiple ex post independent clusters of relevant attributes at the same time. In the presence of disagreement, moreover, local extrapolation limits both the extent to which irrelevant attributes are taken into account in the principal's decision and the extent to which they contribute to better alignment between players. Second, among attributes relevant for some player, sampling of idiosyncratic attributes (i.e., attributes that are relevant for only one of the players) is minimal: the optimal sample includes at most two idiosyncratic attributes. Moreover, we show that the agent optimally either (i) selects a small sample of at most three attributes, at most one of which is common (i.e., equally relevant for both players) and the rest are idiosyncratic, or (ii) exhausts the available sampling capacity, in which case more than half of the optimal sample consists of common attributes.

Our third contribution is to use a parametric class of distance-based attribute covariances—the *powered-exponential* covariances—in order to investigate both how the optimal sample varies with the strength of attribute correlation and the extent to which the aforementioned results are robust to violations of NAP. For any attribute covariance within this class, correlation across attributes weakens exponentially with distance.[3] Whenever NAP is satisfied within this class, the single-player sample expands with stronger attribute correlation in order to offset the increased redundancy within the sample; the optimal attributes in the left half of the attribute space move closer to the leftmost relevant attribute and those in the right half move closer to the rightmost relevant attribute. This continues to be the case qualitatively also in the strategic setting if the mass of common attributes is sufficiently large. By contrast, if the common attributes are few, stronger attribute correlation makes the agent increasingly more concerned about the sample being too informative for the principal, which strengthens incentives to sample closer to the agent's idiosyncratic attributes. Furthermore, this class of covariance allows us to examine how violations of NAP might make it optimal to sample irrelevant attributes. Through a stylized example with only two relevant attributes, we show that sampling of irrelevant attributes is optimal precisely for those powered-exponential covariances for which the

[3] As section 2.3 makes clear, the Ornstein-Uhlenbeck process (Figure 1b) and the squared-exponential process (Figure 1c) both have powered-exponential covariances. However, only the covariance of the first satisfies NAP.

two attributes covary negatively conditional on any irrelevant attribute. This is the case in both the single-player setting and the strategic setting.

It is worth highlighting two further technical contributions. First, our approach generalizes the Brownian motion approach in the experimentation and search literature (Callander (2011), Jovanovic and Rob (1990)). The Brownian covariance satisfies NAP, therefore the tools that we develop to analyze NAP covariances are transportable to problems beyond attribute sampling. A thorough discussion of related work is postponed until section 5.3. Second, our characterization clarifies a key difference between learning through correlated *attributes* and learning through correlated *signals*.[4] Attributes are at the same time constituent components of the value of the project and feasible signals about that value. Remark 1 fleshes out what this dual role of attributes implies for the optimal sample and the value of sampling. The framework also has empirical implications for the consumer research literature on evaluation of multi-attribute products, further discussed in section 5.1.

Section 2 sets up the model, establishes preliminary results on extrapolation, and introduces NAP and the class of powered-exponential covariances. Sections 3.1 and 4.1 provide a general characterization of optimal sampling in the single-player benchmark and in the principal-agent game, respectively. Sections 3.2 and 4.2 illustrate the characterization for the class of powered-exponential covariances. Section 5 discusses implications and extensions of our framework.[5]

2 Framework

2.1 Model

Players. A principal (P, she) and an agent (A, he) jointly evaluate a multi-attribute project of unknown quality. The players have separate authorities: the agent chooses which attributes to sample, whereas the principal makes a decision regarding the project based on the observed attributes. The players differ in their respective weighting of attributes, which gives rise to agency conflict.

Attributes. The project is characterized by a compact attribute space $\mathcal{A} = [\underline{\alpha}, \bar{\alpha}] \subset \mathbb{R}$. The realization of attribute $a \in \mathcal{A}$ is denoted by $f(a) \in \mathbb{R}$. Attribute realizations follow an unknown mapping $f : \mathcal{A} \to \mathbb{R}$, drawn from the space of sample paths of a Gaussian process with *prior mean* function $\mu : \mathcal{A} \to \mathbb{R}$ and symmetric positive semi-definite *covariance* function $\sigma : \mathcal{A} \times \mathcal{A} \to \mathbb{R}$.[6] Per

[4] The existing literature has taken these to be equivalent problems. E.g., Klabjan, Olszewski and Wolinsky (2014) state that "The x_i's can be viewed as realizations of signals about the value of the object, rather than as actual attributes."

[5] Appendix A provides auxiliary results supporting section 2.1. Appendix B collects proofs of results in sections 2.2, 2.3, 3.1, and 4.1. Appendices C and D collect proofs and auxiliary results for sections 3.2 and 4.2 respectively. Appendix E supports the discussion in section 5.

[6] Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a stochastic process $f := \{f(a, \omega)\}_{a \in \mathcal{A}, \omega \in \Omega}$ with index set $\mathcal{A}$ is a Gaussian process if and only if $(f(a_1), \ldots, f(a_n))$ are jointly Gaussian for any $a_1, \ldots, a_n \in \mathcal{A}$ and $n \geq 1$. See Appendix A and Rasmussen and Williams (2006) for a technical introduction to Gaussian processes.

standard notation, the distribution over possible attribute mappings is

$$f \sim \mathcal{GP}\left(\mu, \sigma\right). \tag{1}$$

Parameters (μ, σ) are commonly known and they pin down the distribution of f. The prior mean $\mu(a) := \mathbb{E}[f(a)]$ specifies the expected realization of attribute $a \in \mathcal{A}$, whereas $\sigma(a, a')$ specifies the covariance between $f(a)$ and $f(a')$ for any attribute pair $a, a' \in \mathcal{A}$. The covariance can be interpreted as a similarity measure over attributes, as discussed in section 5.2. Figure 2a illustrates the attribute mapping and the prior mean function for a familiar Gaussian process, namely the Brownian motion.

To impose regularity on the attribute structure, we require that attributes that are close in the attribute space have comparable realizations in the attribute mapping. An implication of this requirement is that μ and σ are both continuous.[7] That is, the realizations of any two attributes arbitrarily close in $\mathcal{A}$ are almost equal in expectation and almost perfectly correlated.

Assumption 1 (Sample-path continuity)**.** *Almost surely any realization of f is continuous.*

Sampling. Because f is a sample path of a Gausian process, the realizations of any sample of k attributes $\mathbf{a} = \{a_1, \ldots, a_k\}$ are jointly Gaussian, that is

$$f(\mathbf{a}) := \begin{pmatrix} f(a_1) \\ \vdots \\ f(a_k) \end{pmatrix} \sim \mathcal{N}\left(\underbrace{\begin{pmatrix} \mu(a_1) \\ \vdots \\ \mu(a_k) \end{pmatrix}}_{:=\mu(\mathbf{a})}, \underbrace{\begin{pmatrix} \sigma(a_1, a_1) & \ldots & \sigma(a_1, a_k) \\ \vdots & \ddots & \vdots \\ \sigma(a_k, a_1) & \ldots & \sigma(a_k, a_k) \end{pmatrix}}_{:=\Sigma(\mathbf{a})} \right).$$

Let $|\mathbf{a}|$ denote the number of distinct attributes in $\mathbf{a}$ and let $a \in \mathbf{a}$ be one such attribute. The sample mean and the sample covariance matrix are denoted by $\mu(\mathbf{a})$ and $\Sigma(\mathbf{a})$ respectively.

The agent samples attributes subject to a sampling capacity $k \in \mathbb{N}$: the cost of drawing n attributes is zero if $n \leq k$ and $+\infty$ otherwise.[8] Upon sampling, the pair $(\mathbf{a}, f(\mathbf{a}))$ is observed publicly and players are symmetrically informed throughout the game. Without loss we restrict attention to non-redundant samples, i.e., samples of attributes the realizations of which are linearly independent. That is, the set of feasible samples consists of

$$\mathcal{A}_k := \bigcup_{n \leqslant k} \left\{ \{a_1, \ldots, a_n\} \subset \mathcal{A} : \Sigma\left(\{a_1, \ldots, a_n\}\right) \text{ is non-singular} \right\}. \tag{2}$$

Decision and payoffs. Player i's value for the project is a weighted sum of attribute realizations

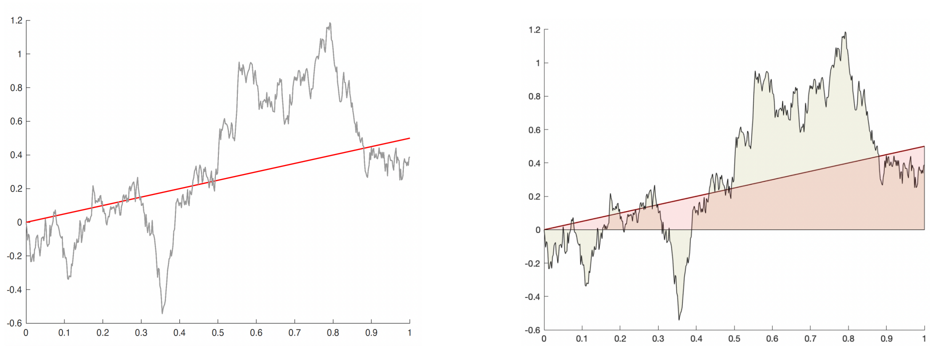
$$v_i = \int_{\mathcal{A}} f(a) \omega_i(a) \, \mathrm{d}a, \tag{3}$$

where ω_i is player i's attribute weight function, assumed to be bounded and continuous almost everywhere, and hence integrable. An attribute is *relevant* for player i if $\omega_i(a) \neq 0$ and it is *irrelevant*

[7]Appendix A provides sufficient conditions on (μ, σ) to guarantee sample-path continuity (Proposition 13) and establishes that sample-path continuity implies continuity of μ and σ (Proposition 14).

[8]Due to the Gaussian distribution of attribute realizations and quadratic loss payoffs (as introduced next), simultaneous and sequential sampling are equivalent here.

otherwise. Correspondingly, the set of relevant attributes $\text{supp}(\omega_i) := \{a \in \mathcal{A} : \omega_i(a) \neq 0\}$ is compact for each player.[9] We let $\underline{a}_i$ and $\bar{a}_i$ denote, respectively, the leftmost and the rightmost relevant attributes for player i. Of particular interest to our analysis is the case in which the set of relevant attributes is an interval; that is, $\omega_i(a) \neq 0$ if and only if $a \in [\underline{a}_i, \bar{a}_i] \subset \mathcal{A}$. A relevant attribute a is said to be *desirable* (or *undesirable*) for player i if $\omega_i(a) > 0$ (or $\omega_i(a) < 0$).



(a) Attribute mapping f (grey) and prior mean $\mu(a) = a/2$ (red).

(b) The player's value v_i (yellow) and prior value ν_0^i (red) for $\omega_i(a) = 1$ for all $a \in \mathcal{A}$.

Figure 2: The Brownian motion with drift $1/2$ and $\sigma(a, a') = \min(a, a')$ for all $a, a' \in \mathcal{A} = [0, 1]$.

The pair of attribute weight functions (ω_A, ω_P) is commonly known. Lemma 3 (in Appendix A) establishes that (i) the attribute weights and the attribute covariance can be normalized without loss so that the variance of each attribute is either zero or one (a normalization that we adopt hereafter), and (ii) in the single-player benchmark it is without loss for all attributes to be desirable and the sum of attribute weights to equal to one. Due to Assumption 1, the value v_i and its distribution are well-defined. The ex ante distribution of the value for player $i \in \{A, P\}$ is

$$v_i \sim \mathcal{N}\left(\underbrace{\int_{\mathcal{A}} \mu(a)\omega_i(a)\,\mathrm{d}a}_{:=\nu_0^i},\ \int_{\mathcal{A}}\int_{\mathcal{A}} \sigma(a, a')\omega_i(a)\omega_i(a')\,\mathrm{d}a\,\mathrm{d}a'\right). \tag{4}$$

Let $\nu_0^i = \mathbb{E}[v_i]$ denote i's prior (expected) value of the project. Graphically, in Figure 2b this prior value corresponds to the area under the linear prior mean μ. Similarly, let $\nu^i(\mathbf{a}, f(\mathbf{a})) = \mathbb{E}[v_i \mid \mathbf{a}, f(\mathbf{a})]$ be the posterior (expected) value of the project for player i given the sample $(\mathbf{a}, f(\mathbf{a}))$. For brevity of notation, we denote it as $\nu^i(\mathbf{a})$.

The principal takes a decision $d \in \mathbb{R}$ based on his posterior value. For a given attribute mapping f with corresponding value v_i, the ex post payoff of player i takes the quadratic loss form:

$$u_i(d, v_i) = -(d - v_i)^2. \tag{5}$$

That is, each player prefers that the decision matches his value from the project. In the absence of any sampling, therefore, the principal takes the status quo decision equal to ν_0^P.

[9] If the set of relevant attributes is finite, the integral (3) becomes a weighted sum.

2.2 Extrapolation and inference

Player i's inferential problem proceeds in two steps, as described in Lemma 1. He first extrapolates from the realizations of the sample to the rest of the attribute space, which results in the extrapolated attribute mapping as depicted in Figure 1. This is the best estimate that the player has for out-of-sample attributes. He then infers the posterior value of the project from the extrapolated mapping.

Lemma 1. *Fix sample* $\mathbf{a} = \{a_1, \ldots, a_n\} \in \mathcal{A}_k$ *with* $n \leq k$ *and realizations* $f(\mathbf{a}) \in \mathbb{R}^n$*.*

(i) For any attribute $\hat{a} \in \mathcal{A}$*, its expected realization is given by*

$$\mathbb{E}[f(\hat{a}) \mid \mathbf{a}, f(\mathbf{a})] = \mu(\hat{a}) + \sum_{j=1}^{n} \tau_j(\hat{a}; \mathbf{a}) \left(f(a_j) - \mu(a_j) \right), \tag{6}$$

where $\tau_j(\hat{a}; \mathbf{a})$ *is the* $(1, j)^{th}$ *entry of the vector* $\begin{pmatrix} \sigma(a_1, \hat{a}) & \ldots & \sigma(a_n, \hat{a}) \end{pmatrix} [\Sigma(\mathbf{a})]^{-1}$*. For any* $j = 1, \ldots, n$ *and* $m \neq j$*,* $\tau_j(a_j; \mathbf{a}) = 1$ *and* $\tau_m(a_j; \mathbf{a}) = 0$*.*

(ii) Player i's posterior value is a linear combination of the sample realizations

$$\nu^i(\mathbf{a}) = \nu_0^i + \sum_{j=1}^{n} \tau_j^i(\mathbf{a}) \left(f(a_j) - \mu(a_j) \right) \tag{7}$$

where realization $f(a_j)$ *is weighted by the sample weight*

$$\tau_j^i(\mathbf{a}) := \int_{\mathcal{A}} \tau_j(\hat{a}; \mathbf{a}) \omega_i(\hat{a}) \, \mathrm{d}\hat{a}. \tag{8}$$

When extrapolating to an out-of-sample attribute $\hat{a} \notin \mathbf{a}$, the player considers how similar that attribute is to each sample attribute, as well as how strongly sample attributes covary with each other. A player's best estimate for $f(\hat{a})$ is a linear combination of the sample realizations, where the j^{th} attribute in the sample is weighted by $\tau_j(\hat{a}; \mathbf{a})$.[10] For instance, if the attribute mapping is a Brownian sample path, equation (6) simplifies to the familiar Brownian-bridge extrapolation in Callander (2011). More generally, the shape of the extrapolated mapping is dictated by the parametric form of the attribute covariance function.

The posterior value of the project is also a linear combination of the sample realizations. The sample weight $\tau_j^i(\mathbf{a})$ corresponds to the marginal change in player i's posterior value as a result of a marginal change in the realization $f(a_j)$ for $a_j \in \mathbf{a}$. Since the players share the same understanding of the attribute covariance but assign different attribute weights, they extrapolate in the same way to out-of-sample attributes but hold generically different posterior values. Whenever appropriate, we simplify notation by writing the sample weight of a singleton sample as $\tau^i(a_1) := \tau_1^i(\{a_1\})$.

[10] Because sample realizations are observed without noise, the extrapolated mapping traverses the realizations. Our focus is on *which* attributes are sampled from a large attribute space rather than *how* precisely each is sampled. Appendix E.3 discusses an extension to noisy observations of attribute realizations.

2.3 Nearest-attribute property

Of particular interest to our analysis is a class of highly tractable attribute covariances for which extrapolation to out-of-sample attributes is *local*, in the sense that it takes into account only the realizations of those sample attributes in the immediate vicinity of the out-of-sample attribute. Given a sample $\mathbf{a} \in \mathcal{A}_k$ and an attribute $a \in \mathcal{A}$, let $n(a; \mathbf{a}) \subset \mathbf{a}$ be the subset of the nearest sample attributes $\max_{\mathbf{a}}\{a_j : a_j < a\}$ and $\min_{\mathbf{a}}\{a_j : a_j > a\}$. Any out-of-sample attribute has at most two such neighbors.

Definition 1. The covariance σ satisfies the *nearest-attribute property* (NAP) if $\mathbb{E}[f(a) \mid \mathbf{a}, f(\mathbf{a})] = \mathbb{E}[f(a) \mid n(a; \mathbf{a}), f(n(a; \mathbf{a}))]$ for any $a \in \mathcal{A}, \mathbf{a} \in \mathcal{A}_k$, and $f(\mathbf{a}) \in \mathbb{R}^{|\mathbf{a}|}$.

Lemma 2 identifies necessary and sufficient conditions on the process and on the attribute covariance that guarantee NAP. This property goes hand in hand with the process from which the attribute mapping is drawn being Markov: for any $a < a' < a''$, observing $f(a')$ renders $f(a)$ and $f(a'')$ conditionally independent. Any sample of k attributes segments the attribute space into $(k+1)$ conditionally independent clusters of attributes. That is, upon sampling, the project looks ex post modular. In turn, this is equivalent to the correlation between any two arbitrary attributes a and a'' being equal to the product of the correlation between a and a' and that between a' and a'' for any $a' \in (a, a'')$. That is, how similar attribute $a' \in (a, a'')$ is to a and to a'' respectively exactly pins down how similar a and a'' are to each other.[11]

Lemma 2. *The following statements are equivalent:*[12]

(i) the attribute covariance σ satisfies NAP;

(ii) for any $a, a', a'' \in \mathcal{A}$ such that $a < a' < a''$, $\sigma(a, a'') = \sigma(a, a')\sigma(a', a'')$;

(iii) the attribute mapping f follows a Markov process.

Moreover, we can obtain explicit expressions for the sample weights that arise from NAP covariances. Because extrapolation is local, the sample weight aggregates the covariance of the sample attribute with any out-of-sample attributes in its immediate vicinity. The tractability of these sample weights allows us to derive general properties of optimal sampling for NAP covariances.[13]

Corollary 1. *Suppose σ satisfies NAP. Then, $\sigma(a, a') \geqslant 0$ for any $a, a' \in \mathcal{A}$. Moreover, for any $\mathbf{a} = \{a_1, \dots, a_n\} \in \mathcal{A}_k$ and each player i, the sample weights are given by*

$$\tau_1^i(\mathbf{a}) = \int_{\underline{a}}^{a_1} \sigma(a, a_1)\omega_i(a)\,\mathrm{d}a + \int_{a_1}^{a_2} \sigma(a, a_1)\frac{1 - \sigma^2(a, a_2)}{1 - \sigma^2(a_1, a_2)}\omega_i(a)\,\mathrm{d}a,$$

$$\tau_j^i(\mathbf{a}) = \int_{a_{j-1}}^{a_j} \sigma(a, a_j)\frac{1 - \sigma^2(a, a_{j-1})}{1 - \sigma^2(a_{j-1}, a_j)}\omega_i(a)\,\mathrm{d}a + \int_{a_j}^{a_{j+1}} \sigma(a, a_j)\frac{1 - \sigma^2(a, a_{j+1})}{1 - \sigma^2(a_j, a_{j+1})}\omega_i(a)\,\mathrm{d}a,$$

[11] This property is reminiscent of the notion of segmental additivity in spatial models of perceptual similarity in cognitive psychology (Beals, Krantz and Tversky, 1968).

[12] Per Lemma 3 in Appendix A, the statement of part (ii) takes it as given that all attributes have unit variance. However, the proof allows for flexible attribute variance and does not rely on Lemma 3.

[13] Sample weights are also arguably easier to elicit from observational data than the entire attribute covariance function. Hence, NAP can be tested indirectly using Corollary 1.

$$\tau_n^i(\mathbf{a}) = \int_{a_n}^{\bar{a}} \sigma(a, a_n)\omega_i(a)\,\mathrm{d}a + \int_{a_{n-1}}^{a_n} \sigma(a, a_n)\frac{1-\sigma^2(a, a_{n-1})}{1-\sigma^2(a_{n-1}, a_n)}\omega_i(a)\,\mathrm{d}a.$$

If all attributes are desirable (undesirable) for player i, then $\tau_j^i(\mathbf{a}) \geqslant (\leqslant) 0$ for any $j \leqslant n$.

Powered-exponential covariances. To illustrate the importance of NAP for sampling, we invoke a parametric class of covariances known as *powered-exponential covariances*. They take the distance-based form $\sigma_p(a, a') = e^{-\left(|a-a'|/\ell\right)^p}$ for $\ell > 0$ and $p \in (0, 2]$. Such covariances share key common features: all attributes are equally uncertain ex ante, the correlation between any two attributes decays exponentially with their distance in the attribute space, and the strength of attribute correlation is parametrized by $\ell > 0$.[14] However, such commonalities belie fundamentally different attribute mappings for different values of p. For instance, Figures 1b and 1c show that the sample paths corresponding to $p = 1$ are nowhere differentiable, whereas those corresponding to $p = 2$ are smooth.

Importantly, the only covariance that satisfies NAP within this class is that corresponding to $p = 1$. This is the covariance of the well-known Ornstein-Uhlenbeck (OU) process. We invoke the OU covariance throughout to sharpen the characterization of optimal sampling under NAP. The equality of Lemma 2.2(ii), which certainly holds for $p = 1$, becomes a strict inequality whenever NAP is violated: for any $a < a' < a''$, $\sigma_p(a, a'') \gtrless \sigma_p(a, a')\sigma_p(a', a'')$ for $p \lesseqgtr 1$. This inequality has a natural interpretation. The ex ante covariance between attributes a and a'' is $\sigma_p(a, a'')$, but the covariance after $f(a')$ is observed decreases to $\sigma_p(a, a'') - \sigma_p(a, a')\sigma_p(a', a'')$. That is, after the sample is observed, the attribute realizations $f(a)$ and $f(a'')$ covary positively for $p < 1$, are independent for $p = 1$, and covary negatively for $p > 1$.

3 Benchmark: Single-player sampling

This section establishes a benchmark for optimal sampling in the absence of conflict, i.e., when $\omega_A = \omega_P = \omega$. We refer to this benchmark as *single-player sampling*. Section 3.1 first provides a general characterization of single-player sampling for any attribute covariance, and then employs this characterization to understand the importance of NAP for whether irrelevant attributes are optimally sampled. Section 3.2 illustrates this characterization for the class of powered-exponential covariances.

3.1 General characterization

Any sample of attributes induces a Gaussian distribution over the player's posterior value for the project. That is, $\nu(\mathbf{a}) \sim \mathcal{N}\left(\nu_0, \psi^2(\mathbf{a})\right)$ for any $\mathbf{a} \in \mathcal{A}_k$, where $\psi^2(\mathbf{a}) \coloneqq \mathrm{var}\left[\mathbb{E}[v \mid \mathbf{a}, f(\mathbf{a})]\right]$ is the *poste-*

[14] The exponential decay as a way of capturing similarity between two uncertain primitives in a spatial model is not without precedent. Shepard (1987) first postulated it as a natural model of generalizations across similar stimuli in a psychological space. More recently, Billot, Gilboa and Schmeidler (2008) offers an axiomatization of such a similarity function in the context of extrapolation from historical data.

rior variance induced by sample **a**. The posterior variance corresponds to the amount of uncertainty resolved by the sample. Upon observing the sample realizations, the player takes the decision that minimizes his expected quadratic loss. Such optimal decision equals the posterior value $\nu(\mathbf{a})$. The expected loss from sample **a** corresponds to the residual uncertainty about the project, since

$$V^s(\mathbf{a}) = \mathbb{E}\left[-\left(\nu(\mathbf{a}) - v\right)^2\right] = -\left(\text{var}[v] - \psi^2(\mathbf{a})\right). \tag{9}$$

Therefore, the posterior variance provides a sufficient statistic based on which the player ranks feasible samples. Samples with higher posterior variance lead to a lower expected loss. A distinctive feature of the attribute problem is that both $\text{var}[v]$ and $\psi^2(\mathbf{a})$—that is, both the amount of uncertainty that there is ex ante about the value of the project and that which is resolved by any feasible sample—are determined by the attribute covariance function. We revisit how the expected loss from the single-player sample varies with the degree of the attribute correlation in section 3.2. Theorem 1 characterizes single-player sampling for any general attribute covariance.

Theorem 1. *Any single-player sample $\mathbf{a}^s \in \mathcal{A}_k$ maximizes the posterior variance ψ^2 over the set of feasible samples $\mathcal{A}_k$, where ψ^2 is given by*

$$\psi^2(\mathbf{a}) = \sum_{j=1}^{|\mathbf{a}|} \tau_j(\mathbf{a})\tau(a_j) \tag{10}$$

for any $\mathbf{a} \in \mathcal{A}_k$. The set of single-player sample does not depend on μ or ν_0.

The player optimally identifies attributes that, when considered in the context of the entire sample, are most useful in predicting the value of the project. If considered in isolation from the rest of the sample, each sample attribute $a_j \in \mathbf{a}$ carries informativeness $\tau^2(a_j)$. So if the player were to ignore correlation within the sample, the posterior variance would be $\sum_{j=1}^{|\mathbf{a}|} \tau^2(a_j)$. When correlation is taken into account, however, $\tau(a_j)$ is scaled by its actual sample weight $\tau_j(\mathbf{a})$ rather than $\tau(a_j)$ in the j^{th} term of this summation. Attribute a_j contributes to posterior variance directly through the term $\tau_j(\mathbf{a})\tau(a_j)$ and indirectly through its impact on the rest of the sample weights $\tau_{-j}(\mathbf{a})$.[15]

Two immediate implications are worth noting. First, a project is evaluated based on the same optimal sample of attributes regardless of its prior value. Hence, if the player were to sample attributes sequentially, the optimal attribute to be sampled next would not depend on the past history of sample realizations. Whether the player samples attributes simultaneously or sequentially is immaterial. Second, any single-player sample uses the entire capacity k, unless the uncertainty about the project can be fully resolved with fewer than k attributes. The latter would be the case if the attribute weight function ω is such that there are fewer than k relevant attributes, or if the attribute covariance σ is such that there are no non-redundant samples of exactly k attributes, i.e., $\mathcal{A}_k = \mathcal{A}_{k-1}$.

[15] An observer need not know the true σ in order to derive the value of a sample V^s: it suffices to elicit the player's sample weights $\tau(a_j)$ and $\tau_j(\mathbf{a})$. In turn, the observer can solve for the pairwise covariances $\sigma(a_i, a_j)$ for each $a_i, a_j \in \mathbf{a}$ by solving the system of $|\mathbf{a}|$ linear equations of the form $\sum_{i=1}^{|\mathbf{a}|} \tau_i(\mathbf{a})\sigma(a_i, a_j) = \tau(a_j)$ (see proof of Theorem 1).

Marginal value of a sample attribute. In the presence of NAP, the posterior variance (10) can be simplified further because each sample weight depends on at most three adjacent sample attributes. Proposition 1 derives the marginal value of each sample attribute, which depends only on the attribute itself and its neighboring sample attributes. Intuitively, the marginal value is higher the more similar the sample attribute is to out-of-sample attributes in its immediate vicinity and the more dissimilar it is to the neighboring sample attributes. NAP reduces the player's sampling problem to a set of local optimization problems.

Proposition 1. *Let σ satisfy NAP. In any single-player sample $\mathbf{a}^s = \{a_1^s, \ldots, a_n^s\} \in \mathcal{A}_k$, where $n \leqslant k$ and $a_j^s < a_{j+1}^s$ for $j = 1, \ldots, n-1$, the attributes a_1^s and a_n^s solve, respectively,*

$$\max_{a_1 \leqslant a_2^s} \quad \left(1 - \sigma^2(a_1, a_2^s)\right) \tau_1^2 \left(\{a_1, a_2^s\}\right), \qquad \max_{a_n \geqslant a_{n-1}^s} \quad \left(1 - \sigma^2(a_{n-1}^s, a_n)\right) \tau_n^2 \left(\{a_{n-1}^s, a_n\}\right),$$

whereas any attribute $a_j^s \in \{a_2^s, \ldots, a_{n-1}^s\}$ solves

$$\max_{a_{j-1}^s \leqslant a_j \leqslant a_{j+1}^s} \quad \tau_j^2(\{a_{j-1}^s, a_j, a_{j+1}^s\}) \frac{(1 - \sigma^2(a_{j-1}^s, a_j))(1 - \sigma^2(a_j, a_{j+1}^s))}{1 - \sigma^2(a_{j-1}^s, a_{j+1}^s)}.$$

No sampling of irrelevant attributes. Does a single-player sample ever include irrelevant attributes? Even though irrelevant attributes do not enter directly the value of the project, their marginal value need not be zero a priori because they might be strongly correlated with relevant attributes. Nonetheless, Proposition 2 establishes that if extrapolation is local, the player only samples among the relevant attributes.

Proposition 2. *If σ satisfies NAP, any single-player sample consists of relevant attributes only.*

To see the intuition for the suboptimality of irrelevant attributes, suppose there are only two relevant attributes $\underline{a} < \bar{a}$, the value of the project is $v = f(\underline{a}) + f(\bar{a})$, and the player can sample at most one attribute. If attribute $a \in \mathcal{A}$ is sampled, the sample weight is $\tau(a) = \sigma(a, \underline{a}) + \sigma(a, \bar{a})$ and the corresponding posterior variance is $\psi^2(a) = \left(\sigma(a, \underline{a}) + \sigma(a, \bar{a})\right)^2$. When unpacked, the posterior variance consists of a sum of (i) the informativeness of the sample attribute about $\underline{a}$, which is $\sigma^2(a, \underline{a})$, (ii) its informativeness about $\bar{a}$, which is $\sigma^2(a, \bar{a})$, and (iii) the covariance of the estimates that the player forms about each relevant attribute, given by $\sigma(a, \underline{a})\sigma(a, \bar{a})$.[16] Such covariance is higher the more informative the sample attribute is about the overlap of the two relevant attributes. Thus, single-player sampling balances informativeness about each relevant attribute with informativeness about the overlap. This is the case for any general attribute covariance.

Now suppose σ satisfies NAP. Sampling an irrelevant attribute in $(\underline{a}, \bar{a})$ renders the realizations of the two relevant attributes independent. The covariance between the estimates equals $\sigma(\underline{a}, \bar{a})$ regardless of which attribute is sampled. So attributes in $(\underline{a}, \bar{a})$ differ only with respect to how informative they are about $\underline{a}$ and $\bar{a}$. Moreover, letting $x \in (\sigma^2(\underline{a}, \bar{a}), 1)$ denote the informativeness

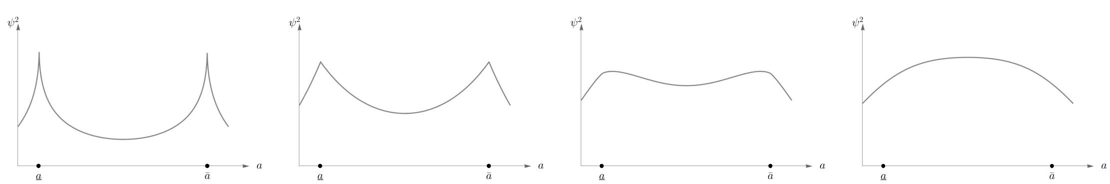
[16] Formally, $\mathbb{E}\left[f(\underline{a}) \mid f(a)\right] = \mu(\underline{a}) + \sigma(a, \underline{a})(f(a) - \mu(a))$ and $\mathbb{E}\left[f(\bar{a}) \mid f(a)\right] = \mu(\bar{a}) + \sigma(a, \bar{a})(f(a) - \mu(a))$. Therefore, the covariance of these two estimates is $\sigma(a, \underline{a})\sigma(a, \bar{a})$.

about one relevant attribute, the informativeness about the other is $\sigma^2(\underline{a}, \bar{a})/x$. The key observation is that the sum $x + \sigma^2(\underline{a}, \bar{a})/x$ is strictly convex in x: there is a first-order loss in the informativeness about one relevant attribute while only a second-order gain in the informativeness about the other. Hence, the player prefers sampling one of the relevant attributes to any irrelevant attributes in between. Reasoning in a similar way, it is also suboptimal to sample attributes to the left of $\underline{a}$ or to the right of $\bar{a}$. The closer any such attribute is to both relevant attributes, the more informative it is about each of them as well as their overlap. The proof of Proposition 2 generalizes this argument to any sampling capacity and any attribute weights.

3.2 Single-player sampling for powered-exponential covariances

We now turn to whether sampling irrelevant attributes is desirable for the class of powered-exponential covariances. Continuing with the stylized example, let the relevant attributes be $\underline{a}$ and $\bar{a}$, weighted equally, and the sampling capacity be one. Because attribute covariance weakens with distance, it is immediate that it is suboptimal to sample irrelevant attributes to the left of $\underline{a}$ or to the right of $\bar{a}$: the closer a is to both relevant attributes, the stronger is its correlation to each of them, hence the greater is the corresponding posterior variance $(\sigma_p(a, \underline{a}) + \sigma_p(a, \bar{a}))^2$.

It is less trivial whether the player might prefer to sample in the interval $(\underline{a}, \bar{a})$. Figure 3 plots the posterior variance induced from sampling an attribute for different values of p. As to be expected from Proposition 2, sampling either $\underline{a}$ or $\bar{a}$ is optimal in the case of the OU covariance, which satisfies NAP. However, plot (a) suggests that, though sufficient, NAP is not necessary: in this stylized example, sampling an irrelevant attribute is in fact suboptimal for any $p \leqslant 1$. The covariance of the estimates that the player holds for the relevant attributes $\sigma_p(\underline{a}, a)\sigma_p(a, \bar{a})$ is strictly below $\sigma_p(\underline{a}, \bar{a})$, so it is maximized at either $a = \underline{a}$ or $a = \bar{a}$. Moreover, the sum of the informativenesses $\sigma_p^2(\underline{a}, a) + \sigma_p^2(a, \bar{a})$ continues to be strictly convex in a, since the covariance σ_p is strictly convex in distance for $p \leqslant 1$. Sampling close to a relevant attribute incurs a first-order loss in informativeness about that attribute but only a second-order gain in informativeness about the other relevant attribute. Hence, an irrelevant attribute contributes neither to higher informativeness about each relevant attribute nor to stronger covariance between their estimates.



(a) $p = 1/2$ *(b) $p = 1$ (NAP)* *(c) $p = 3/2$* *(d) $p = 2$*

Figure 3: Posterior variance of singleton sample $\{a\}$ given $\mathcal{A} = [0, 1]$, $\underline{a} = 1/10$, $\bar{a} = 9/10$, $\ell = 3/5$.

In contrast, plots (c) and (d) show that sampling an irrelevant attribute between $\underline{a}$ and $\bar{a}$ can be optimal for $p > 1$. On the one hand, $\sigma_p(\underline{a}, a)\sigma_p(a, \bar{a})$ is strictly above $\sigma_p(\underline{a}, \bar{a})$ for such attribute covariances. This means that the realizations of the relevant attributes are negatively correlated ex post, which reduces the residual uncertainty about the project. On the other hand, the sum of the informativeness for each relevant attribute is not necessarily convex. Since the attribute covariance σ_p for $p > 1$ is concave for small distances and convex for large ones, sampling in a neighborhood of a relevant attribute induces a first-order gain in the informativeness about the more distant relevant attribute while only a second-order loss in the informativeness about the closer one. These two forces in tandem make it desirable to sample irrelevant attributes in $(\underline{a}, \bar{a})$. Proposition 3 formalizes these observations.

Proposition 3. *Fix $k = 1$. Let the attribute covariance be a powered-exponential covariance. Let $\omega(a) = 1$ if $a \in \{\underline{a}, \bar{a}\}$, where $\underline{a} < \bar{a}$, and $\omega(a) = 0$ otherwise.*

(i) For any $p \in (0, 1]$, the single-player sample consists of a relevant attribute, i.e., $a^s \in \{\underline{a}, \bar{a}\}$.

(ii) For any $p \in (1, 2]$, the single-player sample consists of an irrelevant attribute $a^s \in (\underline{a}, \bar{a})$. For any fixed $p \in (1, 2]$, a^s converges to either relevant attribute as $\ell \to 0$; $a^s = (\underline{a} + \bar{a})/2$ for ℓ sufficiently large; and the distance of a^s from $(\underline{a} + \bar{a})/2$ decreases in ℓ. For any fixed $\ell > 0$, a^s converges to either relevant attribute as $p \downarrow 1$.

Sampling an irrelevant attribute not only might be desirable, but it is in fact generically so for any $p > 1$ and any degree of attribute correlation $\ell > 0$. As the attribute correlation weakens, it becomes more difficult to extrapolate from an irrelevant attribute to a relevant one. The overlap between the relevant attributes shrinks, and so does the extent to which an irrelevant attribute can be informative about this overlap. Irrelevant attributes that are far from either relevant attribute lose their appeal. In response, the the player samples more closely to a relevant attribute.

Single-player sampling for the OU covariance. It is generally intractable to obtain closed-form expressions for the posterior variance and the single-player sample for any powered-exponential covariance. However, the case of $p = 1$ is fully tractable. In what follows, we characterize which relevant attributes are sampled among an interval of equally weighted attributes.

Assumption 2. *Let $\omega(a) = 1$ if $a \in [0, 1]$ and $\omega(a) = 0$ otherwise.*

Thanks to Corollary 1, the sample weights can be obtained in closed form. All sample weights are positive: a higher realization for any sample attribute is good news for the project. Because extrapolation is local, the closer a sample attribute is to neighboring sample attributes, the smaller is the set of out-of-sample attributes about which it is informative, and hence the lower is its sample weight. Extrapolation through the sample realizations is non-linear and convex towards the prior mean μ, as illustrated in Figure 1b. This tendency towards the prior mean, which is due to the mean-reverting property of the OU process, is stronger the less correlated the attributes are.

For $k = 1$, the player optimally samples $a_1^s = 1/2$. All attributes are equally uncertain, so the most informative attribute is naturally the one that is on average closest to all other attributes. Proposition 4 generalizes this intuition. The unique single-player sample has a simple structure for any sampling capacity: it is symmetric around $a = 1/2$, any two adjacent sample attributes are equidistant, and all sample attributes are equally weighted in the posterior value.

Proposition 4. *Let Assumption 2 hold. The single-player sample $\mathbf{a}^s = \{a_1^s, \ldots, a_k^s\}$ is unique, and it is:*

(i) symmetric with respect to the median attribute, i.e., $a_j^s = 1 - a_{k-j+1}^s$ for any j;

(ii) the unique solution to the system of equations

$$1 - e^{-a_1^s/\ell} = \tanh\left(\frac{1-2a_1^s}{2\ell(k-1)}\right), \quad a_j^s = a_1^s + (j-1)\frac{1-2a_1^s}{k-1}; \tag{11}$$

(iii) such that all sample realizations are equally weighted, i.e., $\tau_j(\mathbf{a}^s) = \tau_m(\mathbf{a}^s)$ for $j, m = 1, \ldots, k$.

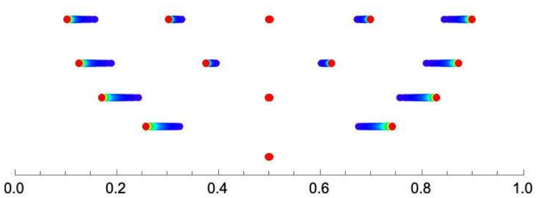


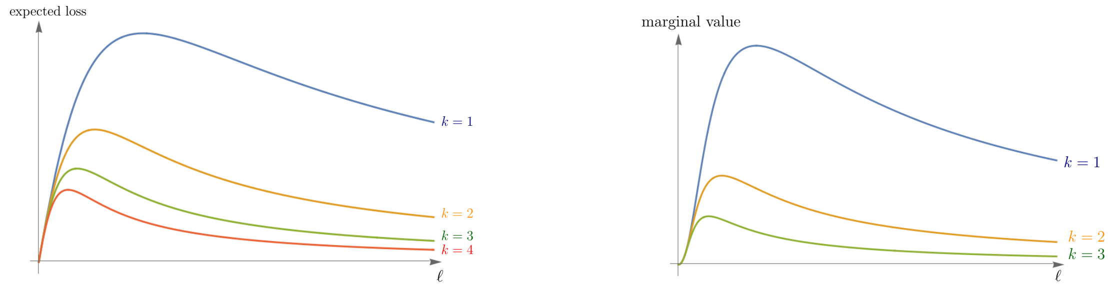
Figure 4: The single-player sample for the OU covariance ($p = 1$), $k \in \{1, 2, 3, 4, 5\}$ (from bottom to top), and $\ell \in (0.02, 2)$ in increments of 0.01 (from blue to red).

This explicit characterization of the single-player sample affords an understanding of how the sample varies with attribute correlation. As the attribute space becomes more correlated, it becomes easier to extrapolate from a sample and the sample attributes become more redundant. To compensate, the single-player sample expands to more peripheral attributes (see Figure 4). However, as attributes approach perfect correlation (resp., independence), the limit single-player sample remains bounded away from the most peripheral attributes $a = 0$ and $a = 1$ (resp., median attribute $a = 1/2$).

Proposition 5. *For any $j \in \{1, \ldots, k\}$, the single-player sample is such that the distance $|a_j^s - 1/2|$ increases in ℓ and the sample converges to $a_j^s \to j/(k+1)$ as $\ell \to 0$ and to $a_j^s \to (2j-1)/(2k)$ as $\ell \to +\infty$.*

Finally, the value of sampling k attributes is single-peaked in the degree of attribute correlation. The prior uncertainty about the value of the project, given by $\text{var}[v] = 2\ell\left(\ell(1 + e^{-1/\ell}) + 1\right) \in (0, 1)$, increases in such correlation.[17] If attributes are either close to independent or close to perfectly correlated, the player's expected loss from a sample is small: in the former case, because the prior uncertainty is negligible, whereas in the latter because one attribute is enough to learn the value almost perfectly. The player is worst off when evaluating a project of moderately correlated attributes (Figure 5a). Moreover, it is for moderately correlated attributes that the marginal value of an additional unit of sampling capacity is highest (Figure 5b).

[17] It approaches zero as $\ell \to 0$ since the uncertainty about the sum of a unit mass of independent attributes is zero. It approaches one as $\ell \to +\infty$ since the uncertainty about the project collapses to uncertainty about a single attribute.



(a) Expected loss from $\mathbf{a}^s$ *for capacity* k *(b) Marginal value of sampling from* k *to* $k+1$

Figure 5: The dependence of the value of sampling on attribute correlation

Remark 1 (Attributes as signals)**.** *Attributes are not only constituent components of the value of the project, but also the means through which the player learns about this value.*[18] *Because of this, the attribute covariance determines simultaneously both (i) how uncertain the value* v *is ex ante, and (ii) how informative any sample of attributes is about* v*. If attributes were instead modeled as signals that are informative about but do not enter directly the value of the project, the attribute covariance would determine (ii) but not (i). This modeling choice—whether attributes enter the value directly or not—matters for how the value of sampling varies with attribute correlation and for the form that the single-player sample takes.*

In contrast to the non-monotonicity in the value of sampling attributes, Clemen and Winkler (1985) show that the value of sampling signals that do not enter directly the value of the project is monotone in the correlation across signals. The stronger such correlation is, the more redundant the signals are. Therefore, both the value of a sample of signals and the marginal value of an additional signal decrease with correlation. Moreover, the player would optimally sample signals that are as independent as possible from each other so as to minimize the redundancy in their informativeness about v*. For an example that stays close to our framework, suppose the value is distributed as* $v \sim \mathcal{N}(\nu_0, 1)$ *and the player has access to a unit mass of correlated signals* $a \in [0,1]$ *with realizations* $f(a) = v + \xi(a)$*, where the noise term* $\xi(a)$ *follows a zero-mean OU process. The player would optimally sample the least correlated signals:* $\{0,1\}$ *for* $k = 2$*,* $\{0, 1/2, 1\}$ *for* $k = 3$*,* $\{0, 1/3, 2/3, 1\}$ *for* $k = 4$ *and so on. That is, the single-player sample not only looks quite different from the one characterized in Proposition 4, but is also independent of* ℓ*, the strength of signal correlation.*

4 Strategic attribute sampling

4.1 General characterization

Upon observing $f(\mathbf{a})$, the principal takes the decision that minimizes her expected loss. As in the single-player benchmark, this optimal decision corresponds to her posterior value $d^*(f(\mathbf{a})) = \nu_P(\mathbf{a})$.

[18] In this regard, our modeling of attributes is akin to how bidders' signals are modeled in interdependent value auctions. Klabjan, Olszewski and Wolinsky (2014) remark that attributes can be reinterpreted as signals or sources of information, the distributions of which depend on the value; this is consistent with our Theorem 1.

Knowing this, the agent seeks a sample of attributes for which the principal's posterior value predicts well his value v_A. That is, he solves

$$\max_{\mathbf{a}\in\mathcal{A}_k} \quad \mathbb{E}\left[-\left(\nu_P(\mathbf{a}) - v_A\right)^2\right]. \tag{12}$$

The agent's value v_A and the principal's decision $\nu_P(\mathbf{a})$ are jointly Gaussian. The sample choice determines both the variance of $\nu_P(\mathbf{a})$, i.e., how informative the sample attributes are for the principal, and the covariance between $\nu_P(\mathbf{a})$ and v_A, i.e., how well the sample aligns the players' interests. The agent prefers the decision to be as close as possible to his value of the project.

The value of a sample, therefore, is determined by two considerations: how much uncertainty the sample introduces in the principal's decision and how strongly the sample correlates the principal's decision with the agent's value. The agent's expected payoff from a sample simplifies to

$$V_A(\mathbf{a}) = -\left(\nu_0^P - \nu_0^A\right)^2 - \text{var}[v_A - \nu_P(\mathbf{a})] = V_A(\emptyset) + \text{cov}[v_A, \nu_P(\mathbf{a})] - \Big(\psi_P^2(\mathbf{a}) - \text{cov}[v_A, \nu_P(\mathbf{a})]\Big),$$

where $V_A(\emptyset)$ denotes the payoff from no sampling. The added value of the sample $V_A(\mathbf{a}) - V_A(\emptyset)$ reflects the agent's two considerations. The term $\text{cov}[v_A, \nu_P(\mathbf{a})]$ captures the alignment between the players' interests,[19] whereas the term $\psi_P^2(\mathbf{a}) - \text{cov}[v_A, \nu_P(\mathbf{a})]$ captures the part of the uncertainty in the principal's decision that is excessive. The agent prefers the sample to be more informative for the principal only to the extent that its informativeness contributes to greater alignment. If the players' attribute weights coincide as in section 3, the term $\text{cov}[v_A, \nu_P(\mathbf{a})]$ equals the players' shared posterior variance, whereas the second term $\psi_P^2(\mathbf{a}) - \text{cov}[v_A, \nu_P(\mathbf{a})]$ vanishes.

Theorem 2 characterizes the agent's optimal sampling problem. Due to quadratic loss, the trade-off between the two relevant statistics $\text{cov}[v_A, \nu_P(\mathbf{a})]$ and $\psi_P^2(\mathbf{a})$ is linear and independent of the players' prior values for the project. It is immediate that any non-empty optimal sample must induce strictly positive covariance between the principal's decision and the agent's value, for otherwise the agent would be better off forgoing sampling altogether. Theorem 2 rewrites the two sufficient statistics in terms of the players' sample weights, which reduces the agent's optimization over the space of all feasible samples of at most k attributes to optimization over a two-dimensional space of feasible values for the pair of sufficient statistics.

Theorem 2. *Any optimal sample solves*

$$\mathbf{a}^* \in \arg\max_{\mathbf{a}\in\mathcal{A}_k} \quad 2\alpha_2(\mathbf{a}) - \alpha_1(\mathbf{a}), \tag{13}$$

where $\alpha_1(\mathbf{a}) \coloneqq \psi_P^2(\mathbf{a})$ *is the posterior variance for the principal given by* (10)*, whereas*

$$\alpha_2(\mathbf{a}) \coloneqq \sum_{j=1}^{n} \tau_j^P(\mathbf{a})\tau^A(a_j) \tag{14}$$

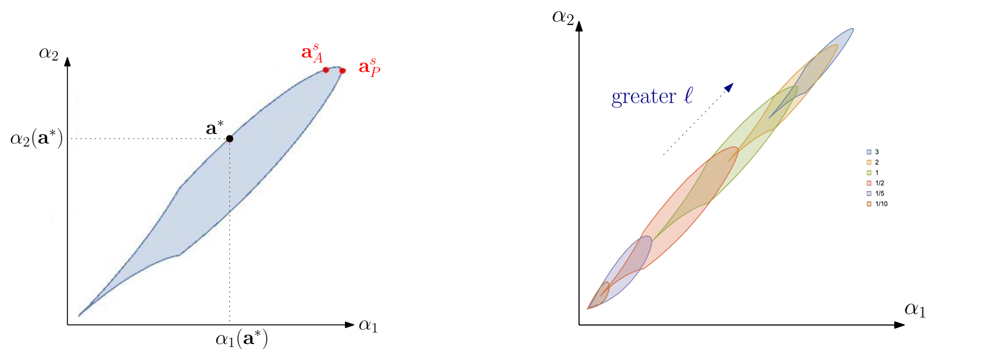
is the covariance between the principal's posterior value $\nu_P(\mathbf{a})$ *and the agent's value* v_A*. The set of optimal samples is independent of* $\left(\nu_0^P, \nu_0^A\right)$*.*

[19] This also corresponds to the covariance in the players' posterior values: $\text{cov}[v_A, \nu_P(\mathbf{a})] = \text{cov}[\nu_A(\mathbf{a}), \nu_P(\mathbf{a})]$.

Feasible region of sufficient statistics. The underlying attribute covariance determines the set of feasible pairs of (α_1, α_2) that are attainable by samples in $\mathcal{A}_k$. This set is given by

$$\mathcal{S}_k \coloneqq \{(\alpha_1', \alpha_2') : \exists \mathbf{a} \in \mathcal{A}_k \text{ such that } \alpha_1(\mathbf{a}) = \alpha_1' \text{ and } \alpha_2(\mathbf{a}) = \alpha_2'\}.$$

By (13), for any fixed α_1, the added value of the sample is increasing in α_2. Hence, it is sufficient for the agent to consider the subset $\bar{\mathcal{S}}_k = \left\{(\alpha_1, \bar{\alpha}_2) \in \mathcal{S}_k : \bar{\alpha}_2 = \max_{(\alpha_1,\alpha_2)\in\mathcal{S}_k} \alpha_2\right\}$. This simplifies the sampling problem to that of maximizing $2\alpha_2 - \alpha_1$ over pairs $(\alpha_1, \alpha_2) \in \bar{\mathcal{S}}_k$. Figure 6a illustrates $\mathcal{S}_k$ for the OU covariance. The optimal sample $\mathbf{a}^*$ corresponds to the point at which the tangent line of slope $1/2$ touches $\bar{\mathcal{S}}_k$. Figure 6b illustrates how $\mathcal{S}_k$ shifts upwards, in terms of both sufficient statistics, as attribute correlation strengthens. It shrinks to a single pair as attributes approach either independence or perfect correlation—at either extreme, all samples are equally valuable. It is for intermediate degrees of attribute correlation that $\mathcal{S}_k$ appears to be largest.



(a) The optimal sample $\mathbf{a}^*$ (black) and the single-player samples $\mathbf{a}_A^s, \mathbf{a}_P^s$ (red).

(b) The set $\mathcal{S}_2$ as a function of $\ell \in \{1/10, 1/5, 1/2, 1, 2, 3\}$.

Figure 6: Feasible set of (α_1, α_2) for the OU covariance, $k = 2$, $\mathcal{A} = [0, 1]$, $\omega_A(a) = a$, $\omega_P(a) = 1$.

Mild disagreement. Under mild disagreement—one in which the players agree about the relative relevance of the attributes but not about the absolute relevance of the project—the optimal sample, whenever non-empty, coincides with the single-player sample of both players. Because the relative relevance $\omega_i(a)/\omega_i(a')$ of any two attributes a and a' is the same for both players, their single-player samples coincide. So whenever sampling is optimal, a single-player sample is drawn. The only departure from the single-player benchmark is that the agent might refrain entirely from sampling. If the project is sufficiently less important to the agent than to the principal, i.e., if $|\omega_A|$ is smaller than $|\omega_P|$, the agent forgoes all sampling because the principal's decision is too unpredictable from the agent's perspective.

Proposition 6. *Suppose the players agree on the relative relevance of attributes:* $\text{supp}(\omega_A) = \text{supp}(\omega_P)$ *and* $\omega_A(a)/\omega_A(a') = \omega_P(a)/\omega_P(a')$ *for any two relevant attributes* $a, a' \in \mathcal{A}$.

(i) The players' attribute weights satisfy $\omega_A(a) = \omega_0 \omega_P(a)$ *for some* $\omega_0 \in \mathbb{R}$, *and the set of single-player samples is the same for both players.*

(ii) If $\omega_0 \geqslant 1/2$, a sample is optimal if and only if it is a single-player sample. Otherwise, the empty sample is uniquely optimal.

Stronger forms of disagreement over attributes are necessary for attributes other than those in the single-player samples to be optimal. In particular, the players might disagree on which attributes are relevant: those in the overlap $\mathrm{supp}(\omega_A) \cap \mathrm{supp}(\omega_P)$ are *common* attributes, whereas the rest of the relevant attributes are *idiosyncratic* to one player. When does the agent find it optimal to sample common attributes, idiosyncratic attributes, or even attributes irrelevant for both himself and the principal?

These questions are tractable for NAP attribute covariances, for which the marginal value of each attribute in the sample depends only on the neighboring sample attributes.[20] For such covariances, the agent does not benefit from sampling attributes which either him or the principal would put zero sample weight on in their respective posterior values. Attributes that are ignored by either player do not contribute to greater alignment between the decision and the agent's value, which is the only benefit of sampling. In fact, if the agent ignores the attribute but the principal does not, that attribute hurts the agent because it provides information to the principal without increasing alignment. Optimality requires that the agent assigns a non-zero sample weight whenever the principal does.

Corollary 2. *Let σ satisfy NAP. Given any sample $\mathbf{a} \in \mathcal{A}_k$, (i) if $\tau_j^P(\mathbf{a}) = 0$ for some $a_j \in \mathbf{a}$, then $V_A(\mathbf{a}) = V_A(\mathbf{a} \setminus \{a_j\})$, whereas (ii) if $\tau_j^P(\mathbf{a}) \neq 0$ but $\tau_j^A(\mathbf{a}) = 0$ for some $a_j \in \mathbf{a}$, then $V_A(\mathbf{a}) < V_A(\mathbf{a} \setminus \{a_j\})$.*

The corollary also clarifies a form of multiplicity that is disregarded without loss in the analysis. A sample $\mathbf{a}^*$ is *optimal up to zero sample weights* if $\mathbf{a}^*$ solves (13) and $\tau_j^P(\mathbf{a}^*) \neq 0$ for each $a_j \in \mathbf{a}^*$. Hereafter, we refer to a sample as optimal if and only if it is optimal up to zero sample weights.

No sampling of attributes irrelevant for both players. Because extrapolation is local, each player assigns non-zero sample weights only to those irrelevant attributes that are closest to some relevant attribute. In particular, if the principal's set of relevant attributes is a single interval $[\underline{a}_P, \bar{a}_P]$, the principal takes into account at most two irrelevant attributes, one on either side of this interval. This suggests that irrelevant attributes possess limited value as a means of influencing the principal's decision. The following proposition establishes that attributes irrelevant for both players are in fact never sampled. If an attribute is sampled, it must be relevant for one of them.

Proposition 7. *Let σ satisfy NAP and $\mathrm{supp}(\omega_i) = [\underline{a}_i, \bar{a}_i]$ for each i. In any optimal sample $\mathbf{a}^* \in \mathcal{A}_k$, no attribute is irrelevant for both players: $\mathbf{a}^* \subset [\underline{a}_A, \bar{a}_A] \cup [\underline{a}_P, \bar{a}_P]$.*

Though this result mirrors that in Proposition 3 in the single-player benchmark, the intuition for why irrelevant attributes are not strategically valuable is quite different. Suppose each player has a single relevant attribute—say, a_P for the principal and $a_A < a_P$ for the agent—and the agent

[20] Lemma 4 in Appendix B, which is analogous to Proposition 1 for the principal-agent setting, derives the marginal value of each attribute in a sample for the agent.

can sample at most one attribute. He chooses between sampling one of the two relevant attributes or sampling an irrelevant one. First, sampling an irrelevant attribute strictly to the left of a_A is suboptimal. The covariance between the decision and the agent's value that such sampling induces is $\sigma(a, a_A)\sigma(a, a_P)$, whereas the informativeness for the principal is $\sigma^2(a, a_P)$. The closer a is to a_A, the more strongly the decision covaries with the agent's value and the more informative sampling is for the principal. However, the covariance between the decision and the agent's value increases faster because $\sigma(a, a_P) = \sigma(a, a_A)\sigma(a_A, a_P) < \sigma(a, a_A)$. The agent is strictly better off sampling a_A instead. Second, sampling any attribute between a_A and a_P leads to the players' values being independent conditional on the sample realization. Hence, the covariance between the decision and the agent's value is the same for any $a \in (a_A, a_P)$; the sample choice cannot affect the alignment among the players. Because all else equal the agent prefers an attribute that is less informative for the principal, he is better off sampling a_A than sampling any $a \in (a_A, a_P)$.

Third, it is suboptimal to sample an irrelevant attribute that is further away from the agent's relevant attribute than the principal's relevant attribute. The agent's expected payoff from any such $a > a_P$ is $\sigma^2(a_P, a)\left(2\sigma(a_A, a_P) - 1\right)$. Sampling in this region is beneficial if and only if $\sigma(a_A, a_P) \geqslant 1/2$, i.e., if the relevant attributes are not too far apart. But whenever this is the case, the agent's payoff decreases in a because $\sigma^2(a_P, a)$ decreases in a. The covariance between the agent's value and the principal's decision is more sensitive than the informativeness of the sample for the principal. The agent is better off sampling a_P instead.

Minimal sampling of idiosyncratic attributes. In a similar vein, attributes relevant to only one player are of limited value to the agent. Depending on how the principal's and the agent's intervals of relevant attributes overlap, each player pays attention to at most two idiosyncratic attributes of the other player. This observation, along with Corollary 2, implies that the optimal sample can feature at most two idiosyncratic attributes—the rest must be common attributes. If the players share no common attributes, the optimal sample consists of at most two attributes, which are idiosyncratic by necessity: one in $\text{supp}(\omega_A)$ and the other in $\text{supp}(\omega_P)$. Therefore, larger sampling capacities $k > 2$ have no added value for the agent.

Proposition 8. *Let σ satisfy NAP and $\text{supp}(\omega_i) = [\underline{a}_i, \bar{a}_i]$ for each i.*

(i) In any optimal sample $\mathbf{a}^* \in \mathcal{A}_k$, at most two sample attributes are idiosyncratic.

(ii) If $[\underline{a}_A, \bar{a}_A] \cap [\underline{a}_P, \bar{a}_P] = \emptyset$, for any capacity $k \geq 1$, any optimal sample consists of at most two attributes, at most one of which is relevant for each player.

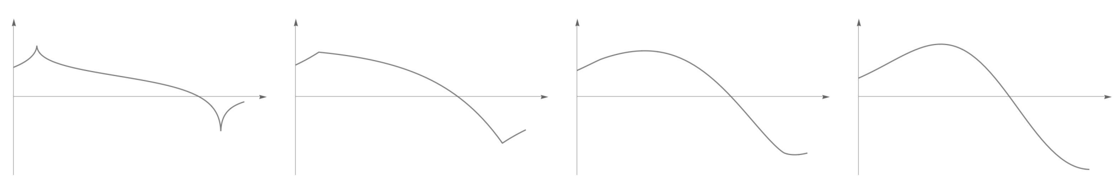
Maximal sampling of common attributes. Proposition 8 established that if the optimal sample has $n^* \geqslant 2$ attributes, at least $(n^* - 2)$ of those must be common. However, if the players also assign the same weights to common attributes—if these attributes are truly common, in the sense that the players agree not only that these attributes are relevant, but also about the extent to which they are relevant—more can be said about the number of common attributes that are optimally sampled. If the agent finds it optimal to sample more than one common attribute, then the optimal sample must exhaust the sampling capacity by including at least $(k-2)$ common attributes.

Proposition 9. *Let σ satisfy NAP. Suppose that* $\text{supp}(\omega_i) = [\underline{a}_i, \bar{a}_i]$ *for each i and $\omega_A(a) = \omega_P(a)$ for any common attribute $a \in [\underline{a}_A, \bar{a}_A] \cap [\underline{a}_P, \bar{a}_P]$. Any optimal sample* $\mathbf{a}^*$ *either (i) does not contain any common attributes and $|\mathbf{a}^*| \leqslant 2$, (ii) or it contains exactly one common attribute and $|\mathbf{a}^*| \leqslant 3$, or (iii) it contains at least $(k-2)$ common attributes and $|\mathbf{a}^*| = k$.*

Put differently, if the sampling capacity is large, that is, $k \geqslant 4$, and the optimal sample uses the entire capacity, then it is guaranteed that the majority of the sample attributes are common. The key observation is that if the sample already includes at least two common attributes, adding a common attribute between those already present in the sample increases α_1 and α_2 by exactly the same amount. The newly added attribute does not introduce any excessive uncertainty in the principal's decision. This is because the players assign the same sample weight to the newly added attribute: this attribute is informative only about other common attributes in its neighborhood and the players weigh common attributes in the same way. Therefore, the marginal value of the newly added attribute equals the increase in the covariance between the decision and the agent's value.

4.2 Strategic sampling for powered-exponential covariances

To see how the sampling of irrelevant attributes can be suboptimal if NAP is violated, we return to a stylized example for the class of powered-exponential covariances. Each player has a single relevant attribute a_i with unit weight, and the agent is constrained to sample at most one attribute. For the case of $p = 1$, for which NAP is satisfied, we know from Proposition 8 that the agent samples either a_A or a_P—in fact it is immediate to show that he samples his own relevant attribute a_A (Figure 7b). As it turns out, this is the case more generally for any $p \leqslant 1$ (Figure 7a). For $p > 1$, however, the optimal attribute is irrelevant for both players, and it is closer to the relevant attribute of the agent than that of the principal (Figure 7c-7d).



(a) $p = 1/2$ *(b) $p = 1$* *(c) $p = 3/2$* *(d) $p = 2$*

Figure 7: Agent's added payoff $2\alpha_2(a) - \alpha_1(a)$ from sampling $\{a\}$ given $\mathcal{A} = [0, 1]$, $a_A = 1/10$, $a_P = 9/10$, $\ell = 3/5$.

Proposition 10. *Suppose $k = 1$ and the attribute covariance is a powered-exponential covariance σ_p. Suppose $\omega_i(a_i) = 1$ for some $a_i \in \mathcal{A}$ and $\omega_i(a) = 0$ for any $a \neq a_i$.*

(i) For any $p \in (0, 1]$, the optimal sample consists of $a^* = a_A$.

(ii) For any $p \in (1, 2]$, the optimal sample consists of an attribute a^* strictly between a_A and $(a_P + a_A)/2$, which is irrelevant for both players. For any fixed $p \in (1, 2]$, a^* converges to

$(a_P + a_A)/2$ as $\ell \to 0$ and it converges to a_A as $\ell \to +\infty$. For any fixed $\ell > 0$, a^* converges to a_A as $p \downarrow 1$.

The key difference between attribute covariances with $p \leqslant 1$ and those with $p > 1$ rests on whether irrelevant attributes between a_A and a_P align the principal's decision with the agent's value. As the sample shifts further away from the agent's relevant attribute and closer to the principal's, the agent's sample weight $\sigma_p(a_A, a)$ decreases and the principal's sample weight $\sigma_p(a, a_P)$ increases. However, for $p \leqslant 1$, the decrease in one sample weight is too large compared to the increase in the other; as a result, the alignment $\alpha_2(a) = \sigma_p(a_A, a)\sigma_p(a, a_P)$ is below the alignment that either relevant attribute attains $\alpha_2(a_A) = \alpha_2(a_P) = \sigma_p(a_A, a_P)$. Sampling closer to the principal's relevant attribute does not lead to stronger alignment, but it costs in terms of greater informativeness for the principal. For $p > 1$, on the other hand, the alignment attained by any irrelevant attribute between a_A and a_P is strictly above that attained by either relevant attribute. As the sample shifts closer to a_P, the gain in alignment is first-order relative to the loss from greater informativeness for the principal. This encourages the agent to sample an irrelevant attribute—a compromise between the players' relevant attributes.

Strategic sampling for the OU covariance. In what follows, we assume that each player has a single interval of equally relevant attributes: $\omega_i(a) = 1$ for any $a \in [\underline{a}_i, \bar{a}_i]$ and 0 otherwise. Let $\Delta_i := \bar{a}_i - \underline{a}_i$ denote the mass of relevant attributes for player i. Proposition D.1 in Appendix D establishes that the optimal sample is non-empty if and only if Δ_A is sufficiently larger than Δ_P, so that the principal's posterior value does not respond as strongly as that of the agent to the realization of any singleton sample.

Strikingly, it is now suboptimal to sample any of the principal's idiosyncratic attributes. This is easiest to see if all relevant attributes are idiosyncratic, so that the optimal sample includes at most one relevant attribute for each player. Consider $\underline{a}_A < a_1^* < \bar{a}_A < \underline{a}_P < a_2^* < \bar{a}_P$. The further to the right a_2^* is, the less aligned the decision is with the agent's value. Sampling such an attribute is justified only if it guarantees lower informativeness for the principal. However, the principal's posterior variance is single-peaked in $[\underline{a}_P, \bar{a}_P]$; due to the presence of a_1^* in the sample, $\underline{a}_P$ is less informative than $\bar{a}_P$ for the principal. This means that the sample $\{a_1^*, \underline{a}_P\}$ is strictly preferred, in terms of both sufficient statistics, to $\{a_1^*, a_2^*\}$. But the sample $\{\bar{a}_A\}$ is even better than $\{a_1^*, \underline{a}_P\}$ because it attains the same alignment while also being less informative for the principal.

Therefore, the agent always samples only one attribute among his relevant attributes, as derived in Proposition 11. The stronger the attribute correlation is, the further away this optimal attribute is from the principal's interval. With almost independent attributes, the agent's dominant concern is generating alignment, hence a_1^* approaches the attribute that is most informative for the principal among the agent's relevant attributes. With highly correlated attributes, on the other hand, suppressing informativeness for the principal takes priority, hence the optimal sample is either empty or it approaches the most distant attribute that the agent is ever willing to sample.

Proposition 11. *Let $[\underline{a}_A, \bar{a}_A] \cap [\underline{a}_P, \bar{a}_P] = \emptyset$. For any $k \geqslant 1$,*

(i) if non-empty, the unique optimal sample is $\mathbf{a}^* = \{a_1^*\}$ with $a_1^* \in [\underline{a}_A, \bar{a}_A]$ given by

$$a_1^* = \begin{cases} \ell \ln\left(e^{\underline{a}_A/\ell} + \frac{e^{\bar{a}_P/\ell} - e^{\underline{a}_P/\ell}}{2}\right) & \text{if } \bar{a}_P < \underline{a}_A \\ -\ell \ln\left(e^{-\bar{a}_A/\ell} + \frac{e^{-\underline{a}_P/\ell} - e^{-\bar{a}_P/\ell}}{2}\right) & \text{if } \bar{a}_A < \underline{a}_P; \end{cases} \tag{15}$$

(ii) if $\mathbf{a}^*$ is non-empty, its distance from $[\underline{a}_P, \bar{a}_P]$ strictly increases in ℓ;

(iii) as $\ell \to 0$, a_1^* approaches $\underline{a}_A$ if $\bar{a}_P < \underline{a}_A$ and it approaches $\bar{a}_A$ if $\bar{a}_A < \underline{a}_P$;

(iv) as $\ell \to +\infty$, then (a) for $\underline{a}_A > \bar{a}_P$, a_1^* approaches $\underline{a}_A + \frac{\Delta_P}{2}$ if $\Delta_A \geqslant \Delta_P/2$ and $\mathbf{a}^*$ is empty otherwise, (b) for $\bar{a}_A < \underline{a}_P$, a_1^* approaches $\bar{a}_A - \frac{\Delta_P}{2}$ if $\Delta_A \geqslant \Delta_P/2$ and $\mathbf{a}^*$ is empty otherwise.

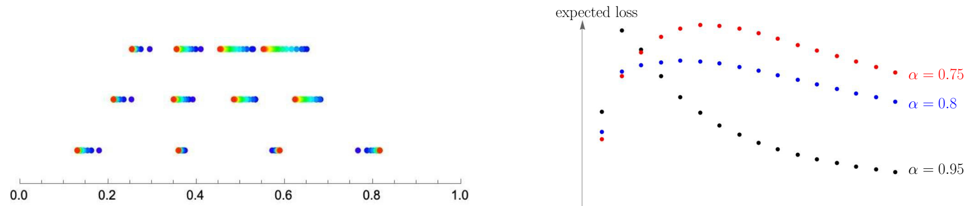
That idiosyncratic attributes are scarcely sampled becomes even clearer in the presence of common attributes. Proposition 9 can be sharpened further for the OU covariance. First, whenever an idiosyncratic attribute is sampled, it is always the agent's, by a similar intuition to that outlined above, and it is moreover the only attribute sampled. By contrast, if the optimal sample is not a singleton—i.e., if the agent is willing to sample among common attributes—the entire sample must consist of common attributes only. The agent resorts to a single idiosyncratic attribute if the principal reacts too strongly to common attributes. This is more likely to be the case if the common attributes are few, the principal's idiosyncratic attributes are much more numerous than those of the agent, and attribute correlation is sufficiently strong. Thus sampling an idiosyncratic attribute is a solution of last resort for the agent.

Proposition 12. *Let $[\underline{a}_A, \bar{a}_A] \cap [\underline{a}_P, \bar{a}_P] \neq \emptyset$. In any non-empty optimal sample $\mathbf{a}^* = \{a_1, \ldots, a_n\}$, either (i) $n = 1 \leq k$ and $a_1^* \in [\underline{a}_A, \bar{a}_A] \setminus [\underline{a}_P, \bar{a}_P]$, or (ii) $n = k$ and all attributes are common: $a_j^* \in [\underline{a}_A, \bar{a}_A] \cap [\underline{a}_P, \bar{a}_P]$ for any $j = 1, \ldots, k$.*

Whenever the optimal sample consists of a single idiosyncratic attribute, the comparative statics in correlation are similar to those in Proposition 11: the optimal attribute moves further away from the principal's relevant attributes. However, how the optimal common attributes vary with attribute correlation is more subtle, since it depends on how large the mass of common attributes is. For a clean illustration, we parametrize $[\underline{a}_A, \bar{a}_A] = [0, \alpha]$ and $[\underline{a}_P, \bar{a}_P] = [1-\alpha, 1]$ so that $\Delta_A = \Delta_P = \alpha$ and $\alpha \in (1/2, 1]$ is a proxy for the degree of agreement. It is straightforward to show that for any degree of correlation the optimal sample is non-empty and it consists of common attributes only.[21] Figure 8a illustrates the comparative statics.

For $\alpha \approx 1$, the relevant attributes for the two players almost coincide. The optimal sample expands with stronger correlation in a similar fashion to the single-player benchmark. As disagreement increases, though, the agent becomes increasingly more concerned about the variability of the principal's decision. Hence, the entire sample shifts towards $(1-\alpha)$ as correlation becomes stronger, in

[21] It is non-empty because $\tau^A(1-\alpha) > \tau^P(1-\alpha) > 0$, hence $2\alpha_2(\{1-\alpha\}) - \alpha_1(\{1-\alpha\}) > 0$. It consists of common attributes only because $V_A(\{a\}) - V_A(\emptyset) = \ell^2 e^{(a-1)/\ell}(e^{\alpha/\ell} - 1)\left(4 + e^{(a-1)/\ell} - 2e^{-a/\ell} - 2e^{(a-\alpha)/\ell} - e^{-(1-\alpha-a)/\ell}\right)$ is strictly increasing in $a \in [0, 1-\alpha)$.



(a) Optimal sample for $k = 4$ from $\ell = 0.02$ (blue) to $\ell = 0.8$ (red) and $\alpha \in \{0.75, 0.8, 0.95\}$ from top to bottom.

(b) Expected loss net the prior bias $(\nu_0^P - \nu_0^A)^2$.

Figure 8: The range of correlation is $\ell \in (0.02, 0.8)$ in increments of 0.05 for both panels.

a similar fashion to the comparative statics of Proposition 11 for the singleton sample with idiosyncratic attributes. In the limit as $\ell \to +\infty$, the sample virtually collapses to $\{1 - \alpha\}$. On the other hand, Figure 8b shows the expected loss attained by the optimal sample as a function of ℓ. As in the single-player benchmark, moderate attribute correlation leads to the greatest loss for the agent.

5 Discussion

5.1 Testable implications for product evaluation

Our findings have implications for empirical research on attribute-based product evaluation. Ample evidence from consumer research suggests that, when evaluating a product, consumers draw correlational inferences across its attributes. They rely on their beliefs about how the realizations of any two attributes covary in order to infer unobserved attributes from observed ones.[22] For example, in the context of consumers evaluating automobiles, Mason and Bequette (1998) establish that experienced consumers hold accurate and stable beliefs about the pairwise covariance of the automobile's attributes.[23] Our results speak directly to two central concerns in this literature: the possibility of reversals in the direction of inference and the value of learning irrelevant attributes.

Inference reversals. Does a higher realization of a *desirable* (resp., undesirable) attribute always lead to a *higher* (resp., lower) estimated value for the product, and how does the direction of such inference depend on the presence of other attributes in the sample? Huber and McCann (1982) address this in a simple setting with only two correlated attributes $\{a, a'\}$. They show, both theoretically and experimentally, that the sample weight that the consumer assigns to the realization of attribute a when estimating the value of the product can have the opposite sign to the attribute weight $\omega(a)$. That is, a higher realization of a desirable attribute can decrease the consumer's willingness to pay for the product. This is due to the inference drawn from a to a': if the two attributes are conflicting, in the sense that $\omega(a)\omega(a')\sigma(a, a') < 0$, the direct contribution $\omega(a)$ to the product's estimated value

[22]For recent surveys on correlational inference in attribute-based decisions, see Scheibehenne, von Helversen and Rieskamp (2015) and Walters and Hershfield (2020).

[23]They elicit the consumers' beliefs about how attributes covary by asking questions such as "If an automobile is rated high in net horsepower, how likely is it also to be rated high in acceleration?" for each attribute pair.

goes against the indirect (inferential) contribution $\omega(a')\sigma(a, a')$. Attribute correlation is critical to such inference reversal, which would not be possible with independent attributes. With a larger attribute space, this insight becomes both more general and more nuanced.

First, as in the two-attribute setting above, the sample weight of an attribute need not have the same sign as its attribute weight, even if that were the only attribute sampled. The sample weight of $\mathbf{a} = \{a_1\}$ is

$$\tau(a_1) = \int_{\mathcal{A}} \omega(a)\sigma(a, a_1)\, \mathrm{d}a,$$

which reflects the inferences that the player draws from a_1 to all other relevant attributes. This generalizes the notion of conflicting attributes in Huber and McCann (1982) to more than two attributes: a_1 conflicts with the rest of the attribute space if $\omega(a)\sigma(a, a_1) < 0$ for a sufficiently large number of relevant attributes. Example E.1 in Online Appendix E illustrates such reversal in a setting with perfectly correlated attributes in which high realizations for some attributes imply low realizations for others. A desirable attribute that is negatively correlated with sufficiently many other relevant attributes is weighted negatively even if it is the only attribute sampled.

With more than two attributes a second form of inference reversal becomes possible: an attribute can be weighted positively (resp., negatively) in the singleton sample but negatively (resp., positively) in a larger sample. Whether an attribute realization implies a higher or lower willingness to pay for the product depends on other observed attributes. Example E.2 illustrates such reversal in a setting with only desirable attributes and an attribute covariance that violates NAP. Any singleton sample carries a positive sample weight, but there exist two-attribute samples for which one of the sample weights is negative. The negative weight arises because the sample features excessive redundancy. Whereas the first form of inference reversal is consistent with the evidence in Huber and McCann (1982), this second form is novel and remains to be tested in future research in settings with more than two attributes. Moreover, our analysis sheds light on properties of the attribute covariance that make such reversals possible: in settings in which the relevant attributes are either all desirable or all undesirable, an inference reversal is possible *only if* the attribute covariance violates NAP.[24]

For strategic settings in which a non-unitary consumer evaluates the product—say, one household member examines attributes and another decides on whether to purchase the product—Theorem 2 suggests that the direction of inference from the optimal sample is on average the same for the two players, because $\tau_j^P(\mathbf{a})$ must have the same sign as $\tau_j^A(\mathbf{a})$ when averaged properly across all sample attributes. More can be said if the underlying attribute covariance satisfies NAP: by Lemma 4, for any optimally observed attribute, a higher realization is either good news for both players' values or bad news for both. How consumers draw attribute inferences in such strategic evaluation is a promising direction for future research on product evaluation.

[24]This is an immediate implication of Corollary 1.

The value of irrelevant attributes. What can be inferred about the consumer's preference for certain attributes from observing the attributes that she focuses on when evaluating the product? With independent attributes, the consumer optimally focuses on attributes that are directly relevant; so it is reasonable to infer that attributes that are observed enter directly the consumer's value for the product. However, this need not be the case in the presence of attribute correlation. Our results clarify that whether extrapolation across attributes is local is an important determinant of whether the consumer decides to learn from irrelevant attributes. If the attribute covariance satisfies NAP in a given empirical setting, then it is reasonable to conclude that the observed attributes are directly relevant for the consumer.[25] This holds for product evaluation by a non-unitary consumer as well: the observed attributes must be relevant for at least one of the players involved in the evaluation.

A second insight that emerges from our analysis is that there exist attribute covariances for which the consumer responds more strongly after observing an irrelevant attribute than a relevant one. This stands in contrast to the finding in Mason et al. (2001) that observing a relevant attribute leads to a more precise product rating than observing an irrelevant one. We provide a theoretical justification for why a Bayesian consumer might be highly sensitive to irrelevant attributes when extrapolation from observed attributes is non-local.

5.2 A similarity interpretation of attribute correlation

Because σ encodes the pairwise correlation across attributes, it is a natural measure of similarity between attributes.[26] When extrapolating to out-of-sample attributes, the player reasons by similarity. Each sample attribute is weighted by its average similarity to out-of-sample attributes, while also accounting for its similarity to the rest of the sample.[27] First, how similar the attributes of a project are dictates how uncertain the project is in its entirety. The ex ante uncertainty about the project's value v can be rewritten as

$$\int_{\mathcal{A}}\int_{\mathcal{A}} \sigma(a,a')\omega(a)\omega(a')\,\mathrm{d}a\,\mathrm{d}a' = \mathbb{E}_{a,a'\sim\omega}[\sigma(a,a')]. \tag{16}$$

The right-hand side corresponds to the similarity of a randomly drawn pair of attributes, where attributes are drawn independently in the pair and the probability of attribute a being drawn is $\omega(a)$.[28] The more similar the relevant attributes are on average, the more their realizations tend to

[25]Beyond *whether* the observed attributes are relevant, Proposition 1 has implications for *the extent to which* they are relevant: all else equal, the relevance of attributes within a neighborhood of an observed attribute must be significant in order for the sample weight of the observed attribute to be significant. The consumer benefits little from sampling an attribute which is of negligible relevance and the attributes close to which are also of negligible relevance. In the context of strategic sampling, Lemma 4 implies that attributes with negligible sample weights for one of the players will not be optimally sampled—their marginal value is too low.

[26]Given covariance σ, a dissimilarity metric can be defined as $d_\sigma(a,a') := \sqrt{\frac{1}{2}(1-\sigma(a,a'))} \in [0,1]$. The higher (resp., lower) $\sigma(a,a')$ is, the more similar (resp., dissimilar) attributes a and a' are.

[27]This is related to case-based reasoning, in particular to statistical procedures such as similarity-weighted averaging in Gilboa, Lieberman and Schmeidler (2006). The key conceptual difference is that here the players know the set of possible mappings f and derive sample weights via Bayesian updating, whereas Gilboa, Lieberman and Schmeidler (2006) posit and axiomatize a particular form for the sample weights without assumptions on the underlying f.

[28]See Lemma 3(ii) for why the attribute weights ω can be interpreted as a well-defined probability density over the attribute space.

move together, hence the more uncertain the project is. Analogously, the ex ante covariance between the principal's and the agent's values is given by

$$\int_{\mathcal{A}}\int_{\mathcal{A}} \sigma(a,a')\omega_A(a)\omega_P(a')\,\mathrm{d}a\,\mathrm{d}a' = \mathbb{E}_{a\sim\omega_A, a'\sim\omega_P}\left[\sigma(a,a')\right], \tag{17}$$

which corresponds to the similarity between a pair of random attributes, one drawn from density ω_A and the other from ω_P. Intuitively, the more similar the attributes that are relevant to the agent are to those of relevance to the principal, the more the players' values for the project covary together.

Second, after a sample $\mathbf{a}$ is discovered, the similarity across the remaining attributes is updated to $\sigma_{\mathbf{a}}(a,a') = \sigma(a,a') - \Big(\sigma(a,a_1),\dots,\sigma(a,a_n)\Big)\,\Sigma^{-1}(\mathbf{a})\,\Big(\sigma(a',a_1),\dots,\sigma(a',a_n)\Big)^{\top}$. The residual uncertainty about the project now corresponds to the average similarity between out-of-sample attributes $\mathbb{E}_{a,a'\sim\omega}[\sigma_{\mathbf{a}}(a,a')]$. Single-player sampling of Theorem 1 leads to the greatest reduction in the average attribute similarity—it makes out-of-sample attributes as dissimilar ex post as possible. As it turns out, for the OU covariance, dividing the attribute space into two equally sized intervals minimizes such attribute similarity. Any singleton sample $\{a\}$ divides the attribute space $\mathcal{A} = [0,1]$ into two independent attribute intervals $[0,a]$ and $[a,1]$. The larger one interval is relative to the other, the more similar the attributes are on average ex post.

The sample weight $\tau_j(\mathbf{a})$ quantifies the similarity of attribute $a_j \in \mathbf{a}$ to the rest of the attribute space. It considers not only how similar a_j is to a random attribute, i.e., $\mathbb{E}_{a\sim\omega}[\sigma(a,a_j)]$, but also how similar a_j is to other sample attributes which in turn might be similar to that random attribute:

$$\tau_j(\mathbf{a}) = \sum_{m=1}^{k}\left(\Sigma^{-1}(\mathbf{a})\right)_{j,m}\mathbb{E}_{a\sim\omega}[\sigma(a,a_m)].$$

The $(m,j)^{th}$ entry of $\Sigma^{-1}(\mathbf{a})$—the partial correlation between $f(a_m)$ and $f(a_j)$—captures how similar attributes a_j and a_m are *conditional* on the rest of the sample. Therefore, the sample weight aggregates all the channels of similarity from a_j to a randomly drawn attribute. It reduces to simply $\mathbb{E}_{a\sim\omega}[\sigma(a,a_j)]$ if a_j is independent of other sample attributes.

5.3 Related work

The paper builds on the literature on costly attribute discovery, to which it contributes the modeling of attribute correlation. Klabjan, Olszewski and Wolinsky (2014) study costly discovery from a finite set of independent attributes. They show that if the attributes are ordered in the sense of second-order stochastic dominance (SOSD) and the heterogeneity in discovery costs is small, the optimal sample consists of SOSD-dominated attributes. In contrast, we focus on correlation within a continuum of attributes and abstract away from heterogeneity in costs. There is, however, a common thread between our characterization and theirs: Theorem 1 establishes that the distribution over posterior values induced by the single-player sample is SOSD-dominated.[29] Branco, Sun and Villas-Boas (2012) also explore learning over a continuum of Gaussian attributes of a single object.

[29] In our framework, attributes are not necessarily ordered in the SOSD sense, but the distributions over posterior values induced by attribute samples are.

However, their attributes—modelled as infinitesimal Brownian increments—are independent and are discovered in an exogenously fixed order.[30] Ke, Shen and Villas-Boas (2016) extend this framework to search over several multi-attribute products, in which there is correlation across products but not across attributes of the same product. Neeman (1995), Olszewski and Wolinsky (2016), Sanjurjo (2017), and Geng, Pejsachowicz and Richter (2017) also consider search across several objects with independent attributes. In recent work, Liang, Mu and Syrgkanis (2021) study dynamic costly learning from a finite set of correlated Gaussian attributes. If the attribute space is arbitrarily large, their main characterization holds only if attributes are approximately independent. By contrast, we have a flexible covariance over a continuum of attributes but no temporal costs. Second, they study sampling by a single forward-looking agent, whereas we focus on strategic sampling.

Second, our analysis of strategic sampling is related to a large literature on agency frictions in information production, in particular to models of persuasion through flexible information design (Brocas and Carrillo (2007), Rayo and Segal (2010), Kamenica and Gentzkow (2011)), Wald persuasion games (Henry and Ottaviani (2019)), and persuasion with multi-dimensional information (Glazer and Rubinstein (2004), Sher (2011)). We share with these papers the assumptions of (i) commitment in information disclosure and (ii) symmetrically informed players but asymmetric authority to produce information. The attribute structure, however, introduces two peculiar features. First, in contrast to flexible information design, the correlation structure constrains and shapes the set of posterior belief distributions that can be generated through sampling. The elegant tools of information design cannot be invoked in this setting. Second, the preferences of both the principal and the agent are attribute-dependent. Another important line of work has focused on the incentives to produce and disclose hard evidence (Shavell (1994), Che and Kartik (2009), Eliaz and Frug (2018)). We abstract away from disclosure incentives in order to tackle the question of *which* attributes are discovered in the presence of attribute correlation. (Di Tillio, Ottaviani and Sørensen, 2017*a*,*b*) also study strategic sample selection, but they do so in a context with state-independent preferences and disclosure of private signals. In contrast to Banerjee et al. (2020), our framework features a single principal and there is no value from randomization over samples. Our analysis also relates to Hirsch (2016) in the agent's tradeoff between learning and influence, however disagreement here arises due to different attribute weights rather than beliefs.

The paper also contributes to a literature starting with Aghion et al. (1991) on optimal experimentation over a rich space of information sources. A productive approach in this literature has modelled unknown payoffs from experimentation through a Brownian sample path (Jovanovic and Rob (1990), Callander (2011), Garfagnini and Strulovici (2016), Callander and Hummel (2014)). Our model departs from this work in two respects. First, the Brownian motion is a special case of the class of Gaussian processes that we introduce.[31] Two recent exceptions are Ilut and Valchev (2023)

[30] Branco, Sun and Villas-Boas (2012) briefly discuss the case of correlated attributes. Notably, in their framework, the case of correlated attributes simplifies to that of independent attributes in which attribute variance decreases the latter the attribute is in the fixed order.

[31] Bardhi (2018) relied on such a Brownian structure.

and Bardhi and Bobkova (2023), which model payoffs through other Gaussian processes. Second, the payoff-relevant statistic here is the area under the sample path rather than a maximum of the path—indeed, this crucial feature differentiates attribute discovery from search. In this respect our model relates to Callander and Clark (2017), Ilut and Valchev (2023), and Bardhi and Bobkova (2023), in that players' payoff depends on the entire sample path. However, in Ilut and Valchev (2023) the problem is not one of sampling; in Callander and Clark (2017) the higher court samples legal cases, but the nature of the adoption problem faced by the lower court differentiates this model from attribute learning[32]; and in Bardhi and Bobkova (2023) each citizen can sample only his own local evidence, which corresponds to a single attribute.

5.4 Robustness and extensions

Multidimensional / finite attribute space. Section 2.1 models the attribute space as a continuum of one-dimensional attributes. The continuum of attributes allows for a continuous covariance—which makes transparent the role of attribute correlation—and fine-tuning in sampling. The one-dimensional attribute space allows for a clean statement of the nearest-attribute property. However, a finite and/or unordered attribute space might be natural when inspecting the attributes of an automobile (e.g., horsepower, color, weight etc.). Moreover, multidimensional attributes are natural, for instance, when an employer seeks to learn a worker's performance across a variety of hypothetical situations that might arise on the job and that invoke a combination of a hard skill $h \in \mathbb{R}$ and a soft skill $s \in \mathbb{R}$—in such a case, each situation (s, h) is an attribute and $f(s, h)$ is the worker's performance in it.

Our basic methodology allows for a finite attribute space and can be extended easily to the case of a multidimensional attribute space: extrapolation in section 2.2, as well as Theorems 1 and 2 are fully general. They only rely on the joint Gaussian distribution of the attributes, be those one- or multidimensional.[33] The feasible region of (α_1, α_2) can be constructed in a similar way, and it is finite if $\mathcal{A}$ is finite. The framework can also accommodate constraints on which attributes are discoverable: e.g., a finite subset of attributes $\tilde{\mathcal{A}} \subset \mathcal{A}$ within a larger, and potentially infinite, attribute space $\mathcal{A}$ that are accessible for sampling.

General prediction problem. A general decision problem that nests our setup has the following structure: player i obtains ex-post payoff $u(d, v_i)$ where $d \in D$ is a decision made by the principal, D is a set of available decisions, and v_i is the value of the project for player i. E.g., $D = \{0, 1\}$ and $(u(1, v_i), u(0, v_i)) = (v_i, r_i)$ where r_i is a reservation value, or $D = \mathbb{R}$ and $u(d, v_i) = -|d - v_i|$. First, any such problem is a prediction problem: the principal takes a decision based on her posterior value $\nu^P(\mathbf{a})$ and the agent's expected payoff $\mathbb{E}_{v_A, f(\mathbf{a})}\left[u(d^*(\nu^P(\mathbf{a})), v_A)\right]$ is determined by the joint

[32] In Callander and Clark (2017) optimal case selection does not consist of the most uncertain legal case. Similarly, the most uncertain attribute is not the most informative in section 3. See also Remark 1.

[33] The Gaussian structure is consistent with the principle of maximum ignorance: given (μ, σ), the players adopt the least informative (entropy-maximizing) prior distribution over attribute realizations, which is a multivariate Gaussian.

distribution of $\big(v_A, \nu^P(\mathbf{a})\big)$. Because this distribution is Gaussian, the two sample-specific moments $\alpha_1(\mathbf{a})$ and $\alpha_2(\mathbf{a})$ (as defined in Theorem 2) are sufficient statistics for the agent's sampling problem.

Second, the attribute covariance still determines the feasible (α_1, α_2) pairs as discussed in section 4.1. What varies with the decision problem is the tradeoff between α_1 and α_2. For the case of $D = \mathbb{R}$ and $u(d, v_i) = -(d - v_i)^2$, this tradeoff is linear and does not depend on the prior values (ν_0^A, ν_0^P), which makes possible a sharp characterization of optimal sampling. In other decision problems the tradeoff might be more complex, e.g., in the binary decision setup, the tradeoff is non-linear and it depends on (ν_0^A, ν_0^P).[34]

Disagreement about the attribute covariance. An alternative plausible source of conflict between the principal and the agent would be disagreement about the attribute covariance rather than the attribute weights, i.e., $\omega_A = \omega_P = \omega$ but $\sigma_A \neq \sigma_P$. In such a case, the extrapolated mapping in Lemma 1 is different across players and the sample weights in equation (8) are instead given by

$$\tau_j^i(\mathbf{a}) = \int_{\mathcal{A}} \tau_j^i(a; \mathbf{a})\omega(a)\, \mathrm{d}a.$$

Given these sample weights, the characterization in Theorem 2 continues to hold. But note that, as long as $\sigma_P(a, a_j) \neq 0$ for all $a \in \mathcal{A}$ and $a_j \in \mathbf{a}$, the agent's singleton sample weight in equation (14) can be rewritten as

$$\tau^A(a_j) = \int_{\mathcal{A}} \left(\omega(a)\frac{\sigma_A(a, a_j)}{\sigma_P(a, a_j)}\right)\sigma_P(a, a_j)\, \mathrm{d}a =: \int_{\mathcal{A}} \hat{\omega}_{A,j}(a)\sigma_P(a, a_j)\, \mathrm{d}a.$$

In this way, the problem of disagreement about the covariance can be reformulated as one of disagreement about attributes weights, where $\omega_P = \omega$, $\omega_A = \hat{\omega}_{A,j}$, and the common covariance is σ_P. However, the caveat is that the agent's attribute weight is now dependent on the sample attribute, hence the subscript j—it is as if the sample choice has a framing effect on how the agent views the relevance of different attributes.

Project choice and complexity. A promising direction for future work is the modeling of choice between multiple projects in this framework. For instance, a consumer might evaluate two or more competing products before deciding whether to purchase one of them. In this more general framework, product j is described by an attribute space $\mathcal{A}_j$ and an attribute covariance σ_j on $\mathcal{A}_j \times \mathcal{A}_j$. The attribute mappings are drawn independently across products. For instance, each σ_j can be an OU covariance defined on $\mathcal{A}_j = [0, 1]$ with parameter $\ell_j > 0$: the available products vary in their prior uncertainty and in how difficult they are to evaluate. The consumer chooses a sample of attributes to be sampled across all products. An immediate implication of our analysis is that the consumer need not ultimately purchase the product with the highest sample realizations, even if all products start with the same prior value, because the corresponding sample weights might be too small due to weak attribute correlation. Moreover, the consumer's choice of sample balances the informativeness of its attributes across available products; for example, this balances the depth of learning that comes from

[34] An earlier version of this paper considered such a binary decision setup. See Appendix E.2.

discovering the most informative attribute of a specific product with the breadth that comes from learning an attribute that is moderately informative across products. Taking this setup further, this is a natural setup to investigate market competition in product design, i.e., the equilibrium choice of attribute correlation by competing firms (Spiegler, 2016).

Implications for strategic site selection. Evaluation through small-scale pilot studies can be understood as a problem of attribute discovery. Allcott (2015) observed empirically the presence of site selection bias in the initial rollout of an energy conservation program: the pilot sites, which were selected by the utility companies involved in the program, were too similar and higher-performing than the average site. The findings of section 4 offer one rationale for such low external validity. Consider a policymaker who cares about the average impact of a program across all demographic groups $a \in \mathcal{A} = [0, 1]$ but has to rely on pilot sites selected by a program advocate (the utility company), whose focus has traditionally been on a subset of demographic groups $[0, \alpha] \subset [0, 1]$—for example, because these groups are expected to experience higher impact and have been traditionally more profitable for the company. If NAP is a plausible feature of the impact of the program across the population, then Corollary 9 suggests that the advocate will focus his selection of pilot sites in the overlap $[0, \alpha]$. This maximizes alignment between their interests while keeping in check the volatility of the policymaker's evaluation of the program.

A Preliminaries on Gaussian processes

Definition A.1 (Validity of the attribute covariance)**.** The covariance $\sigma(a,a')$ is *symmetric* if $\sigma(a,a')=\sigma(a',a)$ for any $(a,a')\in\mathcal{A}^2$. It is *positive semi-definite* if for any $f\in\mathcal{L}^2(\mathcal{A})$,

$$\int_{\mathcal{A}}\int_{\mathcal{A}}\sigma(a,a')f(a)f(a')\,\mathrm{d}a\,\mathrm{d}a'\geqslant 0.$$

Alternatively, σ is positive semi-definite if and only if for any $n\in\mathbb{N}$, any finite sample $\{a_1,\dots,a_n\}\subset\mathcal{A}$ and any vector $\mathbf{x}=(x_1,\dots,x_n)\in\mathbb{R}^n\setminus\{\mathbf{0}\}$, $\sum_{i=1}^n\sum_{i=1}^n x_ix_j\sigma(a_i,a_j)\geq 0$.

Definition A.2. Let Ω be an outcome space. A process f is *sample-path continuous* at $a_0\in\mathcal{A}$ if, for almost all $\omega\in\Omega$, $a\to a_0$ implies $f(\omega,a)\to f(\omega,a_0)$. The process is sample-path continuous if it is sample-path continuous at any $a_0\in\mathcal{A}$.

It is straightforward that the continuity of μ is necessary for sample path continuity. Without loss, the following two propositions normalize μ to zero. Moreover, Proposition 13 normalizes the attribute space without loss to $[0,1]^d$, where $d\geqslant 1$.

Proposition 13. *Let $\mathcal{A}=[0,1]^d, d\geq 1$ and f be a zero-mean Gaussian process with covariance σ. If there exist $\beta,K>0$ such that $\sigma(a,a)+\sigma(a',a')-2\sigma(a,a')\leqslant K|a-a'|^{d+\beta}$ for all $a,a'\in\mathcal{A}$, then f has a modification on $\mathcal{A}$ that is sample-path continuous.*

Proof of Proposition 13. By Kolmorogov's Continuity Theorem, such a continuous modification exists if there exist $\alpha,\beta,K>0$ such that $\mathbb{E}[|f(a)-f(a')|^\alpha]\leq K|a-a'|^{d+\beta}$ for all $a,a'\in\mathcal{A}$. Letting $\alpha=2$ and using the fact that $\mu(a)=0$ for any $a\in\mathcal{A}$, the LHS becomes $\mathbb{E}[|f(a)-f(a')|^2]=\mathbb{E}[f(a)^2]+\mathbb{E}[f(a')^2]-2\mathbb{E}[f(a)f(a')]=\sigma(a,a)+\sigma(a',a')-2\sigma(a,a')$. If this is less than $K|a-a'|^{d+\beta}$ for some $\beta,K>0$, a continuous modification of f exists. □

Proposition 14. *Let $\mu(a)=0$ for all $a\in\mathcal{A}$. If $f(a)$ is sample-path continuous at $a=a_1,a_2\in\mathcal{A}$, then $\sigma(a,a')$ is continuous at $(a,a')=(a_1,a_2)$.*

Proof of Proposition 14. First, if f is sample-path continuous at some $a_1,a_2\in\mathcal{A}$, then

$$\lim_{a\to a_1}\mathbb{E}\left[(f(a)-f(a_1))^2\right]=\lim_{a\to a_2}\mathbb{E}\left[(f(a)-f(a_2))^2\right]=0.$$

Therefore, f is mean-square continuous at $a=a_1,a_2$.[35] Also, note that $\sigma(a,a')-\sigma(a_1,a_2)=(\sigma(a,a')-\sigma(a_1,a'))+(\sigma(a_1,a')-\sigma(a_1,a_2))$. But $|\sigma(a,a')-\sigma(a_1,a')|=|\mathbb{E}\left[(f(a)-f(a_1))f(a')\right]|\leq\sqrt{\mathbb{E}\left[(f(a)-f(a_1))^2\right]}\sqrt{\mathbb{E}\left[f(a')^2\right]}$, where the inequality follows from the Cauchy-Schwarz inequality for expectations. Because f is mean-square continuous at a_1, the term $\sqrt{\mathbb{E}\left[(f(a)-f(a_1))^2\right]}$ vanishes to zero as $a\to a_1$. Hence, $\lim_{a\to a_1}|\sigma(a,a')-\sigma(a_1,a')|=0$. By a similar logic, $\lim_{a'\to a_2}|\sigma(a_1,a')-\sigma(a_1,a_2)|=0$. Therefore, $\lim_{a\to a_1}\lim_{a'\to a_2}\sigma(a,a')=\sigma(a_1,a_2)$ which means that $\sigma(a,a')$ is continuous at (a_1,a_2). □

Lemma 3. *Fix σ and (ω_A,ω_P).*

(i) There exist an outcome-equivalent covariance $\tilde{\sigma}$ and pair of weights $(\tilde{\omega}_A,\tilde{\omega}_P)$ such that $\tilde{\sigma}(a,a)\in\{0,1\}$ for any $a\in\mathcal{A}$. Moreover, for any $a,a'\in\mathcal{A}$ for which $\sigma(a,a),\sigma(a',a')>0$,

$$\tilde{\sigma}(a,a'):=\frac{\sigma(a,a')}{\sqrt{\sigma(a,a)\sigma(a',a')}},\quad \tilde{\omega}_i(a):=\omega_i(a)\sqrt{\sigma(a,a)}\quad\forall i\in\{A,P\}$$

[35] For Gaussian processes, sample-path continuity implies mean-square continuity, but the converse need not hold. See Theorem 2 in Cambanis (1973).

and for any $a \in \mathcal{A}$ for which $\sigma(a,a) = 0$ and any other $a' \in \mathcal{A}$, $\tilde{\sigma}(a,a') := \sigma(a,a') = 0$ and $\tilde{\omega}(a) := \omega(a)$.

(ii) In the single-player benchmark, without loss $\omega(a) \geqslant 0$ for any $a \in \mathcal{A}$ and

$$\int_{\mathcal{A}} \omega(a)\,\mathrm{d}a = 1.$$

Proof of Lemma 3. (i) Let τ^i and $\tilde{\tau}^i$ denote player i's sample weights corresponding to σ and $\tilde{\sigma}$ respectively. Let $R(\mathbf{a})$ denote the correlation matrix for sample $\mathbf{a}$ and $D(\mathbf{a})$ the diagonal matrix with the j^{th} entry being $\sqrt{\sigma(a_j,a_j)}$. Then $\Sigma(\mathbf{a}) = D(\mathbf{a})R(\mathbf{a})D(\mathbf{a})$ and therefore $\Sigma^{-1}(\mathbf{a}) = D^{-1}(\mathbf{a})R^{-1}(\mathbf{a})D^{-1}(\mathbf{a})$. By the expression in Lemma 1, we have

$$\tilde{\tau}^i_j(a;\mathbf{a}) = \frac{\tau^i_j(a;\mathbf{a})}{\sqrt{\sigma(a,a)}}\sqrt{\sigma(a_j,a_j)}.$$

From here, we obtain that $\tilde{\tau}^i_j(\mathbf{a}) = \tau^i_j(\mathbf{a})\sqrt{\sigma(a_j,a_j)}$, which implies that $\tilde{\alpha}_1(\mathbf{a}) = \alpha_1(\mathbf{a})$ and $\tilde{\alpha}_2(\mathbf{a}) = \alpha_2(\mathbf{a})$ in Theorem 2. Hence, $(\tilde{\sigma}, \tilde{\omega}_A, \tilde{\omega}_P)$ is outcome-equivalent to $(\sigma, \omega_A, \omega_P)$.

(ii) First, for any $a \in \mathcal{A}$ such that $\omega(a) < 0$, redefine its realization as $\tilde{f}(a) = -f(a)$. Moreover, let

$$\Omega := \int_{\mathcal{A}} |\omega(a)|\,\mathrm{d}a, \quad \tilde{\omega}(a) := \frac{|\omega(a)|}{\Omega}.$$

The integrability of ω implies the absolute integrability of ω, hence $\Omega < \infty$. Therefore, $\tilde{\omega}$ is well-defined, everywhere positive, and its integral over $\mathcal{A}$ equals one. □

B Main proofs

Proof of Lemma 1. (i) The joint Gaussian distribution is given by

$$\begin{pmatrix} f(\hat{a}) \\ f(\mathbf{a}) \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu(\hat{a}) \\ \mu(\mathbf{a}) \end{pmatrix}, \begin{pmatrix} \sigma(\hat{a},\hat{a}) & \Sigma(\hat{a},\mathbf{a}) \\ \Sigma(\hat{a},\mathbf{a})^\top & \Sigma(\mathbf{a}) \end{pmatrix} \right)$$

where $\Sigma(\hat{a},\mathbf{a}) = \operatorname{cov}(f(\hat{a}), f(\mathbf{a})) = \begin{pmatrix} \sigma(a_1,\hat{a}) & \ldots & \sigma(a_n,\hat{a}) \end{pmatrix}$ and $\Sigma(\mathbf{a})$ is the sample covariance matrix. Hence, $\mathbb{E}[f(\hat{a}) \mid \mathbf{a}, f(\mathbf{a})] = \mu(\hat{a}) + \Sigma(\hat{a},\mathbf{a})[\Sigma(\mathbf{a})]^{-1}(f(\mathbf{a}) - \mu(\mathbf{a}))$. Moreover, sample realizations are observed perfectly. So, for any $(\mathbf{a}, f(\mathbf{a}))$ and any $a_j \in \mathbf{a}$, $\mathbb{E}[f(a_j) \mid \mathbf{a}, f(\mathbf{a})] = f(a_j)$. Therefore, $\tau_j(a_j;\mathbf{a}) = 1$ and $\tau_m(a_j;\mathbf{a}) = 0$ for $m \neq j$.

(ii) First suppose that $\mu(a) = 0$ for all $a \in \mathcal{A}$. Applying part (i),

$$\begin{aligned} \nu^i(\mathbf{a}) &= \int_{\mathcal{A}} \mathbb{E}[f(a) \mid \mathbf{a}, f(\mathbf{a})]\omega_i(a)\,\mathrm{d}a \\ &= \int_{\mathcal{A}} (\tau_1(a;\mathbf{a})f(a_1) + \ldots \tau_n(a;\mathbf{a})f(a_n))\,\omega_i(a)\,\mathrm{d}a = \sum_{j=1}^n f(a_j)\left(\int_{\mathcal{A}} \tau_j(a;\mathbf{a})\omega_i(a)\,\mathrm{d}a\right). \end{aligned}$$

For any $j \in \{1,\ldots,k\}$, $(\nu^i(\mathbf{a}) - \nu^i_0)$ and $(f(a_j) - \mu(a_j))$ are centered at zero. Therefore, for an arbitrary mean $\mu : [0,1] \to \mathbb{R}$, we have

$$\nu^i(\mathbf{a}) - \nu^i_0 = \sum_{j=1}^n \left(\int_{\mathcal{A}} \tau_j(a;\mathbf{a})\omega_i(a)\,\mathrm{d}a\right)(f(a_j) - \mu(a_j)).$$

□

Proof of Lemma 2. (iii) ⇒ (ii): Suppose f is Markov. Then, for any finite selection of realizations $f(a_1), \ldots, f(a_n)$ for $a_1 < a_2 < \ldots < a_n$ and any $i \leqslant m \leqslant j$,

$$\sigma(a_i, a_j) = \mathbb{E}_{\mathcal{F}_m}\left[\mathbb{E}[f(a_i)f(a_j) \mid \mathcal{F}_m]\right] = \mathbb{E}_{\mathcal{F}_m}\left[f(a_i)\mathbb{E}[f(a_j) \mid \mathcal{F}_m]\right]$$
$$= \mathbb{E}_{\mathcal{F}_m}\left[f(a_i)\mathbb{E}[f(a_j) \mid f(a_m)]\right] = \mathbb{E}_{\mathcal{F}_m}\left[f(a_i)\frac{\sigma(a_j, a_m)}{\sigma(a_m, a_m)}f(a_m)\right] = \frac{\sigma(a_j, a_m)}{\sigma(a_m, a_m)}\sigma(a_i, a_m),$$

where $\mathcal{F}_m$ is the filtration generated by $\{f(a) : 0 \leqslant a \leqslant a_m\}$. The first equality uses the definition of the covariance and the Law of Iterated Expectations, the second equality follows from $i \leqslant m \leqslant j$, the third equality follows from the Markov property of f, the fourth equality follows from the jointly Gaussian distribution of $f(a_j)$ and $f(a_m)$, and the last equality uses $\mathbb{E}[f(a_i)f(a_m)] = \sigma(a_i, a_m)$. From Lemma B.3, $\sigma(a_m, a_m) = 1$.

(ii) ⇒ (iii): Suppose σ satisfies the triangle property in (ii), and let $i \leqslant m \leqslant j$. For f to be Markov, it is sufficient to show that $\mathbb{E}[f(a_j) \mid \mathcal{F}_m] = \mathbb{E}[f(a_j) \mid f(a_m)]$. Consider

$$\mathbb{E}\left[\left(\mathbb{E}[f(a_j) \mid \mathcal{F}_m] - \frac{\sigma(a_m, a_j)}{\sigma(a_m, a_m)}f(a_m)\right)f(a_i)\right] = \mathbb{E}\left[\mathbb{E}[f(a_i)f(a_j) \mid \mathcal{F}_m] - \frac{\sigma(a_m, a_j)}{\sigma(a_m, a_m)}f(a_m)f(a_i)\right]$$
$$= \sigma(a_i, a_j) - \mathbb{E}[f(a_i)f(a_m)]\frac{\sigma(a_m, a_j)}{\sigma(a_m, a_m)} = 0,$$

where the first equality is because $f(a_i)$ for $i \leqslant m$ is a constant with respect to $\mathcal{F}_m$, the second equality uses the Law of Iterated Expectations, and the last equality uses (ii). Therefore, $\mathbb{E}[f(a_j) \mid \mathcal{F}_m] - \frac{\sigma(a_m,a_j)}{\sigma(a_m,a_m)}f(a_m)$ is independent of $f(a_1), \ldots, f(a_m)$. But attribute realizations are jointly Gaussian, so $\mathbb{E}[f(a_j) \mid \mathcal{F}_m] - \frac{\sigma(a_m,a_j)}{\sigma(a_m,a_m)}f(a_m)$ is a linear combination of $f(a_1), \ldots, f(a_m)$. Therefore,

$$\mathbb{E}[f(a_j) \mid \mathcal{F}_m] - \frac{\sigma(a_m, a_j)}{\sigma(a_m, a_m)}f(a_m) = 0 \Rightarrow \mathbb{E}[f(a_j) \mid \mathcal{F}_m] = \frac{\sigma(a_m, a_j)}{\sigma(a_m, a_m)}f(a_m) = \mathbb{E}[f(a_j) \mid f(a_m)].$$

(i) ⇒ (ii): Suppose that σ satisfies NAP. Then, for any $a_1 < a_2 < \ldots < a_n$ and any $a > a_n$, $\mathbb{E}[f(a) \mid \mathcal{F}_n] = \mathbb{E}[f(a) \mid f(a_n)]$ where $\mathcal{F}_n$ is the filtration generated by $\{f(\hat{a}) : 0 \leqslant \hat{a} \leqslant a_n\}$. Hence, f is Markov. By the argument above, σ satisfies the triangle property.

(ii) ⇒ (i): We use the following result from Barrett and Feinsilver (1978).

Claim 1 (Theorem 1, Barrett and Feinsilver (1978))**.** *A positive definite symmetric matrix has the triangle property if and only if its inverse is tridiagonal.*

Fix $\mathbf{a} \in \mathcal{A}_k$ and let $\Sigma := \Sigma(\mathbf{a})$ be its positive definite and symmetric covariance matrix. The theorem implies that Σ^{-1} is tridiagonal for any sample $\mathbf{a}$, i.e., $\Sigma^{-1}_{ij} = 0$ for any $|i-j| > 1$. Fix $\hat{a} \in \mathcal{A}$. Equation (6) simplifies to

$$\tau_j(\hat{a}; \mathbf{a}) = \begin{cases} \sigma(\hat{a}, a_1)\Sigma^{-1}_{11} + \sigma(\hat{a}, a_2)\Sigma^{-1}_{12} & \text{if } j = 1 \\ \sigma(\hat{a}, a_{j-1})\Sigma^{-1}_{j-1,j} + \sigma(\hat{a}, a_j)\Sigma^{-1}_{jj} + \sigma(\hat{a}, a_{j+1})\Sigma^{-1}_{j+1,j} & \text{if } j = 2, \ldots, n-1 \\ \sigma(\hat{a}, a_{n-1})\Sigma^{-1}_{n-1,n} + \sigma(\hat{a}, a_n)\Sigma^{-1}_{nn} & \text{if } j = n. \end{cases}$$

From Barrett and Feinsilver (1978), the entries of the inverse of the covariance matrix are

$$\Sigma^{-1}_{11} = \frac{\sigma(a_2, a_2)}{d_{12}}, \Sigma^{-1}_{jj} = \sigma(a_j, a_j)\frac{d_{j-1,j+1}}{d_{j,j+1}d_{j-1,j}}, \Sigma^{-1}_{nn} = \frac{\sigma(a_{n-1}, a_{n-1})}{d_{n-1,n}}, \Sigma^{-1}_{ij} = -\frac{\sigma(a_i, a_j)}{d_{ij}}$$

if $|i-j| = 1$ for the latter, where $d_{ij} = \sigma(a_i, a_i)\sigma(a_j, a_j) - \sigma^2(a_i, a_j)$. All other entries of Σ^{-1} are zero. Substituting these expressions into $\tau_j(\hat{a}; \mathbf{a})$ we obtain that (1) if $\hat{a} < a_1$, then $\tau_1(\hat{a}; \mathbf{a}) =$

$\sigma(\hat{a}, a_1)/\sigma(a_1, a_1)$ and $\tau_j(\hat{a}; \mathbf{a}) = 0$ for $j \neq 1$; (2) if $a_i < \hat{a} < a_{i+1}$, where $i = 1, \dots, n-1$, then

$$\tau_j(\hat{a}; \mathbf{a}) = \begin{cases} \dfrac{\sigma(\hat{a}, a_i)\sigma(a_{i+1}, a_{i+1}) - \sigma(\hat{a}, a_{i+1})\sigma(a_i, a_{i+1})}{d_{i,i+1}} & \text{if } j = i \\ \dfrac{\sigma(\hat{a}, a_{i+1})\sigma(a_i, a_i) - \sigma(\hat{a}, a_i)\sigma(a_i, a_{i+1})}{d_{i,i+1}} & \text{if } j = i+1 \\ 0 & \text{otherwise;} \end{cases} \tag{18}$$

(3) if $\hat{a} > a_n$, then $\tau_n(\hat{a}; \mathbf{a}) = \sigma(\hat{a}, a_n)/\sigma(a_n, a_n)$ and $\tau_j(\hat{a}; \mathbf{a}) = 0$ for $j \neq n$. □

Proof of Corollary 1. The sample weights follow immediately from summing up the weights $\tau_j(\hat{a}; \mathbf{a})$ derived in the proof of Lemma 2 over $\hat{a} \in \mathcal{A}$. A useful fact for such simplification is that NAP implies $\sigma(a, a_i) - \sigma(a, a_{i-1})\sigma(a_i, a_{i-1}) = \sigma(a, a_i)\left(1 - \sigma^2(a, a_{i-1})\right)$ for any $a \in (a_{i-1}, a_i)$.

To show that $\sigma(a, a') \geqslant 0$ for any $a, a' \in \mathcal{A}$, we suppose, by contradiction, that there exist distinct attributes $a_1 < a_2$ such that $\sigma(a_1, a_2) < 0$. Because $\sigma(a_1, a_1) \geqslant 0$ by the definition of attribute variance and $\sigma(a_1, a)$ is continuous in a by Assumption 1, there exists $a_0 \in (a_1, a_2)$ such that $\sigma(a_1, a_0) = 0$. Given that σ satisfies NAP, $\sigma(a_1, a_2) = \sigma(a_1, a_0)\sigma(a_0, a_2) = 0$, which leads to a contradiction. Hence, $\sigma(a, a') \geqslant 0$ for any $a, a' \in \mathcal{A}$.

By Lemma B.3, for any j, $1 - \sigma^2(a_{j-1}, a_j) \geqslant 0$ because $\sigma(a_{j-1}, a_j) \leqslant 1$. Moreover, for $a \in (a_{j-1}, a_j)$, $\sigma(a, a_j)\left(1 - \sigma^2(a_{j-1}, a)\right) \geqslant 0$ and $\sigma(a_{i-1}, a)\left(1 - \sigma^2(a, a_i)\right) \geqslant 0$ by the argument above. Therefore, if ω_i does not switch sign over $\mathcal{A}$, the sample weights τ_j^i have the same weight as ω_i. □

Proof for Theorem 1. Fix $\mathbf{a} \in \mathcal{A}_k$. Given $(\mathbf{a}, f(\mathbf{a}))$, the player's expected value for the project is $\nu(\mathbf{a})$, from Lemma 1. The player's expected payoff from sampling $\mathbf{a}$ is the negative of the mean squared error, hence it takes the variance-bias form:

$$\begin{aligned} \mathbb{E}\left[-(\nu(\mathbf{a}) - v)^2\right] &= -\mathbb{E}_v\left[\text{var}[\nu(\mathbf{a}) \mid v] + (\mathbb{E}[\nu(\mathbf{a}) \mid v] - v)^2\right] \\ &= -\mathbb{E}_v\left[\text{var}[\nu(\mathbf{a}) \mid v]\right] = \text{var}[\nu(\mathbf{a})] - \text{var}[\mathbb{E}[\nu(\mathbf{a}) \mid v]] = \psi^2(\mathbf{a}) - \text{var}[v], \end{aligned}$$

where the second equality follows from $\mathbb{E}[\nu(\mathbf{a}) \mid v] = v$, the third equality follows from the law of total variance, and $\psi^2(\mathbf{a}) = \text{var}[\nu(\mathbf{a})]$. This strictly increases in $\psi^2(\mathbf{a})$, hence any single-player sample maximizes $\psi^2(\mathbf{a})$.

Decomposed another way, the player's expected payoff also equals $\mathbb{E}[-\nu(\mathbf{a})^2] + 2\mathbb{E}[\nu(\mathbf{a})v] - \mathbb{E}[v^2] = -\psi^2(\mathbf{a}) - \text{var}[v] + 2\text{cov}[\nu(\mathbf{a}), v]$. Hence, it must be that $\text{cov}[\nu(\mathbf{a}), v] = \psi^2(\mathbf{a})$. Therefore,

$$\begin{aligned} \psi^2(\mathbf{a}) &= \text{cov}\left[\sum_{j=1}^n \tau_j(\mathbf{a}) f(a_j), \int_{\mathcal{A}} \omega(a) f(a)\, \text{d}a\right] = \sum_{j=1}^n \tau_j(\mathbf{a}) \text{cov}\left[f(a_j), \int_{\mathcal{A}} \omega(a) f(a)\, \text{d}a\right] \\ &= \sum_{j=1}^n \tau_j(\mathbf{a}) \left(\int_{\mathcal{A}} \omega(a) \text{cov}[f(a_j), f(a)]\, \text{d}a\right) = \sum_{j=1}^n \tau_j(\mathbf{a}) \left(\int_{\mathcal{A}} \omega(a) \sigma(a_j, a)\, \text{d}a\right) \end{aligned}$$

which gives equation (10). On the other hand, posterior variance is alternatively equal to $\psi^2(\mathbf{a}) = \text{cov}\left[\sum_{j=1}^n \tau_j(\mathbf{a}) f(a_j), \sum_{i=1}^n \tau_i(\mathbf{a}) f(a_i)\right] = \sum_{i=1}^n \sum_{j=1}^n \tau_i(\mathbf{a}) \tau_j(\mathbf{a}) \sigma(a_i, a_j)$, therefore equalizing the two expressions for $\psi^2(\mathbf{a})$, we obtain $\tau_j(\mathbf{a}) + \sum_{i \neq j} \tau_i(\mathbf{a}) \sigma(a_i, a_j) = \tau(a_j)$. The expression for $\psi^2(\mathbf{a})$ is independent of μ, and hence ν_0. Hence, the set of single-player samples does not depend on (μ, ν_0). □

Proof of Proposition 1. Suppose $\mathbf{a}_{-1}^s = \{a_2^s, \dots, a_k^s\}$ has been discovered with realizations $f(\mathbf{a}_{-1}^s)$ and posterior value $\nu(\mathbf{a}_{-1}^s)$. By Lemma 2, if σ satisfies NAP then f is Markov. Therefore, the distribution of $f(a_1)$ given $f(\mathbf{a}^s_{-1})$ depends only on $f(a_2)$. The difference in the posterior values generated from the samples with and without a_1 is $\nu(\mathbf{a}) - \nu(\mathbf{a}^s_{-1}) = \tau_1(\mathbf{a})f(a_1) + \left(\tau_2(\mathbf{a}) - \tau_1(\mathbf{a}^s_{-1})\right) f(a_2)$. The difference $\tau_2(\mathbf{a}) - \tau_1(\mathbf{a}^s_{-1})$ in the sample weight of a^s_2 across the two samples equals

$$-\int_0^{a_1} \sigma(a, a^s_2)\omega(a)\,\mathrm{d}a - \int_{a_1}^{a^s_2} \sigma(a, a^s_2)\frac{\sigma^2(a_1, a) - \sigma^2(a_1, a^s_2)}{1 - \sigma^2(a_1, a^s_2)}\omega(a)\,\mathrm{d}a = -\sigma(a_1, a^s_2)\tau_1(\mathbf{a}),$$

where the second equality uses the fact that for $a < a_1$, $\sigma(a, a^s_2) = \sigma(a_1, a^s_2)\sigma(a, a_1)$ and for $a \in (a_1, a^s_2)$, $\sigma(a, a^s_2)\sigma(a, a_1) = \sigma(a_1, a^s_2)$. Therefore,

$$\nu(\mathbf{a}) - \nu(\mathbf{a}^s_{-1}) = \tau_1(\mathbf{a})\left(f(a_1) - \sigma(a_1, a_2)f(a_2)\right).$$

The difference $f(a_1) - \sigma(a_1, a_2)f(a_2)$ is uncorrelated with $f(a_2)$. Therefore, $f(a_1) - \sigma(a_1, a_2)f(a_2)$ is also uncorrelated with $\nu(\mathbf{a}^s_{-1})$, which implies that $\mathrm{cov}(\nu(\mathbf{a}), \nu(\mathbf{a}^s_{-1}) = \psi^2(\mathbf{a}^s_{-1})$. Hence,

$$\psi^2(\mathbf{a}) - \psi^2(\mathbf{a}^s_{-1}) = \mathrm{var}(\nu(\mathbf{a}) - \nu(\mathbf{a}^s_{-1}) = \tau_1^2(\mathbf{a})\left(1 - \sigma^2(a_1, a_2)\right),$$

which is what a^s_1 maximizes. By the same argument,

$$\psi^2\left(\mathbf{a}^s_{-k} \cup \{a_k\}\right) - \psi^2(\mathbf{a}^s_{-k}) = \tau_k^2\left(\mathbf{a}^s_{-k} \cup \{a_k\}\right)\left(1 - \sigma^2(a^s_{k-1}, a_k)\right).$$

Consider now $\mathbf{a}' := \mathbf{a}^s_{-j} \cup \{a_j\}$ for an interior $j = 2, \ldots, k-1$. Using NAP, the difference in the sample weights assigned to a^s_{j-1} and a^s_{j+1} due to the presence of a_j in the sample is

$$\begin{aligned}
\tau_{j-1}(\mathbf{a}') - \tau_{j-1}(\mathbf{a}^s_{-j}) &= \int_{a^s_{j-1}}^{a_j} \sigma(a, a^s_{j-1})\frac{1 - \sigma^2(a, a_j)}{1 - \sigma^2(a_j, a^s_{j-1})}\omega(a)\,\mathrm{d}a \\
&\qquad - \int_{a^s_{j-1}}^{a^s_{j+1}} \sigma(a, a^s_{j-1})\frac{1 - \sigma^2(a, a^s_{j+1})}{1 - \sigma^2(a^s_{j-1}, a^s_{j+1})}\omega(a)\,\mathrm{d}a \\
&= -\sigma(a_j, a^s_{j-1})\frac{1 - \sigma^2(a_j, a^s_{j+1})}{1 - \sigma^2(a^s_{j-1}, a^s_{j+1})}\left(\int_{a^s_{j-1}}^{a_j} \sigma(a, a_j)\frac{1 - \sigma^2(a, a^s_{j-1})}{1 - \sigma^2(a_j, a^s_{j-1})}\omega(a)\,\mathrm{d}a\right) \\
&:= -\sigma(a_j, a^s_{j-1})\frac{1 - \sigma^2(a_j, a^s_{j+1})}{1 - \sigma^2(a^s_{j-1}, a^s_{j+1})}\tau^l_j(\mathbf{a})
\end{aligned}$$

and by similar reasoning

$$\begin{aligned}
\tau_{j+1}(\mathbf{a}') - \tau_j(\mathbf{a}^s_{-j}) &= -\sigma(a_j, a^s_{j+1})\frac{1 - \sigma^2(a^s_{j-1}, a_j)}{1 - \sigma^2(a^s_{j-1}, a^s_{j+1})}\left(\int_{a_j}^{a^s_{j+1}} \sigma(a, a_j)\frac{1 - \sigma^2(a_j, a^s_{j+1})}{1 - \sigma^2(a_j, a^s_{j+1})}\omega(a)\,\mathrm{d}a\right) \\
&:= -\sigma(a_j, a^s_{j+1})\frac{1 - \sigma^2(a^s_{j-1}, a_j)}{1 - \sigma^2(a^s_{j-1}, a^s_{j+1})}\tau^r_j(\mathbf{a}).
\end{aligned}$$

The difference in posterior values is, therefore,

$$\begin{aligned}
\nu(\mathbf{a}) - \nu(\mathbf{a}^s_{-j}) &= \tau_j(\mathbf{a})f(a_j) + \left(\tau_{j-1}(\mathbf{a}') - \tau_{j-1}(\mathbf{a}^s_{-j})\right) f(a_{j-1}) + \left(\tau_{j+1}(\mathbf{a}') - \tau_j(\mathbf{a}^s_{-j})\right) f(a_{j+1}) \\
&= \tau^l_j(\mathbf{a})\left(f(a_j) - \sigma(a_j, a^s_{j-1})\frac{1 - \sigma^2(a_j, a^s_{j+1})}{1 - \sigma^2(a^s_{j-1}, a^s_{j+1})}f(a^s_{j-1})\right) \\
&\qquad + \tau^r_j(\mathbf{a})\left(f(a_j) - \sigma(a_j, a^s_{j+1})\frac{1 - \sigma^2(a^s_{j-1}, a_j)}{1 - \sigma^2(a^s_{j-1}, a^s_{j+1})}f(a^s_{j+1})\right).
\end{aligned}$$

The difference $f(a_j) - \sigma(a_j, a^s_{j-1})\frac{1-\sigma^2(a_j,a^s_{j+1})}{1-\sigma^2(a^s_{j-1},a^s_{j+1})}f(a^s_{j-1})$ is uncorrelated with $f(a^s_{j-1})$ but correlated with $f(a^s_{j+1})$. The difference $f(a_j) - \sigma(a_j, a^s_{j+1})\frac{1-\sigma^2(a^s_{j-1},a_j)}{1-\sigma^2(a^s_{j-1},a^s_{j+1})}f(a^s_{j+1})$ is uncorrelated with $f(a^s_{j+1})$ but correlated with $f(a^s_{j-1})$. By a similar argument to the one above, the difference in posterior variances simplifies to

$$\psi^2(\mathbf{a}) - \psi^2(\mathbf{a}^s_{-j}) = (\tau^l_j(\mathbf{a}) + \tau^r_j(\mathbf{a}))^2 \frac{(1-\sigma^2(a^s_{j-1}, a_j))(1-\sigma^2(a_j, a^s_{j+1}))}{1-\sigma^2(a^s_{j-1}, a^s_{j+1})}.$$

□

Proof of Proposition 2. Let $\underline{a}$ and $\bar{a}$ be, respectively, the minimum and the maximum attribute in $\mathrm{supp}(\omega)$. Because $\mathrm{supp}(\omega)$ is compact, it is either a single interval $[\underline{a}, \bar{a}]$ or a union of (potentially degenerate) intervals $[\underline{a}, \bar{a}] \setminus \bigcup_n I_n$, where $\{I_n\}$ is a family of countably many disjoint open intervals of the form $(\underline{a}_n, \bar{a}_n)$, ordered from left to right. Then, $\underline{a} < \underline{a}_1$ and $\sup(\bigcup_n I_n) < \bar{a}$.

Case I: $a_1 < \underline{a}$ or $a_k > \bar{a}$. Let $\mathbf{a}^s = \{a_1, a_2, \ldots, a_k\}$. We first observe that there is at most one sample attribute to the left of $\underline{a}$: if $a_j < \underline{a}$ for $j > 1$, then $\tau_k(\mathbf{a}^s) = 0$ for any $k \le j-1$. Similarly, $a_{k-1} \le \bar{a}$. By way of contradiction, suppose $a_1 < \underline{a}$. By Proposition 1, a_1 maximizes

$$\begin{aligned}\tau_1^2(\mathbf{a}^s)\left(1-\sigma^2(a_1,a_2)\right) &= \left(\int_{\underline{a}}^{a_2} \sigma(a,a_1)\left(1-\sigma^2(a,a_2)\right)\omega(a)\,\mathrm{d}a\right)^2 / (1-\sigma^2(a_1,a_2)) \\ &= \frac{\sigma^2(a_1,\underline{a})}{1-\sigma^2(a_1,\underline{a})\sigma^2(\underline{a},a_2)}\left(\int_{\underline{a}}^{a_2} \sigma(a,\underline{a})\left(1-\sigma^2(a,a_2)\right)\omega(a)\,\mathrm{d}a\right)^2 \\ &=: \frac{\sigma^2(a_1,\underline{a})}{1-\sigma^2(a_1,\underline{a})\sigma^2(\underline{a},a_2)} C_0^2,\end{aligned}$$

where C_0 does not depend on a_1. But the RHS is strictly increasing in $\sigma^2(a_1, \underline{a})$, which contradicts the optimality of a_1. The player is better off sampling $a_1 = \underline{a}$ instead. By a similar argument, $a_k \le \bar{a}$ as well.

Case II: $a_1 \in I_n$ or $a_k \in I_n$ for some n. Without loss, consider $a_1 \in I_n = (\underline{a}_n, \bar{a}_n)$. Suppose first that $a_1 < \bar{a}_n < a_2$, so a_1 is the only sample attribute in the interval I_n. Then,

$$\begin{aligned}\tau_1^2(\mathbf{a}^s)\left(1-\sigma^2(a_1,a_2)\right) &= \left(\int_{\underline{a}}^{\underline{a}_n} \sigma(a,a_1)\omega(a)\,\mathrm{d}a + \int_{\bar{a}_n}^{a_2} \sigma(a,a_1)\frac{1-\sigma^2(a,a_2)}{1-\sigma^2(a_1,a_2)}\omega(a)\,\mathrm{d}a\right)^2 \left(1-\sigma^2(a_1,a_2)\right) \\ &:= \left(\sigma(\underline{a}_n,a_1)C_1 + \frac{\sigma(a_1,\bar{a}_n)}{1-\sigma^2(a_1,\bar{a}_n)\sigma^2(\bar{a}_n,a_2)}C_2\right)^2 \left(1-\sigma^2(a_1,a_2)\right) \\ &= \sigma^2(\underline{a}_n,a_1)C_1^2 + \frac{\sigma^2(a_1,\bar{a}_n)}{1-\sigma^2(a_1,\bar{a}_n)\sigma^2(\bar{a}_n,a_2)}C_2^2 - \sigma^2(\underline{a}_n,a_2)C_1^2 + 2C_1C_2\sigma(\underline{a}_n,\bar{a}_n) \\ &:= \frac{\sigma^2(\underline{a}_n,\bar{a}_n)}{\sigma^2(a_1,\bar{a}_n)}C_1^2 + \frac{\sigma^2(a_1,\bar{a}_n)}{1-\sigma^2(a_1,\bar{a}_n)\sigma^2(\bar{a}_n,a_2)}C_2^2 + C_3,\end{aligned}$$

where C_1, C_2, C_3 do not depend on a_1. This can be rewritten in terms of $x := \sigma^2(a_1, \bar{a}_n)$ as

$$f(x) = \frac{\sigma^2(\underline{a}_n,\bar{a}_n)}{x}C_1^2 + \frac{x}{1-x\sigma^2(\bar{a}_n,a_2)}C_2^2 + C_3,$$

the second derivative of which with respect to x is

$$\frac{\partial^2 f(x)}{\partial x^2} = 2\left(\frac{C_1^2\sigma^2(\underline{a}_n,\bar{a}_n)}{x^3} + \frac{C_2^2\sigma^2(\bar{a}_n,a_2)}{(1-x\sigma^2(\bar{a}_n,a_2))^3}\right) > 0.$$

That is, given $\mathbf{a}^s \setminus \{a_1\}$, the marginal value of sampling a_1 is convex in $\sigma^2(a_1,\bar{a}_n) \in [\sigma^2(\underline{a}_n,\bar{a}_n),1]$. The player is better off sampling either $a_1 = \underline{a}_n$ or $a_1 = \bar{a}_n$ instead. A similar argument applies to $a_k \in I_n$ for some n.

Next, suppose that $a_1, a_2 \in I_n$ for some n. It is immediate that no more than two sample attributes can be in any I_n. Then,

$$\begin{aligned}\tau_1^2(\mathbf{a}^s)\left(1-\sigma^2(a_1,a_2)\right) &= \left(\int_{\underline{a}}^{\underline{a}_n} \sigma(a,a_1)\omega(a)\,\mathrm{d}a\right)^2\left(1-\sigma^2(a_1,a_2)\right)\\ &:= \sigma^2(\underline{a}_n,a_1)C_4^2 + \sigma^2(\underline{a}_n,a_2)C_4^2\end{aligned}$$

where C_4 does not depend on a_1. It is immediate that this is increasing in $\sigma^2(\underline{a}_n,a_1)$ so the player is better off sampling $a_1 = \underline{a}_n$.

Following a similar argument, Proposition 1 implies that a_2 maximizes

$$\begin{aligned}\tau_2^2(\mathbf{a}^s)\frac{\left(1-\sigma^2(a_1,a_2)\right)\left(1-\sigma^2(a_2,a_3)\right)}{(1-\sigma^2(a_1,a_3))} &= \left(\int_{\bar{a}_n}^{a_3}\sigma(a,a_2)\frac{1-\sigma^2(a,a_3)}{1-\sigma^2(a_2,a_3)}\omega(a)\,\mathrm{d}a\right)^2\frac{\left(1-\sigma^2(a_1,a_2)\right)\left(1-\sigma^2(a_2,a_3)\right)}{(1-\sigma^2(a_1,a_3))}\\ &= \left(\frac{\sigma(a_2,\bar{a}_n)}{1-\sigma^2(a_2,a_3)}C_5\right)^2\frac{\left(1-\sigma^2(a_1,a_2)\right)\left(1-\sigma^2(a_2,a_3)\right)}{(1-\sigma^2(a_1,a_3))}\\ &= \frac{\sigma^2(a_2,\bar{a}_n)\left(1-\sigma^2(a_1,a_2)\right)}{1-\sigma^2(a_2,a_3)}\frac{C_5^2}{1-\sigma^2(a_1,a_3)}\\ &= \frac{\sigma^2(a_2,\bar{a}_n)-\sigma^2(a_1,\bar{a}_n)}{1-\sigma^2(a_2,\bar{a}_n)\sigma^2(\bar{a}_n,a_3)}\frac{C_5^2}{1-\sigma^2(a_1,a_3)}\end{aligned}$$

But note that the first term is strictly increasing in $\sigma^2(a_2,\bar{a}_n)$, which contradicts the optimality of any $a_2 < \bar{a}_n$. The player is better off sampling $a_2 = \bar{a}_n$. We have therefore established that it cannot be that $a_1 \in I_n$ or $a_1, a_2 \in I_n$ for some n.

Case III: $a_j \in I_n$ for $j = 2, \ldots, k-1$. To establish that this cannot be the case, we show that either $a_2 \in I_n$ by itself or $a_2, a_3 \in I_n$. The impossibility of $a_1, a_2 \in I_n$ was established in case II.

First suppose that for some n, $a_2 \in I_n$ and $a_1, a_3 \notin I_n$. By a similar argument to the one above, and letting

$$C_6 := \int_{a_1}^{\underline{a}_n}\sigma(a,\underline{a}_n)(1-\sigma^2(a,a_1))\omega(a)\,\mathrm{d}a, \quad C_7 := \int_{\bar{a}_n}^{a_3}\sigma(a,\bar{a}_n)(1-\sigma^2(a,a_3))\omega(a)\,\mathrm{d}a,$$

the objective in Proposition 1 simplifies to

$$\begin{aligned}&\left(\frac{\sigma(\underline{a}_n,a_2)}{1-\sigma^2(a_1,a_2)}C_6 + \frac{\sigma(a_2,\bar{a}_n)}{1-\sigma^2(a_2,a_3)}C_7\right)^2\frac{\left(1-\sigma^2(a_1,a_2)\right)\left(1-\sigma^2(a_2,a_3)\right)}{1-\sigma^2(a_1,a_3)}\\ &= \frac{\sigma^2(\underline{a}_n,a_2)}{1-\sigma^2(a_1,a_2)}(1-\sigma^2(a_2,a_3))\frac{C_6^2}{1-\sigma^2(a_1,a_3)} + \frac{\sigma^2(a_2,\bar{a}_n)}{1-\sigma^2(a_2,a_3)}(1-\sigma^2(a_1,a_2))\frac{C_7^2}{1-\sigma^2(a_1,a_3)} +\\ &+ 2\frac{\sigma(\underline{a}_n,\bar{a}_n)}{1-\sigma^2(a_1,a_3)}C_6C_7.\end{aligned}$$

Letting $x := \sigma^2(\underline{a}_n, a_2)$ and supressing constants that do not depend on x, this can be expressed as

$$f(x) = \kappa_0 + \frac{x}{1 - x\sigma^2(a_1, \underline{a}_n)} \left(1 - \frac{\sigma^2(\underline{a}_n, a_3)}{x}\right) \kappa_1 + \frac{\sigma^2(\underline{a}_n, \bar{a}_n)/x}{1 - \sigma^2(\underline{a}_n, a_3)/x} \left(1 - x\sigma^2(a_1, \underline{a}_n)\right) \kappa_2.$$

This function is convex in x because

$$\frac{\partial^2 f(x)}{\partial x^2} = 2(1 - \sigma^2(a_1, a_3)) \left(\frac{\sigma^2(\underline{a}_n, \bar{a}_n)\kappa_2}{\sigma^2(\underline{a}_n, a_2) - \sigma^2(\underline{a}_n, a_3)} + \frac{\sigma^2(a_1, \underline{a}_n)\kappa_1}{1 - x\sigma^2(a_1, \underline{a}_n)} \right) > 0.$$

Therefore, it is maximized at either $a_2 = \underline{a}_n$ or $a_2 = \bar{a}_n$, which contradicts the optimality of $a_2 \in I_n$.

Finally, suppose $a_2, a_3 \in I_n$. By an argument identical to that in Case II, it follows that the player is better off setting $a_3 = \bar{a}_n$. On the other hand, a_2 maximizes

$$\begin{aligned}
&\left(\int_{a_1}^{\underline{a}_n} \sigma(a, a_2) \frac{1 - \sigma^2(a, a_1)}{1 - \sigma^2(a_1, a_2)} \omega(a)\, \mathrm{d}a\right)^2 \frac{\left(1 - \sigma^2(a_1, a_2)\right)\left(1 - \sigma^2(a_2, a_3)\right)}{\left(1 - \sigma^2(a_1, a_3)\right)} \\
&= \frac{\sigma^2(\underline{a}_n, a_2)}{1 - \sigma^2(a_1, a_2)} (1 - \sigma^2(a_2, a_3)) \frac{C_8^2}{1 - \sigma^2(a_1, a_3)} \\
&= \frac{\sigma^2(\underline{a}_n, a_2)}{1 - \sigma^2(\underline{a}_n, a_2)\sigma^2(a_1, \underline{a}_n)} \left(1 - \frac{\sigma^2(\underline{a}_n, a_3)}{\sigma^2(\underline{a}_n, a_2)}\right) \frac{C_8^2}{1 - \sigma^2(a_1, a_3)},
\end{aligned}$$

which is strictly increasing in $\sigma^2(\underline{a}_n, a_2)$. Therefore, the player is better off setting $a_2 = \underline{a}_n$. □

Proof of Theorem 2. Fix $\mathbf{a} = \{a_1, \ldots, a_n\} \in \mathcal{A}_k$. Given $f(\mathbf{a})$, the principal's best reply is $d^*(f(\mathbf{a})) = \mathbb{E}[v_P \mid f(\mathbf{a})] = \nu_P(\mathbf{a})$. Therefore, the agent's payoff $V_A(\mathbf{a})$ is

$$\begin{aligned}
\mathbb{E}\left[-(\nu_P(\mathbf{a}) - v_A)^2\right] &= \mathbb{E}\left[-\nu_P(\mathbf{a})^2\right] + 2\mathbb{E}\left[\nu_P(\mathbf{a})v_A\right] - \mathbb{E}\left[v_A^2\right] \\
&= -\left(\psi_P^2(\mathbf{a}) + (\nu_0^P)^2\right) + 2\left(\mathrm{cov}[\nu_P(\mathbf{a}), v_A] + \nu_0^P \nu_0^A\right) - \left(\mathrm{var}[v_A] + (\nu_0^A)^2\right) \\
&= -(\nu_0^P - \nu_0^A)^2 - \psi_P^2(\mathbf{a}) - \mathrm{var}[v_A] + 2\mathrm{cov}[\nu_P(\mathbf{a}), v_A].
\end{aligned}$$

If the agent does not sample any attributes, his payoff is

$$\begin{aligned}
V_A(\emptyset) \coloneqq \mathbb{E}\left[-(\nu_0^P - v_A)^2\right] &= -(\nu_0^P)^2 + 2\nu_0^P \mathbb{E}[v_A] - \mathbb{E}[v_A^2] \\
&= -(\nu_0^P)^2 + 2\nu_0^P \nu_0^A - \left(\mathrm{var}[v_A] + (\nu_0^A)^2\right) \\
&= -(\nu_0^P - \nu_0^A)^2 - \mathrm{var}[v_A],
\end{aligned}$$

therefore $V_A(\mathbf{a}) = V_A(\emptyset) + 2\mathrm{cov}[\nu_P(\mathbf{a}), v_A] - \psi_P^2(\mathbf{a})$. The agent maximizes the added value $2\mathrm{cov}[\nu_P(\mathbf{a}), v_A] - \psi_P^2(\mathbf{a})$, which is constant in $\left(\nu_0^A, \nu_0^P\right)$. But $\alpha_2(\mathbf{a}) = \mathrm{cov}[\nu_P(\mathbf{a}), v_A]$ equals

$$\mathrm{cov}\left[\nu_0^P + \sum_{j=1}^n \tau_j^P(\mathbf{a})(f(a_j) - \mu(a_j)), \int_{\mathcal{A}} f(a)\omega_A(a)\, \mathrm{d}a\right] = \sum_{j=1}^n \tau_j^P(\mathbf{a}) \left(\int_{\mathcal{A}} \sigma(a, a_j)\omega_A(a)\, \mathrm{d}a\right),$$

whereas the expression for α_1 follows from Theorem 1. □

Proof of Proposition 6. Suppose players agree on the relative relevance of attributes. First, it must be that $\omega_i(a) = 0 \Leftrightarrow \omega_{-i}(a) = 0$. Second, for any relevant attributes a, a', $\omega_P(a)/\omega_P(a') = \omega_A(a)/\omega_A(a')$. Rearranged, $\omega_P(a)/\omega_A(a) = \omega_P(a')/\omega_A(a')$ constant for any a, a'. Hence, it must be that $\omega_A(a) = \omega_0 \cdot \omega_P(a)$, where $\omega_0 \in \mathbb{R}$ a constant, for all $a \in \mathcal{A}$.

(i) Because $\omega_A(a) = \omega_0 \cdot \omega_P(a)$ for all $a \in \mathcal{A}$, for any $\mathbf{a} \in \mathcal{A}_k$,

$$\tau_j^A(\mathbf{a}) = \int_{\mathcal{A}} \tau_j(a;\mathbf{a})\omega_A(a)\,\mathrm{d}a = \int_{\mathcal{A}} \tau_j(a;\mathbf{a})\omega_0\omega_P(a)\,\mathrm{d}a = \omega_0\tau_j^P(\mathbf{a}).$$

By similar reasoning, $\tau^A(a_j) = \omega_0\tau^P(a_j)$. This implies that $\psi_A^2(\mathbf{a}) = \omega_0^2\psi_P^2(\mathbf{a})$. Hence, the single-player samples coincide.

(ii) Plugging $\tau^A(a_j) = \omega_0\tau^P(a_j)$ into (14), we have that $\alpha_2(\mathbf{a}) = \omega_0\psi_P^2(\mathbf{a})$. The optimal sample maximizes $2\alpha_2(\mathbf{a}) - \alpha_1(\mathbf{a}) = (2\omega_0 - 1)\psi_P^2(\mathbf{a})$. If $\omega_0 \geqslant 1/2$, then $V_A(\mathbf{a}) - V_A(\emptyset) \geqslant 0$. The agent's objective is maximized at any $\mathbf{a}^* = \mathbf{a}_P^s = \mathbf{a}_A^s$. Otherwise, if $\omega_0 < 1/2$, then $V_A(\mathbf{a}) - V_A(\emptyset) < 0$, hence the empty sample is the unique optimum. □

Lemma 4. *Fix a sample $\mathbf{a} = \{a_1, \ldots, a_n\} \in \mathcal{A}_k$ and suppose that σ satisfies NAP. The marginal value $V_A(\mathbf{a}) - V_A(\mathbf{a} \setminus \{a_j\})$ of any sample attribute $a_j \in \mathbf{a}$ equals*

$$\begin{cases} \tau_1^P(\mathbf{a})\left(2\tau_1^A(\mathbf{a}) - \tau_1^P(\mathbf{a})\right)\left(1 - \sigma^2(a_1,a_2)\right) & \text{if } j = 1 \\ \tau_j^P(\mathbf{a})\left(2\tau_j^A(\mathbf{a}) - \tau_j^P(\mathbf{a})\right)\dfrac{(1-\sigma^2(a_{j-1},a_j))(1-\sigma^2(a_j,a_{j+1}))}{1-\sigma^2(a_{j-1},a_{j+1})} & \text{if } 1 < j < n \\ \tau_n^P(\mathbf{a})\left(2\tau_n^A(\mathbf{a}) - \tau_n^P(\mathbf{a})\right)\left(1 - \sigma^2(a_{n-1},a_n)\right) & \text{if } j = n. \end{cases}$$

For any optimal sample $\mathbf{a}^*$ and $a_j \in \mathbf{a}^*$, $\tau_j^P(\mathbf{a}^*)$ and $\tau_j^A(\mathbf{a}^*)$ have the same sign.

Proof of Lemma 4. Suppose σ satisfies NAP, and let $\mathbf{a} \setminus \{a_j\} =: \mathbf{a}_{-j}$ for $a_j \in \mathbf{a}$. First, removing a_1 from $\mathbf{a}^*$ only affects the sample weights of a_1 and a_2. Therefore, $\alpha_2(\mathbf{a}^*) - \alpha_2(\mathbf{a}^*_{-1})$ equals

$$\begin{aligned}
&\tau_1^P(\mathbf{a}^*)\tau^A(a_1) + \tau_2^P(\mathbf{a}^*)\tau^A(a_2) - \tau_1^P(\mathbf{a}^*_{-1})\tau^A(a_2) \\
&= \tau_1^P(\mathbf{a}^*)\tau^A(a_1) - \sigma(a_1,a_2)\tau_1^P(\mathbf{a}^*)\tau^A(a_2) \\
&= \tau_1^P(\mathbf{a}^*)\left(\int_0^{a_1}\sigma(a,a_1)\omega_A(a)\,\mathrm{d}a + \int_{a_1}^{a_2}\sigma(a,a_1)\frac{1-\sigma^2(a,a_2)}{1-\sigma^2(a_1,a_2)}\omega_A(a)\,\mathrm{d}a\right)\left(1-\sigma^2(a_1,a_2)\right) \\
&= \tau_1^P(\mathbf{a}^*)\tau_1^A(\mathbf{a}^*)\left(1-\sigma^2(a_1,a_2)\right),
\end{aligned}$$

where the first equality substitutes for $\tau_2^P(\mathbf{a}^*) - \tau_2^P(\mathbf{a}^*_{-1})$, as in the proof of Proposition 1, and the second equality uses the fact that for any $a > a_2 > a_1$, $\sigma(a,a_1) - \sigma(a_1,a_2)\sigma(a,a_2) = 0$, and for any $a < a_1 < a_2$, $\sigma(a,a_2) - \sigma(a_1,a_2)\sigma(a,a_1) = 0$. Proposition 1 gives that $\alpha_1(\mathbf{a}^*) - \alpha_1(\mathbf{a}^*_{-1}) = \tau_1^P(\mathbf{a}^*)^2\left(1-\sigma^2(a_1,a_2)\right)$ as well. Therefore,

$$V_A(\mathbf{a}^*) - V_A(\mathbf{a}^*_{-1}) = \tau_1^P(\mathbf{a}^*)\left(2\tau_1^A(\mathbf{a}^*) - \tau_1^P(\mathbf{a}^*)\right)\left(1-\sigma^2(a_1,a_2)\right).$$

The mirror argument is used for the case of $V_A(\mathbf{a}^*) - V_A(\mathbf{a}^*_{-n})$. Similarly, for $j = 2, \ldots, n-1$,

$$\alpha_2(\mathbf{a}^*) - \alpha_2(\mathbf{a}^*_{-j}) = \tau_j^P(\mathbf{a}^*)\tau_j^A(\mathbf{a}^*)\frac{(1-\sigma^2(a_{j-1},a_j))(1-\sigma^2(a_j,a_{j+1}))}{1-\sigma^2(a_{j-1},a_{j+1})}.$$

Combining this with Proposition 1, we have

$$V_A(\mathbf{a}^*) - V_A(\mathbf{a}^*_{-j}) = \tau_j^P(\mathbf{a}^*)\left(2\tau_j^A(\mathbf{a}^*) - \tau_j^P(\mathbf{a}^*)\right)\frac{(1-\sigma^2(a_{j-1},a_j))(1-\sigma^2(a_j,a_{j+1}))}{1-\sigma^2(a_{j-1},a_{j+1})}.$$

□

Proof of Corollary 2. Immediate from Lemma 4. □

Proof of Proposition 7. Let $\underline{a} := \min\{\underline{a}_A, \underline{a}_P\}$ and $\bar{a} := \max\{\bar{a}_A, \bar{a}_P\}$. Suppose $\mathbf{a}^* = \{a_1, \ldots, a_n\}$ is optimal. First, we establish that $a_1 \geqslant \underline{a}$ and $a_n \leqslant \bar{a}$. Second, we show that if $[\underline{a}_A, \bar{a}_A] \cap [\underline{a}_P, \bar{a}_P] = \emptyset$, there exists no $a \in \mathbf{a}^*$ such that $a \in (\bar{a}_A, \underline{a}_P)$ or $a \in (\bar{a}_P, \underline{a}_A)$.

Towards a contradiction, suppose $a_1 < \underline{a}$. If $n = 1$, then $\tau^i(a_1) = \sigma(a_1, \underline{a})\tau^i(\underline{a})$ for $i = A, P$. Therefore, $0 < \alpha_2(a_1) = \sigma(a_1, \underline{a})^2\alpha_2(\underline{a}) < \alpha_2(\underline{a})$, where the first inequality follows from the optimality of a_1 and the last one follows from $a_1 < \underline{a}$. Similarly, $0 \leqslant \alpha_1(a_1) = \sigma(a_1, \underline{a})^2\alpha_1(\underline{a}) < \alpha_1(\underline{a})$. But $2\alpha_2(a_1) - \alpha_1(a_1) = \sigma^2(\underline{a}, a_1)(2\alpha_2(\underline{a}) - \alpha_1(\underline{a})) < 2\alpha_2(\underline{a}) - \alpha_1(\underline{a})$. Hence, $\{\underline{a}\}$ is strictly preferred to $\mathbf{a}^*$. If $n > 1$, Corollary 2 implies that $a_2 \geqslant \underline{a}$. Then, $\tau_1^A(\mathbf{a}^*)\tau_1^P(\mathbf{a}^*)\left(1 - \sigma^2(a_1, a_2)\right)$ simplifies to

$$\tau_1^A(\mathbf{a}^*)\tau_1^P(\mathbf{a}^*)\left(1 - \sigma^2(a_1, a_2)\right) = \frac{\sigma^2(a_1, \underline{a})}{1 - \sigma^2(a_1, \underline{a})\sigma^2(\underline{a}, a_2)} \prod_{i=A,P} \left(\int_{\underline{a}}^{a_2} \sigma(\underline{a}, a)\left(1 - \sigma^2(a, a_2)\right)\omega_i(a)\,\mathrm{d}a\right)$$

$$:= \frac{\sigma^2(a_1, \underline{a})}{1 - \sigma^2(a_1, \underline{a})\sigma^2(\underline{a}, a_2)} C_A C_P,$$

where C_A, C_P do not depend on a_1. Applying Lemma 4, the marginal value of a_1 is then

$$\frac{\sigma^2(a_1, \underline{a})}{1 - \sigma^2(a_1, \underline{a})\sigma^2(\underline{a}, a_2)}(2C_A C_P - C_P^2).$$

If it is optimal to sample at all, it must be that $(2C_A C_P - C_P^2) \geqslant 0$, hence this marginal value is increasing in a_1. The agent is better off sampling $a_1 = \underline{a}$ instead.

Second, suppose $[\underline{a}_A, \bar{a}_A] \cap [\underline{a}_P, \bar{a}_P] = \emptyset$ and let $\bar{a}_A < \underline{a}_P$ without loss. By Corollary 2, if there exists $a_j \in \mathbf{a}^*$ such that $a_j \in (\bar{a}_A, \underline{a}_P)$, then it must be that $a_1 \in (\bar{a}_A, \underline{a}_P)$ and $n = 1$. By NAP, $\alpha_2(a_1)$ does not depend on a_1 because $\sigma(a_1, \underline{a}_P)\sigma(a_1, \bar{a}_A) = \sigma(\bar{a}_A, \underline{a}_P)$, hence

$$\tau^A(a_1)\tau^P(a_1) = \sigma(\bar{a}_A, \underline{a}_P)\left(\int_{\underline{a}_A}^{\bar{a}_A} \omega_A(a)\sigma(a, \bar{a}_A)\,\mathrm{d}a\right)\left(\int_{\underline{a}_P}^{\bar{a}_P} \omega_P(a)\sigma(a, \underline{a}_P)\,\mathrm{d}a\right).$$

By a similar reasoning, $\alpha_1(a_1) = \sigma^2(a_1, \underline{a}_P)\alpha_1(\underline{a}_P)$, which decreases in a_1 because $\sigma^2(a_1', \underline{a}_P) = \sigma^2(a_1', a_1)\sigma^2(a_1, \underline{a}_P) < \sigma^2(a_1, \underline{a}_P)$ for any $a_1' < a_1 < \underline{a}_P$. This contradicts the optimality of $\mathbf{a}^*$. ☐

Proof of Proposition 8. (i) From Proposition 7 for any $a_j \in \mathbf{a}^*$, $a_j \in [\underline{a}_i, \bar{a}_i]$ for some $i = A, P$. By Corollary 2 and NAP, there can exist at most one attribute $a_1 \in (\underline{a}_i, \underline{a}_{-i})$ such that $\tau_1^{-i}(\mathbf{a}^*) \neq 0$, for otherwise player $-i$ would ignore all but the one closest to $\underline{a}_{-i}$. Analogously, there exists at most one attribute in $(\bar{a}_i, \bar{a}_{-i})$.

(ii) Suppose $[\underline{a}_A, \bar{a}_A] \cap [\underline{a}_P, \bar{a}_P] = \emptyset$. Because all relevant attributes are idiosyncratic, part (i) implies that $|\mathbf{a}^*| \leq 2$ for any optimal sample $\mathbf{a}^*$ and k. By contradiction, suppose there exists an optimal sample $\mathbf{a}^* = \{a_1^*, a_2^*\}$ such that $a_1^*, a_2^* \in \mathrm{supp}(\omega_i)$ for some player i. Then, either $\tau_1^{-i}(\mathbf{a}^*) = 0$ or $\tau_2^{-i}(\mathbf{a}^*) = 0$, which violates Corollary 2. ☐

Proof of Proposition 9. Without loss, let $\underline{a}_P \leqslant \bar{a}_A$ and $\underline{a}_A \leqslant \bar{a}_P$, so that the set of common attributes has positive measure. If the optimal sample $\mathbf{a}^*$ includes either no common attributes or one common attribute, then by Proposition 8, either $|\mathbf{a}^*| \leqslant 2$ or $|\mathbf{a}^*| \leqslant 3$ respectively. Suppose the optimal sample consists of at least two common attributes $a_j^* < a_{j+1}^*$ and $|\mathbf{a}^*| < k$. Consider sampling $\tilde{a} \in (a_j, a_{j+1})$. The sample weight of such $\tilde{a}$ in $\mathbf{a}^* \cup \{\tilde{a}\}$ is generically non-zero and it is identical for both players because $\omega_A(a) = \omega_P(a)$ for any $a \in (a_j, a_{j+1})$. By Lemma 4, the marginal value of $\tilde{a}$ is strictly positive, which contradicts the optimality of $\mathbf{a}^*$. Hence, if $\mathbf{a}^*$ has at least two common attributes, then $|\mathbf{a}^*| = k$ and $\mathbf{a}^*$ contains at least $k - 2$ common attributes. ☐

References

Aghion, Philippe, Patrick Bolton, Christopher Harris, and Bruno Jullien. 1991. "Optimal Learning by Experimentation." *The Review of Economic Studies*, 58(4): 621–654.

Allcott, Hunt. 2015. "Site Selection Bias in Program Evaluation." *Quarterly Journal of Economics*, 130(3): 1117–1165.

Banerjee, Abhijit V, Sylvain Chassang, Sergio Montero, and Erik Snowberg. 2020. "A theory of experimenters: Robustness, randomization, and balance." *American Economic Review*, 110(4): 1206–30.

Bardhi, Arjada. 2018. "Optimal Discovery and Influence Through Selective Sampling."

Bardhi, Arjada, and Nina Bobkova. 2023. "Local evidence and diversity in minipublics." *Journal of Political Economy*, 131(9): 000–000.

Barrett, Wayne W., and Philip J. Feinsilver. 1978. "Gaussian Families and a Theorem on Patterned Matrices." *Journal of Applied Probability*, 15(3): 514–522.

Beals, Richard, David H Krantz, and Amos Tversky. 1968. "Foundations of Multidimensional Scaling." *Psychological Review*, 75(2): 127.

Billot, Antoine, Itzhak Gilboa, and David Schmeidler. 2008. "Axiomatization of an Exponential Similarity Function." *Mathematical Social Sciences*, 55(2): 107–115.

Branco, Fernando, Monic Sun, and J. Miguel Villas-Boas. 2012. "Optimal Search for Product Information." *Management Science*, 58(11): 2037–2056.

Brocas, Isabelle, and Juan D. Carrillo. 2007. "Influence Through Ignorance." *The RAND Journal of Economics*, 38(4): 931–947.

Callander, Steven. 2011. "Searching and Learning by Trial and Error." *American Economic Review*, 101: 2277–2308.

Callander, Steven, and Patrick Hummel. 2014. "Preemptive Policy Experimentation." *Econometrica*, 82(4): 1509–1528.

Callander, Steven, and Tom S. Clark. 2017. "Precedent and Doctrine in a Complicated World." *American Political Science Review*, 111(1): 184–203.

Cambanis, Stamatis. 1973. "On Some Continuity and Differentiability Properties of Paths of Gaussian Processes." *Journal of Multivariate Analysis*, 3(4): 420–434.

Che, Yeon-Koo, and Navin Kartik. 2009. "Opinions as Incentives." *Journal of Political Economy*, 117(5): 815–860.

Clemen, Robert T., and Robert L. Winkler. 1985. "Limits for the Precision and Value of Information from Dependent Sources." *Operations Research*, 33(2): 427–442.

Di Tillio, Alfredo, Marco Ottaviani, and Peter Norman Sørensen. 2017*a*. "Persuasion Bias in Science: Can Economics Help?" *The Economic Journal*, 127: 266–304.

Di Tillio, Alfredo, Marco Ottaviani, and Peter Norman Sørensen. 2017*b*. "Strategic Sample Selection."

Eliaz, Kfir, and Alexander Frug. 2018. "Bilateral Trade with Strategic Gradual Learning." *Games and Economic Behavior*, 107: 380–395.

Garfagnini, Umberto, and Bruno Strulovici. 2016. "Social Experimentation with Interdependent and Expanding Technologies." *The Review of Economic Studies*, 83(4): 1579–1613.

Geng, Sen, Leonardo Pejsachowicz, and Michael Richter. 2017. "Breadth versus Depth."

Gilboa, Itzhak, Offer Lieberman, and David Schmeidler. 2006. "Empirical Similarity." *The Review of Economics and Statistics*, 88(3): 433–444.

Glazer, Jacob, and Ariel Rubinstein. 2004. "On Optimal Rules of Persuasion." *Econometrica*, 72(6): 1715–1736.

Henry, Emeric, and Marco Ottaviani. 2019. "Research and the Approval Process: The Organization of Persuasion." *American Economic Review*, 109(3): 911–55.

Hirsch, Alexander V. 2016. "Experimentation and Persuasion in Political Organizations." *American Political Science Review*, 110(1): 68–84.

Huber, Joel, and John McCann. 1982. "The Impact of Inferential Beliefs on Product Evaluations." *Journal of Marketing Research*, 19(3): 324–333.

Ilut, Cosmin, and Rosen Valchev. 2023. "Economic agents as imperfect problem solvers." *The Quarterly Journal of Economics*, 138(1): 313–362.

Jovanovic, Boyan, and Rafael Rob. 1990. "Long Waves and Short Waves: Growth Through Intensive and Extensive Search." *Econometrica*, 58(6): 1391–1409.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101(6): 2590–2615.

Keeney, Ralph L., and Howard Raiffa. 1976. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* New York:Wiley.

Ke, T. Tony, Zuo-Jun Max Shen, and J. Miguel Villas-Boas. 2016. "Search for Information on Multiple Products." *Management Science*, 62(12): 3576–3603.

Klabjan, Diego, Wojciech Olszewski, and Asher Wolinsky. 2014. "Attributes." *Games and Economic Behavior*, 88: 190–206.

Lancaster, Kelvin J. 1966. "A New Approach to Consumer Theory." *Journal of Political Economy*, 74: 132–157.

Liang, Annie, Xiaosheng Mu, and Vasilis Syrgkanis. 2021. "Dynamically Aggregating Diverse Information."

Mason, Kevin, and Joyce Bequette. 1998. "Product Experience and Consumer Product Attribute Inference Accuracy." *Journal of Consumer Marketing*.

Mason, Kevin, Thomas Jensen, Scot Burton, and Dave Roach. 2001. "The Accuracy of Brand and Attribute Judgments: The Role of Information Relevancy, Product Experience, and Attribute-Relationship Schemata." *Journal of the Academy of Marketing Science*, 29(3): 308–318.

Neeman, Zvika. 1995. "On Determining the Importance of Attributes with a Stopping Problem." *Mathematical Social Sciences*, 29: 195–212.

Nevo, Aviv. 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica*, 69(2): 307–342.

Olszewski, Wojciech, and Asher Wolinsky. 2016. "Search for an Object with Two Attributes." *Journal of Economic Theory*, 161: 145–160.

Rasmussen, Carl Edward, and Christopher K. I. Williams. 2006. *Gaussian Processes for Machine Learning.* Cambridge, Massachusetts:MIT Press.

Rayo, Luis, and Ilya Segal. 2010. "Optimal Information Disclosure." *Journal of Political Economy*, 118(5): 949–987.

Sanjurjo, Adam. 2017. "Search with Multiple Attributes: Theory and Empirics."

Scheibehenne, Benjamin, Bettina von Helversen, and Jörg Rieskamp. 2015. "Different Strategies for Evaluating Consumer Products: Attribute-and Exemplar-based Approaches Compared." *Journal of Economic Psychology*, 46: 39–50.

Shavell, Steven. 1994. "Acquisition and Disclosure of Information Prior to Sale." *The RAND Journal of Economics*, 25(1): 20–36.

Shepard, Roger N. 1987. "Toward a Universal Law of Generalization for Psychological Science." *Science*, 237(4820): 1317–1323.

Sher, Itai. 2011. "Credibility and Determinism in a Game of Persuasion." *Games and Economic Behavior*, 71: 409–419.

Spiegler, Ran. 2016. "Choice Complexity and Market Competition." *Annual Review of Economics*, 8: 1–25.

Srinivasan, Venkat, and Allan D Shocker. 1973. "Estimating the Weights for Multiple Attributes in a Composite Criterion Using Pairwise Judgments." *Psychometrika*, 38(4): 473–493.

Walters, Daniel J, and Hal E Hershfield. 2020. "Consumers Make Different Inferences and Choices When Product Uncertainty Is Attributed to Forgetting Rather than Ignorance." *Journal of Consumer Research*, 47(1): 56–78.