NBER WORKING PAPER SERIES

MORE PATHS OR MORE CONTRAST?
A THEORY OF EXPERIMENTATION BREADTH

Joshua S. Gans
Luca Gius

Working Paper 35207
http://www.nber.org/papers/w35207

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2026

Thanks to Nel Dutt, Hanna Halaburda, and Scott Stern for valuable feedback. Thanks to Refine.ink, ChatGPT 5.5, and Claude Opus 4.7 for research assistance. Gans thanks the SSHRC for funding. Responsibility for all errors remains our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w35207

More Paths or More Contrast? A Theory of Experimentation Breadth
Joshua S. Gans and Luca Gius
NBER Working Paper No. 35207
May 2026
JEL No. D81, D83, L26, O31

ABSTRACT

How should an organisation choose the breadth of its experimentation portfolio? Breadth has two distinct margins—the number of paths kept alive, and the degree of contrast among them—and prior research has largely studied them in isolation. We bring them into a single framework and show that they need not move together. Under a fixed experimentation budget, adding paths creates more chances to find a strong direction, but it also dilutes learning across paths and weakens the strongest feasible contrast. When the task is primarily ranking among already-viable alternatives, broader portfolios become more attractive as the budget rises. When paths share common viability uncertainty, and experimental signals track payoff relatedness, however, additional paths partly repeat the same viability test rather than provide independent information. We identify conditions under which testing exactly two sharply contrasting paths is optimal, dominating both a single deep test and broader portfolios. The framework reconciles competing prescriptions—many parallel shots versus a few sharp comparisons—by clarifying when each applies, and shows why empirical measures of breadth should not treat the number of options and their relatedness as separable margins.

Joshua S. Gans
University of Toronto
Rotman School of Management
and NBER
joshua.gans@rotman.utoronto.ca

Luca Gius
Bocconi University
luca.gius@unibocconi.it

1 Introduction

Consider Biobot Analytics, a start-up built around the idea that wastewater could be transformed into population-health data (Kluender et al., 2021). The underlying technology had shown promise, but the commercial opportunity was broad. Even after focusing on prescription and illicit-drug monitoring, the founders faced several possible customer paths. One path was to sell national surveillance data to federal agencies such as the CDC, DEA, or FDA. A second was to help state and municipal governments, or public-health departments, monitor opioid consumption and evaluate outreach programs. A third was to sell wastewater-derived data to private customers such as hedge funds. The first two paths are close variants: both test public-sector demand for opioid-related public-health data. The private-sector path uses the same wastewater platform, but tests a different buyer, value proposition, and data-scale requirement. The founders can run only a limited number of serious pilots before deciding whether to scale the venture or abandon it: should they test all three paths, or focus on fewer? And if so, which ones?

This paper studies how an organisation should choose the breadth of such an experimentation portfolio. Breadth has two distinct margins. It can mean the number of live paths (N). It can also mean the degree of contrast among those paths (ρ). A path is a strategic hypothesis: a market, use case, customer segment, technology configuration, channel, or business-model variant. Contrast captures how differently those hypotheses resolve uncertainty. Two paths are close when they tend to succeed or fail for the same reasons. They are contrasting when their outcomes separate competing explanations of where value lies.[1]

The distinction matters because expanding breadth along one margin can weaken it along the other. With a fixed experimentation budget, keeping more paths alive creates more chances to discover a strong direction, but it also leaves less information to be generated about each individual path. And a larger menu need not produce proportionately richer learning: if additional paths are close variants of existing ones, they add options while contributing little new information about where value lies. By contrast, a smaller set of deliberately differentiated paths can generate sharper inference, though at the cost of exploring fewer possibilities. The design problem is, therefore, not simply whether to experiment more or less broadly, but which combination of number and contrast makes best use of the available tests. Breadth is thus a point on a number–contrast frontier rather than a single scalar.

[1] In their textbook on entrepreneurial strategy, Gans et al. (2024) argue that startups should formulate two distinct strategies that are distinguished by value creation and value capture hypotheses based on disjoint assumptions about uncertain states.

We build a Gaussian model to study this frontier. Before learning, the agent chooses an experimentation set, including both the number of live paths and their dependence structure. The agent then observes noisy signals generated by a fixed budget T, forms posterior means, and implements the best surviving path. Managers choose markets, prototypes, customer segments, and experimental designs, and the correlation matrix summarises how those choices make payoffs and signals move together.

Our central benchmark is the regular-simplex frontier: N paths that are maximally negatively equicorrelated. On this frontier three named forces operate. A *selection effect* favours more paths, because the maximum of more draws is larger in expectation. A *precision effect* favours fewer paths, because the fixed budget is spread thinner. A *contrast effect* also favours fewer paths, because maximal feasible negative dependence becomes less extreme as N grows. As the budget T rises, precision dilution becomes less costly, so the optimum tilts toward broader portfolios at the cost of milder contrast; when resources are scarce, it shifts toward fewer, more sharply contrasting paths.

We then relax the assumption that payoff differences are purely path-specific and instead assume that each path's payoff contains a common component, representing whether the underlying idea is viable, and a residual component, representing path-specific fit. This limits total payoff contrast, raises the possibility that no path will be ultimately viable, and makes abandonment after experimentation relevant. Under the covariance-preserving signal technology and the residual-simplex benchmark, common uncertainty tilts the optimal choice toward lower N: for intermediate common uncertainty, the agent reconciles the need to learn viability and to rank paths by testing exactly two contrasting options.

Together, these results help reconcile seemingly different prescriptions on how many options organisations should explore. The reconciliation is conditional. A Nelson-style logic of parallel experimentation, which emphasises many independent shots on goal (Nelson, 1961), is better suited to environments in which experimentation is primarily a ranking problem and the budget is large. Entrepreneurial strategy advice emphasising a small number of sharply contrasting options (Gans et al., 2019) is better suited to environments in which viability uncertainty is substantial and the experiments' signals are themselves related in the same way as the payoffs, so additional paths partly repeat the same viability test rather than provide independent assays of the common component.

The paper makes three contributions. First, it models experimentation breadth as the joint design of number and relatedness. This differs from extant work that studies the return to increasing one breadth margin while taking the other as fixed. Second, it characterises the trade-off between these margins through the independent benchmark, the regular-simplex frontier, and the corresponding fixed-N correlation-design upper envelope. The frontier shows why independent-shots reasoning can overstate optimal breadth when the agent can select or design contrast, and why empirical studies should not treat the number of projects, markets, or knowledge sources and their variety as separable measures of breadth: expanding breadth along one margin can compress the other. Third, it characterises conditions that favour different sides of the trade-off and identifies a formal region in which testing exactly two contrasting paths is optimal. This provides a benchmark microfoundation for the "test two, choose one" prescription in Gans et al. (2019): exact for the $\{1, 2, 3\}$ comparison and benchmark-based for larger residual-simplex menus. We also use a reduced-form cost discussion and the empirical implications to discipline, rather than broaden, the managerial reading.

Our analysis is related to several literatures. The closest is the literature on entrepreneurial experimentation and strategy under uncertainty (Gruber et al., 2008; Levinthal, 2017; Gans et al., 2019; Koning et al., 2022; Camuffo et al., 2020; Felin et al., 2020; Agrawal et al., 2026). That work studies how entrepreneurs learn about uncertain opportunities, design tests, and revise strategic commitments. We focus on a prior design choice: which paths enter the experiment, how many of them are kept alive, and how much contrast the tested set contains. Including an outside option connects the model to work on staged commitment and abandonment. Our main focus, however, is not the timing of continuation or exit, but the ex ante composition of the experimentation portfolio.

Second, we contribute to the broader innovation-search and portfolio-strategy literatures by shifting attention from trade-offs within a single notion of breadth to trade-offs across distinct notions of breadth (Harrison and Klein, 2007). Although prior work recognises that breadth is multifaceted—whether in the form of more draws, broader knowledge sourcing, broader innovation objectives, or greater variety of perspectives—it typically studies the returns to expanding one breadth margin at a time (Leiponen and Helfat, 2010; Boudreau et al., 2011; Klingebiel and Rammer, 2014; Dutt and Lawrence, 2022; Shermon and Moeen, 2022; Chattopadhyay et al., 2025). We instead place two classic rationales for breadth in a common framework: breadth as more draws (Nelson, 1961) and breadth as greater dissimilarity among retained alternatives (Prahalad and Bettis, 1986). Our model shows that these margins need not be complementary: expanding breadth along one margin can compress it along the other, making some combinations difficult to sustain. Organisations with different budgets, cost structures, or opportunity menus may therefore occupy different points on the number–contrast frontier. Echoing prior work showing that larger menus need not be better menus (Goldreich and Hałaburda, 2013), organisations may learn no more, and sometimes less, from many loosely differentiated alternatives than from a smaller menu of sharper tests. This helps explain why different dimensions of breadth often appear beneficial in isolation yet weak or negatively interacting when pursued simultaneously (e.g., Leiponen and Helfat (2010); Srinivasan et al. (2021)). The implication is that the relevant strategic question is not (only) whether to search broadly, but how.

Third, we connect two literatures that each capture one side of our problem. Work on exploration and adaptive learning emphasises that organisations can be trapped in familiar regions of the opportunity space, biased against risky alternatives, and that unconventional options may be valuable precisely because they reveal possibilities that would otherwise remain untested (Denrell and March, 2001; Denrell et al., 2004; Csaszar and Levinthal, 2016; Bernardo and Welch, 2001; Rao, 2009; Gius, 2025; Manso, 2011). A separate literature on parallel experimentation and selectionism studies the costs and benefits of pursuing multiple approaches at once (Thomke et al., 1998; Loch et al., 2001; Sommer and Loch, 2004). We bring these arguments together by showing that the number of parallel experiments and the unconventionality of the alternatives tested need not be complements: under a fixed budget, adding more paths can dilute learning and reduce contrast, so a broader parallel portfolio may make it harder, not easier, to learn from genuinely unconventional alternatives.

The rest of the paper is organised as follows. Section 2 presents the model and the finite-menu interpretation. Section 3 studies the baseline trade-off between the number of paths and the contrast among them. Section 4 adds common viability uncertainty and the outside option, and characterises the region in which two paths are optimal among one, two, and three. Section 5 discusses experimentation-cost incidence in reduced form. Section 6 draws out managerial and empirical implications, and concludes.

2 Set-up

We consider an agent choosing an experimentation portfolio before observing any signals. The available opportunities are normalised to have common prior mean zero and unit variance.[2]

For a tested set of size N, payoffs are

$$\theta = (\theta_1, \dots, \theta_N)^\top \sim \mathsf{N}(0, R), \qquad R \in \mathcal{C}_N,$$

[2]The mean-zero normalisation centres payoffs relative to the relevant outside option introduced below. The unit-variance normalisation keeps attention on dependence rather than scale heterogeneity across paths.

where

$$\mathcal{C}_N = \{R \in \mathbb{R}^{N \times N} : R = R^\top,\ R \succeq 0,\ R_{ii} = 1 \text{ for all } i\}.$$

Here, R is the correlation matrix among the live opportunities. A high off-diagonal element means that two paths tend to succeed or fail for similar reasons. A low, or negative, off-diagonal element means that the paths are more contrasting. Economically, negative dependence should be read as a pair of competing hypotheses about where value lies, not merely as different labels attached to similar pilots.

Conditional on (N, R), the agent spends an experimentation budget $T > 0$ and observes

$$Y = \theta + \varepsilon, \qquad \varepsilon \perp \theta, \qquad \varepsilon \sim \mathsf{N}\left(0, \frac{N}{T} R\right).$$

This signal structure captures equal depth across live paths: when more paths are kept alive, each receives precision T/N. The assumption that signal noise has the same correlation structure as payoffs is a covariance-preserving reduced form: it makes posterior uncertainty a scalar contraction of prior uncertainty and lets us isolate the portfolio-design problem without introducing a separate experimental-technology choice.[3] In the special case where the menu consists of independent paths ($R = I_N$), the noise matrix is diagonal and each path is observed with independent precision T/N – the standard equal-budget benchmark in the parallel-experimentation literature (Nelson, 1961).

Gaussian updating implies that the posterior mean vector is distributed as[4]

$$\mu \sim \mathsf{N}\left(0, \frac{T}{T+N} R\right).$$

Thus, in the baseline ranking problem, the value of a tested portfolio is

$$V(N, R; T) = \sqrt{\frac{T}{T+N}}\, \mathbb{E}\left[\max_{1 \le i \le N} Z_i\right], \qquad Z \sim \mathsf{N}(0, R).$$

which depends both on N and the contrast embedded in R. At the boundary, $V(N, R; 0) = 0$: with no experimentation budget, the agent's posterior coincides with the prior and the live-path maximum carries no value. In the limit $T \to \infty$, the precision factor approaches

[3] For a general noise covariance Σ_N, the posterior mean satisfies $\mu \sim \mathsf{N}(0,\ R(R + \Sigma_N)^\dagger R)$, where $\dagger$ denotes the Moore–Penrose inverse. The scalar contraction below obtains only in the covariance-preserving case $\Sigma_N = (N/T)R$. The common-viability results should therefore be read as a benchmark for experiments whose signal relatedness tracks payoff relatedness; they do not cover technologies in which each pilot provides an independent direct measurement of a common viability factor.

[4] This expression remains valid when R is singular. In that case θ, ε, and Y are supported on the range of R, and the Gaussian projection is understood on that range.

Model object	Entrepreneurial interpretation
Path i	A live strategic hypothesis, such as a market, use case, customer segment, product variant, channel, or business-model configuration.
Breadth N	The number of hypotheses kept alive during the experiment.
Correlation R_{ij}	The extent to which two paths succeed or fail for the same underlying reasons.
Low or negative R_{ij}	A pair of contrasting hypotheses that separate competing accounts of where value lies.
Budget T	Scarce experimental capacity: runway, founder attention, engineering time, customer pilots, or managerial bandwidth.
Common component ν	Uncertainty about whether the underlying idea, technology, or business model is viable at all.
Outside option	Abandoning, pivoting, or conserving resources rather than scaling any tested path.

Table 1: Managerial interpretation of the main primitives. The model treats correlations as summaries of selected strategic hypotheses and experimental designs, not as literal objects managers manipulate directly.

one and value depends only on selection and contrast, $\mathbb{E}[\max_i Z_i]$. The interesting region of the model lies between these two extremes.

2.1 Managerial primitives and feasible menus

The primitive managerial choice is not a correlation matrix by itself: an organisation chooses a set of strategic hypotheses and designs pilots around them, and the correlation matrix is a reduced-form summary of how those choices make payoffs and signals covary. Table 1 records the mapping; Table 2 records the hierarchy of analytical objects this paper uses.

A finite-menu version makes the mapping explicit. Suppose the organisation has a menu $\mathcal{M} = \{1, \dots, K\}$ of candidate paths with correlation matrix $R^{\mathcal{M}}$. If it tests a subset $S \subseteq \mathcal{M}$, the relevant correlation matrix is the principal submatrix $R^{\mathcal{M}}_S$, and the baseline value of that subset is

$$V(S;T) = \sqrt{\frac{T}{T+|S|}}\,\mathbb{E}\left[\max_{i\in S} Z_i\right], \qquad Z_S \sim \mathsf{N}(0, R^{\mathcal{M}}_S).$$

The finite-menu problem is therefore $\max_{S\subseteq\mathcal{M}} V(S;T)$, or, if breadth is fixed at N, the same maximisation subject to $|S| = N$. The fixed-N correlation-design problem studied below

Object	Role in the paper	Formal status
Finite-menu subset problem	The closest representation of an organisation choosing actual paths from an opportunity menu.	Exact for any specified menu, but no universal closed-form solution is claimed.
Fixed-N correlation-design upper envelope	The fixed-N problem over all feasible $R \in \mathcal{C}_N$.	Proposition 1 identifies it with a Gaussian mean-width problem.
Regular-simplex frontier	Symmetric maximal-contrast benchmark for ranking problems.	Exact in known simplex cases; used as a benchmark where the general mean-width problem is unresolved.
Residual-simplex frontier with common viability	Symmetric benchmark for joint viability-and-ranking problems.	Pair optimality is exact over $N \in \{1, 2, 3\}$ and extends to larger menus under the residual-simplex restriction.

Table 2: Hierarchy of analytical claims. The paper's managerial interpretation uses the finite-menu object, while the main comparative statics are benchmark results unless exactness is stated.

replaces the finite set of feasible submatrices with all of $\mathcal{C}_N$; it is an upper envelope, and actual opportunity menus lie weakly below it. For applications, the finite-menu problem says what the manager must assess not only the expected payoff of each path, but also whether the paths test distinct assumptions. A menu containing three close variants may be effectively narrower than a menu containing two sharply opposed hypotheses.

Three maintained assumptions keep the benchmark focused. First, for fixed N, the correlation-design results treat contrast as selectable from a rich menu of feasible opportunities or experimental designs; under a finite menu, the value functions still apply, but the feasible correlation matrices are restricted to the submatrices generated by that menu. Second, the covariance-preserving signal structure makes payoff relatedness and signal relatedness coincide, so additional paths do not generate independent direct measurements of viability. Third, equal budget allocation rules out asymmetric designs, such as one deep viability pilot paired with a lighter contrast pilot. These assumptions make the model a clean benchmark for the number–contrast frontier, not a claim that every organisation can freely implement every covariance matrix or every experimental technology.

Running Example — Set-up

For Biobot, let

$$\mathsf{Fed} = \text{federal agencies}, \qquad \mathsf{Mun} = \text{municipal governments}, \qquad \mathsf{HF} = \text{hedge funds}.$$

The full menu has correlation matrix

$$R_{\mathsf{FedMunHF}} = \begin{pmatrix} 1 & \rho_{\mathsf{FedMun}} & \rho_{\mathsf{FedHF}} \\ \rho_{\mathsf{FedMun}} & 1 & \rho_{\mathsf{MunHF}} \\ \rho_{\mathsf{FedHF}} & \rho_{\mathsf{MunHF}} & 1 \end{pmatrix}.$$

We would expect ρ_{FedMun} to be relatively high because federal agencies and municipal governments both test public-sector demand for opioid-related wastewater data. They differ in scale and decision problem: federal agencies use the data for national epidemiology and surveillance, while municipal governments use it to monitor local consumption, evaluate outreach, and allocate resources. The hedge-fund path is more distant, so ρ_{FedHF} and ρ_{MunHF} are likely to be lower: it tests private-sector willingness to pay for alternative data and a different monetisation logic.

Choosing a tested set $S \subseteq \{\mathsf{Fed}, \mathsf{Mun}, \mathsf{HF}\}$ selects the corresponding principal submatrix R_S. Thus $S = \{\mathsf{Mun}, \mathsf{HF}\}$ and $S = \{\mathsf{Fed}, \mathsf{Mun}, \mathsf{HF}\}$ differ not only in the number of paths, but also in the covariance structure through which the evidence is compared.

3 Baseline: More Paths versus More Contrast

We first set common viability uncertainty to zero. Under the regular-simplex portfolios characterised below, posterior means sum to zero; hence the outside option is never exercised. The baseline is, therefore, a pure ranking problem – after experimentation, the agent chooses the best live path – and, as we show below, the environment most favourable to broad parallel experimentation.

3.1 Independent paths

A natural benchmark treats the live paths as independent. Setting $R = I_N$, the agent chooses

$$N_0^\star(T) \in \arg\max_{N\geq 2} \underbrace{m_N}_{\text{selection effect}} \underbrace{\sqrt{\frac{T}{T+N}}}_{\text{precision effect}}, \qquad m_N = \mathbb{E}\left[\max_{1\leq i\leq N} Z_i\right],$$

where $Z_1, \ldots, Z_N$ are independent standard normal random variables. The selection effect m_N rises with N, because the maximum is taken over more draws. The precision effect falls with N, because the fixed budget is divided across more paths. The independent benchmark, therefore, formalises the usual breadth–depth trade-off.

Lemma 1 (Independent benchmark)**.** *For $T > 0$, let*

$$W_N^0(T) = m_N\sqrt{\frac{T}{T+N}}, \qquad \mathcal{A}_0(T) = \arg\max_{N\geq 2} W_N^0(T).$$

The set $\mathcal{A}_0(T)$ is nonempty and finite. If

$$\underline{N}_0(T) = \min \mathcal{A}_0(T), \qquad \overline{N}_0(T) = \max \mathcal{A}_0(T),$$

then both $\underline{N}_0(T)$ and $\overline{N}_0(T)$ are weakly increasing in T. Moreover, every independent-benchmark optimiser satisfies $N \geq 5$.

The proofs of all results are in Appendix A. The interpretation of Lemma 1 is that, even with independent paths, the breadth–depth trade-off is decided by which of two opposing forces dominates. The selection effect favours larger N, because the maximum of more draws is larger in expectation; the precision effect favours smaller N, because each added path reduces the precision with which any one path is measured. Scarce resources alone are not enough to drive the optimum to the smallest feasible menus: the lemma shows that every independent-benchmark optimiser keeps at least five paths alive, regardless of how tight the budget becomes. To put it differently, an independent-shots logic, on its own, never recommends a narrow portfolio.

That said, the independent benchmark is incomplete. It treats independence as the natural expression of breadth. In many strategic settings, however, the decision maker can choose or select paths that are more or less related. A founder can test two related public-sector customer paths, such as federal agencies and municipal governments, or can pit a municipal-government path against a hedge-fund data-product path. The next subsection shows why this second margin matters.

3.2 The value and feasibility of contrast

Holding N fixed, contrast is valuable because it increases the dispersion of posterior means. But contrast is not freely scalable: the feasible amount of negative dependence depends on how many paths are kept alive. The following lemma records these two facts.

Lemma 2 (Value and feasibility of contrast)**.** *Fix $N \geq 2$, and let $m_N = \mathbb{E}[\max_{1 \leq i \leq N} X_i]$, where $X_1, \ldots, X_N$ are independent standard normal random variables.*

1. *If $Z \sim \mathsf{N}(0, R)$ and R is equicorrelated — that is, with all off-diagonal entries equal to a common ρ — so that $R_{ii} = 1$ and $R_{ij} = \rho$ for $i \neq j$, then*

$$\mathbb{E}\left[\max_i Z_i\right] = m_N \sqrt{1-\rho}.$$

2. *For any $R \in \mathcal{C}_N$, if*

$$\bar{\rho} \equiv \frac{1}{N(N-1)} \sum_{i \neq j} R_{ij},$$

then

$$\bar{\rho} \geq -\frac{1}{N-1}.$$

For a fixed number of paths in the equicorrelated case, a lower ρ raises value. It makes the paths move less together, so the best posterior mean is more likely to stand out from the rest. But contrast is not freely scalable. Lemma 2 shows that a pair can have correlation -1, a triple can do no better than average correlation $-1/2$, and as N grows, the lowest feasible average correlation moves towards zero. Thus, more paths reduce the strongest attainable average contrast.

Running Example — Contrast

A two-path design can deliberately pit municipal governments against hedge funds. At maximal contrast, the two posterior means move in opposite directions: evidence favouring one customer hypothesis weakens the other. With three paths, the strongest symmetric design can no longer make every pair perfectly opposed: the best equicorrelated benchmark has pairwise correlation $-1/2$. If the federal-agency and municipal-government paths are close variants in the actual menu, the realised contrast of the triple is weaker still.

3.3 The regular-simplex benchmark

Once both N and R are choice variables, the agent's problem ranges over all portfolio breadths and all feasible correlation structures. Holding N fixed, the unrestricted choice of R is equivalent to a longstanding mean-width problem in convex geometry.

Proposition 1 (Fixed-breadth correlation design as Gaussian mean width)**.** *Fix $N \geq 2$. Then*

$$\sup_{R \in \mathcal{C}_N} \mathbb{E}\left[\max_{1 \leq i \leq N} Z_i\right] = \sup_{u_1,\ldots,u_N \in S^{N-1}} \mathbb{E} h_{\mathrm{conv}\{u_1,\ldots,u_N\}}(G),$$

where $Z \sim \mathsf{N}(0, R)$, $G \sim \mathsf{N}(0, I_N)$, and $h_K(g) = \sup_{x \in K} \langle g, x \rangle$ is the support function of K. Equivalently, the fixed-N correlation-design problem is the problem of maximising the mean width of the convex hull of N unit vectors.

The interpretation of Proposition 1 is that the unrestricted fixed-breadth design problem is, in geometric language, the problem of arranging N unit vectors so that the convex hull they span is as wide as possible from a typical Gaussian direction. Proposition 1 thereby separates the unrestricted fixed-breadth design problem from the regular-simplex benchmark used below. The unrestricted problem is the simplex mean-width problem. In our notation, the regular-simplex candidate has value

$$m_N \sqrt{\frac{N}{N-1}},$$

with the value attained by the covariance matrix whose off-diagonal entries are all $-1/(N-1)$. The regular-simplex solution is known for $N \leq 4$, and a recent preprint gives a proposed proof in all dimensions; the published literature, however, still often treats the general statement as a conjecture. None of the comparative statics below requires resolving that external convex-geometric problem for every N. For $N \geq 5$, the analysis should be read as a regular-simplex benchmark: the symmetric design with maximal feasible negative equicorrelation.[5]

For fixed $N \geq 2$, the regular-simplex correlation matrix is

$$R^{\mathrm{simp}}_{N,ii} = 1, \qquad R^{\mathrm{simp}}_{N,ij} = -\frac{1}{N-1} \quad (i \neq j).$$

[5] For the equivalence between the simplex mean-width conjecture and the Gaussian expected-maximum formulation, see Litvak (2018). For the $N = 4$ Gaussian maximum result, see Sun et al. (2020). For computational evidence for the $N \leq 8$ cases in the related Gaussian signal-set-design and weak-simplex literature, see Sun (1997). For asymptotic evidence, see Kabluchko et al. (2017). For a proposed general proof, see Goldsmith (2023). The regular simplex also maximises the determinant of the centred information matrix for fixed N, and is therefore optimal when the objective is to maximise information gain rather than the expected maximum.

This matrix attains the average-correlation lower bound in Lemma 2, and hence places the N-path portfolio at the strongest symmetric contrast allowed by feasibility. Applying Lemma 2 with $\rho = \rho_N^* = -1/(N-1)$ gives

$$\mathbb{E}\left[\max_{1\le i\le N} Z_i\right] = m_N\sqrt{1-\rho_N^*} = m_N\sqrt{\frac{N}{N-1}}.$$

Under this benchmark, the remaining problem is one-dimensional: the agent chooses where to sit on the regular-simplex frontier. Combining the simplex value above with the posterior precision factor gives the following decomposition.

$$\overline{V}_N(T) = \underbrace{m_N}_{\text{selection effect}} \underbrace{\sqrt{\frac{T}{T+N}}}_{\text{precision effect}} \underbrace{\sqrt{\frac{N}{N-1}}}_{\text{contrast effect}}, \qquad N \ge 2.$$

The selection effect favours more paths: the maximum of more draws is larger in expectation. The precision effect favours fewer paths: each added path reduces learning precision. The contrast effect also favours fewer paths: maximal feasible contrast is strongest for small N. These three named forces recur throughout the rest of the paper; the analysis to follow is largely a study of which of them dominates as the environment changes.

Two implications follow directly. First, independence is not the right benchmark when contrast can be chosen. Since lowering correlation raises value at fixed N in the symmetric case, and the regular simplex is the natural benchmark for maximal symmetric contrast, an independent-shots benchmark misses a margin of choice. Second, a higher budget makes broader portfolios more attractive.

Proposition 2 (Budget and breadth on the regular-simplex frontier)**.** *For $T > 0$, let*

$$\mathcal{A}(T) = \arg\max_{N\ge 2} \overline{V}_N(T).$$

The set $\mathcal{A}(T)$ is nonempty and finite. Moreover, if

$$\underline{N}(T) = \min \mathcal{A}(T), \qquad \overline{N}(T) = \max \mathcal{A}(T),$$

then both $\underline{N}(T)$ and $\overline{N}(T)$ are weakly increasing in T. Equivalently, the optimal-breadth correspondence is increasing: as the budget rises, the set of optimal breadths cannot move to smaller N.

The interpretation of Proposition 2 is that the selection and contrast effects in $\overline{V}_N$ do not depend on T; only the precision effect does. As T grows, precision becomes locally cheap, and the resulting reweighting tilts the optimum toward larger N. The contrast cost of larger N — through the $\sqrt{N/(N-1)}$ factor — does not vanish; it is simply less binding once each path is well measured. Formally, for any $K > N$, the ratio $\overline{V}_K(T)/\overline{V}_N(T)$ has one term independent of T and one term, $\sqrt{(T+N)/(T+K)}$, that increases in T. Hence larger portfolios become relatively more attractive as experimentation resources grow, and an organisation that can afford more experimentation will, on the regular-simplex frontier, prefer to spend that capacity on more paths rather than on sharper contrast.

Before turning to viability, it is useful to record how stark the regular-simplex baseline is in the absence of common uncertainty. Can resource scarcity alone justify testing only two paths? The following result shows that it cannot.

Corollary 1 (Three beats two on the regular-simplex frontier)**.** *In the baseline regular-simplex benchmark,*

$$\overline{V}_3(T) > \overline{V}_2(T) \qquad \textit{for every } T > 0.$$

The interpretation of Corollary 1 is that, in the pure ranking problem, the single contrast lost from moving from a pair to a triple is more than offset by the extra ranking draw, regardless of the budget. Put differently, when contrast can be chosen freely from a rich menu, the selection effect dominates the contrast effect at $N = 3$ versus $N = 2$. This result is in tension with both entrepreneurial guidance — which suggests testing two options before committing (Gans et al., 2019) — and entrepreneurial practice: in Gruber et al. (2008), over three quarters of entrepreneurs considered at most two total market opportunities before committing. To explain the discrepancy, in the next section we focus on the case where experimentation needs to establish viability in addition to ranking paths.

4 Establishing Viability

We now add the feature that is central in many entrepreneurial settings. The agent may not know whether the underlying idea is worth pursuing at all. Experimentation must therefore serve two purposes: it must rank the live paths, and it must determine whether any path should be implemented.

4.1 Common uncertainty and exit

For a tested set of size $N \geq 2$, let $s = (s_1, \ldots, s_N)^\top \sim \mathsf{N}(0, Q)$, with $Q \in \mathcal{C}_N$. We model viability by decomposing each payoff into a common component and a path-specific component:

$$\theta_i = \alpha\nu + \sqrt{1-\alpha^2}\, s_i, \qquad \nu \sim \mathsf{N}(0,1), \qquad \nu \perp s, \qquad \alpha \in [0,1]. \tag{1}$$

Here, ν is the value of the underlying idea. The residual s_i is the path-specific payoff component. The parameter α is the common-viability loading; the common component's share of each path's prior variance is α^2. When $\alpha = 0$, experimentation is purely about ranking paths. When $\alpha = 1$, all uncertainty is common, so the live paths rise or fall together. The equal-loading assumption means that every live path is equally informative about the common viability component. In the Biobot mapping, equal loading says that the federal-agency, municipal-government, and hedge-fund pilots are each as informative as the others about whether wastewater analytics is viable as a business at all; they differ only in residual customer-fit.

The outside option is normalised to zero. After learning, the agent implements the best live path only if its posterior mean is positive. Thus the decision rule is

$$\max\{0, \mu_1, \ldots, \mu_N\}.$$

Both ingredients matter. Without the outside option, common uncertainty simply shifts all paths together and does not by itself make fewer paths attractive. Without common uncertainty, the outside option does not bind under the simplex design because the posterior means sum to zero. Together, the two ingredients turn experimentation into a joint viability-and-ranking problem. The model assumes that exit is available after experimentation and that the outside option is not distorted by escalation pressures, sunk-cost commitments, or other organisational frictions.

Running Example — Viability

For Biobot, ν is whether wastewater analytics can become a viable business at all. For $j \in \{\mathsf{Fed}, \mathsf{Mun}, \mathsf{HF}\}$, write

$$\theta_j = \alpha\nu + \sqrt{1-\alpha^2}\, s_j.$$

The residuals s_{Fed}, s_{Mun}, and s_{HF} capture customer-specific fit: federal surveillance demand, municipal adoption and resource-allocation value, and hedge-fund willingness to pay for alternative data.

A negative common signal lowers all three posterior means and can make exit optimal; conditional on continuing, the residual terms determine which customer path to scale.

Throughout this section, the signal technology remains the covariance-preserving one introduced in Section 2. Thus, adding a path does not create an independent assay of ν. It adds another payoff-relevant signal whose common and residual components inherit the same dependence structure as the payoff vector. If a separate viability experiment directly measured ν, or if path-level pilots produced independent viability information, the value of broader portfolios could be higher and the thresholds below would no longer describe the relevant design problem.

The equal-loading structure implies

$$R = \alpha^2 \mathbf{1}\mathbf{1}^\top + (1-\alpha^2)Q.$$

Common uncertainty, therefore, limits feasible contrast. If residual components can be arranged as a regular simplex, the lowest feasible total equicorrelation at breadth N is

$$\underline{\rho}_N(\alpha) = \alpha^2 - \frac{1-\alpha^2}{N-1} = \frac{N\alpha^2 - 1}{N-1}. \tag{2}$$

In words, when more of the payoff variance is common, it becomes harder to make live paths sharply opposed in total payoff: the common component is shared and cannot be cancelled out, so even residuals that are perfectly opposed leave the total payoffs partly aligned. As α rises, the lower bound $\underline{\rho}_N(\alpha)$ rises accordingly. Under the residual-simplex design, residual uncertainty is pure ranking uncertainty: residual posterior means sum to zero. Residual "all paths are bad" risk is therefore represented only through the common component in this benchmark.

4.2 Choosing two paths over one and three

We now solve the minimal comparison in which the test-two prescription has content. A singleton can test viability but cannot compare paths. A pair can both test viability and preserve a residual comparison. A triple adds residual ranking value, but also dilutes precision. Thus, testing two is optimal only if it beats both the singleton and the first broader portfolio. The comparison below is exact for $N \in \{1, 2, 3\}$; larger menus are discussed below under the residual-simplex benchmark.

For a singleton, the posterior mean is normally distributed with the usual learning scale, so

$$V_1(T, \alpha) = \sqrt{\frac{T}{T+1}} \frac{1}{\sqrt{2\pi}}.$$

For a pair, the optimal residual design uses perfect negative dependence. Its value is

$$V_2(T,\alpha) = \sqrt{\frac{T}{T+2}}F_2(\alpha), \qquad F_2(\alpha) = \frac{1+\sqrt{1-\alpha^2}}{\sqrt{2\pi}}.$$

The first term in F_2 is viability value; the second is residual ranking value. As α rises, residual ranking value falls.

The triple comparison requires one additional step. With exit, it is not immediate that the best three-path design remains symmetric: one might suspect that the agent should keep a sharply contrasting pair and add a third, less related path. The next result rules this out.

Proposition 3 (Optimal triple under common viability)**.** *Fix $\alpha \in (0,1)$. Among all feasible three-path correlation matrices consistent with* (1)*, the unique optimal triple is*

$$R_3^\star(\alpha) = \alpha^2 \mathbf{1}\mathbf{1}^\top + (1-\alpha^2)\begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{pmatrix}.$$

The optimised triple value is

$$V_3(T,\alpha) = \sqrt{\frac{T}{T+3}}F_3(\alpha),$$

where

$$F_3(\alpha) = \frac{3}{\pi\sqrt{2\pi}}\left[\arctan(\sqrt{3}\,\alpha) + \sqrt{3(1-\alpha^2)}\left(\frac{\pi}{4} + \frac{1}{2}\arctan\frac{\sqrt{1-\alpha^2}}{2\alpha}\right)\right].$$

Proposition 3 shows that common uncertainty does not make the optimal triple a pair plus an unrelated third path. After the common component is removed, the three residual paths are still placed as evenly as possible. What changes with α is not the shape of the optimal triple, but its residual ranking value.

We can now compare one, two, and three paths. The pair value $F_2(\alpha)$ and the triple value $F_3(\alpha)$ trade off in opposite directions as common uncertainty rises: the pair preserves a sharp residual contrast but loses one ranking draw, while the triple adds the draw at the cost of a milder residual contrast and a smaller precision share. The next proposition pins down the exact thresholds at which the optimal breadth shifts from three to two and from two to one.

Proposition 4 (Viability and the return to two)**.** *Fix $T > 0$. Define*

$$\Gamma(\alpha) = \frac{F_3(\alpha)}{F_2(\alpha)}.$$

There is a unique threshold $\alpha^\star_{23}(T) \in (0,1)$ satisfying

$$\Gamma(\alpha^\star_{23}(T)) = \sqrt{\frac{T+3}{T+2}},$$

and the optimised pair beats the optimised triple if and only if

$$\alpha > \alpha^\star_{23}(T).$$

There is also a unique threshold

$$\alpha^\star_{12}(T) = \sqrt{1 - \left(\sqrt{\frac{T+2}{T+1}} - 1\right)^2},$$

and the optimised pair beats the singleton if and only if

$$\alpha < \alpha^\star_{12}(T).$$

Moreover,

$$\alpha^\star_{23}(T) < \alpha^\star_{12}(T).$$

Proposition 4 gives the exact region in which testing two is optimal among one, two, and three paths. Ignoring knife-edge ties, the optimal breadth among $\{1,2,3\}$ is

$$N^\star_{\{1,2,3\}}(T,\alpha) = \begin{cases} 3, & \alpha < \alpha^\star_{23}(T), \\ 2, & \alpha^\star_{23}(T) < \alpha < \alpha^\star_{12}(T), \\ 1, & \alpha > \alpha^\star_{12}(T). \end{cases}$$

At the two thresholds, the adjacent portfolios tie. For low common uncertainty, the problem is mostly about ranking, and three paths dominate. For intermediate common uncertainty, the third path contributes too little residual ranking value to justify the loss of precision, while a singleton is still too narrow. This is the test-two region. For very high common uncertainty, even the pair has little residual ranking value, and a single deep viability test dominates. Figure 1 illustrates the economic pattern: as the common loading α rises, the

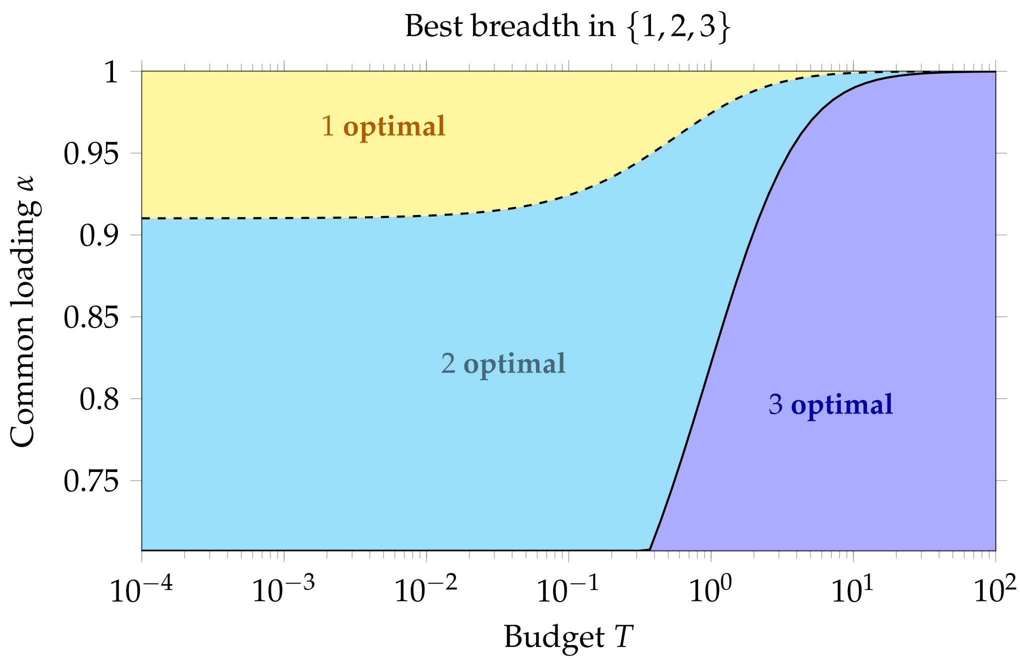

Figure 1: Optimal breadth among one, two, and three paths as a function of the experimentation budget T and common uncertainty α. The solid boundary is the pair–triple indifference curve; the dashed boundary is the pair–singleton indifference curve. The vertical axis begins at $1/\sqrt{2}$; below that value, three paths are optimal.

optimal breadth moves from three to two to one.

Running Example — Why Two

Let the candidate pair be $S_2 = \{\mathsf{Mun}, \mathsf{HF}\}$, and let the triple be $S_3 = \{\mathsf{Fed}, \mathsf{Mun}, \mathsf{HF}\}$. In the pair-optimal region, municipal governments and hedge funds provide enough contrast to rank customer paths if wastewater analytics is viable, while avoiding the precision loss from also testing a second public-sector variant. If α is low, the triple wins because the task is mostly ranking. If α is very high, a singleton wins because the task is mostly viability.

4.3 Larger menus

The comparison above is exact for the minimal nontrivial menu: one path, two paths, and the first broader portfolio. To ask whether a pair can beat all larger portfolios, we extend the residual-simplex benchmark to every $N \leq |\mathcal{N}|$. For $N \geq 2$, let M_N^{simp} denote the maximum of an N-dimensional regular-simplex Gaussian vector, i.e. a centred Gaussian vector with unit variances and pairwise correlations $-1/(N-1)$. Let $Z_0 \sim \mathsf{N}(0,1)$ be independent of M_N^{simp}. Define

$$V_N^{\text{simp}}(T,\alpha) = \sqrt{\frac{T}{T+N}}\mathbb{E}\left[\left(\alpha Z_0 + \sqrt{1-\alpha^2}\, M_N^{\text{simp}}\right)^+\right], \qquad N \geq 2,$$

and keep

$$V_1(T) = \sqrt{\frac{T}{T+1}}\frac{1}{\sqrt{2\pi}}.$$

The residual-simplex optimal-breadth correspondence is

$$\mathcal{A}_{\text{RS}}(T,\alpha) = \underset{1\leq N\leq|\mathcal{N}|}{\arg\max}\, V_N^{\text{simp}}(T,\alpha),$$

where $V_1^{\text{simp}}(T,\alpha) \equiv V_1(T)$.

Proposition 5 (Larger menus under the residual-simplex benchmark)**.** *Fix a finite menu size $|\mathcal{N}| \geq 3$ and $T > 0$. Under the residual-simplex benchmark:*

1. *the optimal-breadth correspondence $\mathcal{A}_{\text{RS}}(T,\alpha)$ is decreasing in α: both its smallest and largest elements are weakly decreasing in α;*

2. *there are parameter regions in which testing exactly two paths is uniquely optimal. More precisely, for all sufficiently large T, there is a nonempty interval $I_T^\infty \subset (0,1)$ on which $\mathcal{A}_{\text{RS}}(T,\alpha) = \{2\}$, and both endpoints of I_T^∞ converge to one as $T \to \infty$. There is also a nonempty interval $I^0 \subset (0,1)$, lying near*

$$\alpha_0 = \sqrt{1-(\sqrt{2}-1)^2} = \sqrt{2\sqrt{2}-2} \approx 0.910,$$

such that $\mathcal{A}_{\text{RS}}(T,\alpha) = \{2\}$ for every $\alpha \in I^0$ and every sufficiently small $T > 0$.

The interpretation of Proposition 5 is that, in contrast with the baseline model, when the agent needs to establish viability, it may be optimal to test exactly two contrasting options even against the full menu. Part (i) highlights that the test-two region survives at high budgets: when common uncertainty is extreme and the budget is large, additional paths buy almost no residual ranking value, and a sharply contrasting pair dominates wider portfolios. Part (ii) shows the complementary force at low budgets: when resources are scarce and common uncertainty is intermediate, the precision effect is biting and the pair preserves the strongest residual contrast that can still be afforded. The second case, in particular, is aligned with many entrepreneurial settings, where agents have scarce resources and learning helps to decide whether to pursue the idea, not only which strategy to pursue (Gans et al., 2019).

5 Experimentation Costs

The formal results above compare portfolios with different N and ρ under a common experimentation budget T. They do not endogenise the cost of activating paths or the cost of making paths contrastive. This section, therefore, has a narrower role: it records how different cost margins enter the design problem in reduced form. The comparative statics in this section should be read as disciplined incidence logic, not as additional structural propositions.

First, activating another path can be costly. A pilot may require a specialised prototype, hiring a dedicated expert, or developing a specific customer relationship. If c is the per-path activation cost, a reduced-form objective is

$$\max_{N,\rho} \; V_N^{\text{eq}}(T, \alpha, \rho) - cN,$$

where V_N^{eq} denotes the value in a symmetric equicorrelation specification and ρ is the *total* equicorrelation among path payoffs. Under the common-factor model, feasibility requires

$$\rho \in \left[\underline{\rho}_N(\alpha), 1\right], \qquad \underline{\rho}_N(\alpha) = \alpha^2 - \frac{1-\alpha^2}{N-1}.$$

For a fixed value function and feasible set, a higher c makes broader experimentation less attractive. This force pushes towards fewer live paths. In the running example, activating the hedge-fund path may require a sales pipeline entirely different from a municipal-government pilot. If that activation cost is large, the hedge-fund path must provide enough additional contrast to justify it.

Second, contrast itself can be costly. Running several approximately independent pilots can often be delegated to separate teams. By contrast, engineering a set of deliberately contrasting pilots requires managerial coordination: someone must specify the hypotheses, separate the tasks, and ensure that teams are not simply exploring close variants of the same logic. This is the cost counterpart of strategic variety: the managerial burden is not the number of activities alone, but the need to sustain distinct strategic logics (Prahalad and Bettis, 1986).

It is useful to state the contrast-effort formulation in residual rather than total correlations. Let e denote effort spent coordinating residual contrast, and suppose the induced residual equicorrelation satisfies

$$q_N(0) = 0, \qquad q_N'(e) < 0, \qquad \lim_{e \to \infty} q_N(e) = -\frac{1}{N-1}.$$

The corresponding total equicorrelation is

$$\rho_N^{\text{tot}}(e; \alpha) = \alpha^2 + (1 - \alpha^2) q_N(e).$$

Thus independent residuals imply $\rho_N^{\text{tot}}(0; \alpha) = \alpha^2$, and maximal residual contrast implies

$$\lim_{e \to \infty} \rho_N^{\text{tot}}(e; \alpha) = \underline{\rho}_N(\alpha).$$

A simple reduced-form objective is therefore

$$\max_{N,e} \; V_N^{\text{eq}}\big(T, \alpha, \rho_N^{\text{tot}}(e; \alpha)\big) - cN - he,$$

where h is the marginal cost of specifying and coordinating diagnostic contrast across the experiments.[6]

For fixed N, higher h lowers the optimal effort devoted to residual contrast and moves $q_N(e)$ back toward zero, which moves total correlation back toward α^2. A general comparative static for N does not follow from the reduced form alone. If independently delegated paths substitute for engineered contrast, higher h may lead to more, less contrasting paths. If broader portfolios themselves require more coordination, the effect on N can instead be muted or reversed.

The implication is that how experimenters solve the N–ρ trade-off depends not only on the level of experimentation costs, but also on their incidence. If costs mainly attach to activating another path—for example, building a prototype, recruiting a customer, or creating a new sales process—experimenters are primarily constrained in their choice of N, and a reduction in those costs expands the number margin. If, instead, costs mainly attach to producing contrast—for example, formulating discriminating hypotheses, designing

[6] One way to microfound $q_N(e)$ is to suppose that the manager chooses the precision of a public signal that helps employees select contrastive strategies (Van den Steen, 2010). Public guidance is correct with probability $p(e)$, with $p' > 0$ and $p'' < 0$. Employees load on the public signal by $m(e) \in [0, 1]$, with $m(e) = 0$ when the signal is too imprecise and $m'(e) > 0$ once it is useful. If correct guidance induces a target residual correlation $q_N^T < 0$ and incorrect guidance induces independent residuals, then $q_N(e) = p(e)m(e)q_N^T$. Taking $q_N^T = -1/(N-1)$ and $p(e)m(e) \to 1$ gives the limiting case above. The same reduced form can also be read as theory-development effort: e helps the manager identify which paths should contrast, by specifying shared assumptions, path-specific causal links, and discriminating evidence. This interpretation is consistent with evidence that theory-of-value training helps ventures to make more coordinated changes across core and operational elements of their business models (Agarwal et al., 2025).

tests that separate mechanisms, or coordinating interdependent experiments so that teams do not converge on close variants—experimenters are primarily constrained in their choice of q_N and hence ρ_N^{tot}, and a reduction in those costs expands the contrast margin. The two cost margins thus combine into one diagnostic observation.

Under the reduced-form objective above, suppose independently delegated paths substitute for engineered contrast and broader portfolios do not themselves create offsetting coordination costs. Then a fall in activation costs without a comparable fall in contrast-coordination costs tends to move the organisation toward higher N and weaker contrast. This is the cheap-parallelism region: many cheap pilots test similar hypotheses. The point is deliberately conditional. It identifies a class of technological and organisational shocks that can expand observed breadth while shrinking effective breadth, but it is not a general comparative-static theorem for all cost technologies.

6 Discussion and Conclusion

The analysis reframes experimentation breadth as a position on a number–contrast frontier. An organisation can broaden search by keeping more paths alive, N, or by choosing paths that contrast more sharply, that is, by lowering ρ. These are distinct margins, and under a fixed experimentation budget they can move in opposite directions. Adding paths spreads learning more thinly and, at the benchmark frontier, weakens the strongest attainable contrast; preserving sharper contrast often requires focusing resources on fewer paths. As Table 3 summarises, the experimentation budget and common viability uncertainty generate formal benchmark comparative statics, while activation costs and contrast-coordination costs identify reduced-form incidence forces. The common point is that breadth is not only a scale choice; it is also a composition choice.

Implications for managers. For managers, the model cautions against a universal answer to the question, "How broad should search be?" When the main uncertainty is ranking among alternatives whose viability is already credible, broad portfolios are attractive, especially when the experimentation budget is large. This is the Nelson-style logic of many shots on goal (Nelson, 1961). When candidate paths share substantial common uncertainty—whether the underlying idea is viable at all—and when pilots do not provide independent direct measurements of that common factor, a single pilot is noisy in a more strategic sense: it confounds the value of the idea with the fit of the particular market, customer, channel, or implementation being tested (Gans et al., 2019). In that region, the

Force	$N^\star$	$\rho^\star$	Interpretation
Budget $T \uparrow$	$\uparrow$	$\uparrow$	Precision dilution weakens: breadth shifts toward more paths and milder contrast.
Common uncertainty $\alpha \uparrow$ *with outside option*	$\downarrow$	mixed	Viability dominates ranking: extra paths add less residual ranking value, while common loading and lower N push correlation in opposite directions.
Activation cost $c \uparrow$	$\downarrow$	$\downarrow$	Active paths become costly: breadth shifts toward fewer, sharper comparisons.
Coordination cost $h \uparrow$	mixed	$\uparrow$	Engineered contrast becomes costly: when delegation substitutes for coordination, breadth shifts toward more paths and weaker contrast.

Table 3: Forces shaping the N–ρ trade-off. $\rho^\star \uparrow$ means weaker contrast. Budget and common-uncertainty rows summarise formal benchmark results; cost rows summarise reduced-form incidence forces.

relevant question is not how many experiments the organisation can afford, but whether the experiments are jointly informative: can they reveal whether the idea should be pursued and, conditional on continuation, which path should be scaled?

The key managerial challenge is therefore not simply securing resources to activate another path. It is recognising whether alternatives are merely different or genuinely *contrasting*, and then designing the experiments so that each result informs the interpretation of the others. Turning experimentation from a collection of isolated trials into a system of contrasts requires both coordination and theory. Coordination ensures that teams do not drift toward adjacent variants of the same bet. Theory specifies the rival hypotheses, identifies which comparison would be diagnostic, and explains why resolving that comparison matters for the venture.

Competitors with different budgets, activation costs, or coordination capabilities will not simply search more or less broadly; they will occupy different positions on the N–ρ frontier. When organisations are close to the frontier, increasing N requires accepting weaker contrast or less depth per path. Environments in which contrast is a poor substitute for additional paths—because the problem is primarily ranking among many viable alternatives—will be harder for resource-constrained entrepreneurs to navigate; environments in which a few sharp comparisons can resolve the key uncertainty are more favourable. Innovations that reduce activation costs (cheaper prototypes, easier customer tests) push toward higher N, while innovations that reduce theory and coordination costs (help formulating rival hypotheses, identifying diagnostic contrasts, running interdependent experiments) push toward lower ρ. The danger is that the first kind of innovation may arrive without the second. AI-enabled development tools, for example, may make MVPs cheaper and faster, raising the number of experiments ventures can run in parallel (Blank, 2026); whether they also lower the cost of ideating discriminating hypotheses or coordinating interdependent tests is a separate question. When activation falls without a comparable fall in coordination, organisations may engage in *cheap parallelism*: many experiments that are too similar to produce effective learning.

Implications for empirical work. Empirical studies frequently measure broad search using counts of projects, markets, knowledge sources, innovation objectives, partners, customer segments, or geographic areas. The model says these are projections of a jointly chosen portfolio, and a one-dimensional estimate of the return to breadth observes only one of those projections. The same N can mean scarce capacity, high activation costs, weak coordination of contrast, or the deliberate selection of a sharply contrasting pair; a large portfolio can mean broad exploration or many close variants of the same bet. The estimated return to a given breadth measure will therefore depend on whether the observed variation reflects movement toward the frontier or along it.

The opportunity is to study these margins jointly. If researchers can observe both the number of alternatives and the dependence among them, they can distinguish observed from effective breadth. Text-embedding measures of patent content (Arts et al., 2023) and landing pages (Guzman and Li, 2023) can help identify whether a large portfolio is semantically dispersed or locally clustered, but semantic dispersion is not the same object as structural contrast: the model's narrowing results require low or negative payoff dependence, so that success on one path weakens the posterior case for another. A stronger design would combine semantic measures with evidence on outcome co-movement, shared assumptions, or coded rival mechanisms.

The model then predicts when different breadth measures should appear complementary or substitutable. The return to more paths should be higher when viability is already credible and the main problem is ranking; the return to structural contrast should be higher when common viability uncertainty is substantial but not overwhelming. Market, knowledge-source, objective, partner, and project counts should interact negatively when they draw on the same experimental budget, and those interactions should weaken or reverse when organisations have larger budgets, better coordination, or stronger theories that let them sustain both more paths and sharper contrast. Weak or negative interactions among breadth measures need not imply that combining forms of breadth is intrinsically harmful; they may simply reveal the frontier along which organisations are choosing.

Limits and future research. The benchmark has clear limits. For fixed N, the correlation-design results are upper envelopes over a rich menu of correlation structures. Actual opportunities may come from a finite menu whose feasible correlations are far from the regular simplex. The common-factor model uses equal loadings; if some paths reveal much more about viability than others, the optimal portfolio need not be symmetric. The covariance-preserving signal technology rules out independent repeated measurements of the common factor and other noise structures. The model also suppresses mean, scale, and cost heterogeneity across opportunities. Those features matter for full-scale path selection, because a high-mean but correlated path or a low-cost but less contrasting path may be worth testing.

Several extensions remain open. We have solved the unrestricted rich-menu problem exactly for one, two, and three paths under equal loading on the common viability component. The corresponding viability problem for $N \geq 4$, and for heterogeneous common-factor loadings, remains outside the exact analysis. Sequential experimentation would allow the agent to drop paths and add new ones over time. A richer experimental technology would distinguish the cost of activating paths from the cost of making them genuinely contrasting, and would allow signal covariance to differ from payoff covariance. These extensions would determine how far the benchmark conclusions travel beyond the symmetric and covariance-preserving environments studied here.

For Biobot, the residual-simplex benchmark says that if the underlying viability of wastewater analytics as a business is highly uncertain and pilots mainly provide covariance-preserving evidence about that viability, the founders may be better off pairing the municipal-government and hedge-fund paths—two sharply contrasting bets—than running all three pilots in parallel. The third public-sector variant adds little residual ranking value while diluting precision and weakening the contrast between the public-sector and private-sector hypotheses. If, instead, viability were already credible and the founders were primarily ranking customer paths, or if each pilot independently measured common viability, the broader triple could dominate.

Breadth is, therefore, not just the number of experiments. It is the joint design of number and contrast. Testing two is not special because the number two has intrinsic value. It is special because, in the viability benchmark studied here, two paths can preserve a sharp comparison while avoiding the precision loss of a broader portfolio.

A Proofs

A.1 Posterior scaling

Let $\theta \sim \mathsf{N}(0, R)$ and $Y = \theta + \varepsilon$, where $\varepsilon \sim \mathsf{N}(0, \frac{N}{T}R)$ is independent of θ. Then

$$\operatorname{Var}(Y) = \left(1 + \frac{N}{T}\right) R, \qquad \operatorname{Cov}(\theta, Y) = R.$$

If R is singular, θ, ε, and Y are supported on $\operatorname{range}(R)$. Hence $RR^{+}Y = Y$ almost surely. The Gaussian projection formula, with the Moore–Penrose inverse on the support, gives

$$\mathbb{E}[\theta \mid Y] = R\left[\left(1 + \frac{N}{T}\right) R\right]^{+} Y = \frac{T}{T+N} RR^{+}Y = \frac{T}{T+N} Y.$$

Therefore

$$\operatorname{Var}(\mathbb{E}[\theta \mid Y]) = \frac{T^2}{(T+N)^2} \operatorname{Var}(Y) = \frac{T}{T+N} R.$$

This proves the posterior scaling used in the main text.

A.2 Proof of Lemma 1

Write

$$W_N^0(T) = m_N \sqrt{\frac{T}{T+N}}.$$

The maximiser exists. For any $\lambda > 0$, the log-sum-exp bound and Jensen's inequality give

$$m_N = \mathbb{E}\left[\max_i Z_i\right] \le \frac{1}{\lambda} \log \mathbb{E}\left[\sum_{i=1}^{N} e^{\lambda Z_i}\right] = \frac{\log N}{\lambda} + \frac{\lambda}{2}.$$

Choosing $\lambda = \sqrt{2 \log N}$ gives $m_N \le \sqrt{2 \log N}$, and therefore $W_N^0(T) \to 0$ as $N \to \infty$. Since $W_2^0(T) > 0$, only finitely many N can maximise the value.

For any $K > N$,

$$\frac{W_K^0(T)}{W_N^0(T)} = \frac{m_K}{m_N} \sqrt{\frac{T+N}{T+K}}.$$

The first factor is independent of T, while the second is strictly increasing in T. The same single-crossing argument used in the proof of Proposition 2 implies that the smallest and largest elements of $\mathcal{A}_0(T)$ are weakly increasing.

It remains to show that an optimiser cannot have $N < 5$. If $N < 5$, then

$$\frac{W_5^0(T)}{W_N^0(T)} = \frac{m_5}{m_N} \sqrt{\frac{T+N}{T+5}}$$

is increasing in T, so it is enough to compare the limit as $T \downarrow 0$. The exact values $m_2 =$

$1/\sqrt{\pi}$ and $m_3 = 3/(2\sqrt{\pi})$, together with the one-dimensional formula

$$m_N = N \int_{-\infty}^{\infty} z\varphi(z)\Phi(z)^{N-1}\,dz,$$

give the certified intervals

N	2	3	4	5
$m_N/\sqrt{N}$	0.39894	0.48860	(0.51468, 0.51469)	(0.52009, 0.52010).

The entries for $N = 4, 5$ are obtained by deterministic interval integration of the displayed normal integral; the tails outside $[-8, 8]$ are bounded by Mills' ratio and are below 10^{-12}. Hence $m_5/\sqrt{5} > m_N/\sqrt{N}$ for $N = 2, 3, 4$, and therefore $W_5^0(T) > W_N^0(T)$ for every $T > 0$ and every $N < 5$. Thus all independent-benchmark optimisers satisfy $N \geq 5$.

A.3 Proof of Lemma 2

For an equicorrelated matrix, feasibility is equivalent to $-1/(N-1) \leq \rho \leq 1$. Let $P = I_N - N^{-1}\mathbf{1}\mathbf{1}^\top$. Since

$$R = (1-\rho)I_N + \rho\mathbf{1}\mathbf{1}^\top,$$

we have

$$PRP = (1-\rho)P.$$

If $X_1, \ldots, X_N$ are independent standard normal random variables and $\bar{X} = N^{-1}\sum_i X_i$, then

$$Z - \bar{Z}\mathbf{1} \stackrel{d}{=} \sqrt{1-\rho}\,(X - \bar{X}\mathbf{1}),$$

where $\bar{Z} = N^{-1}\sum_i Z_i$. Since

$$\max_i Z_i = \bar{Z} + \max_i (Z_i - \bar{Z})$$

and $\mathbb{E}[\bar{Z}] = 0$,

$$\mathbb{E}\max_i Z_i = \sqrt{1-\rho}\,\mathbb{E}\max_i (X_i - \bar{X}) = \sqrt{1-\rho}\,\mathbb{E}\max_i X_i = m_N\sqrt{1-\rho}.$$

For the average-correlation bound, positive semidefiniteness gives

$$0 \leq \mathbf{1}^\top R\mathbf{1} = N + \sum_{i \neq j} R_{ij} = N + N(N-1)\bar{\rho}.$$

Rearranging yields $\bar{\rho} \geq -1/(N-1)$.

A.4 Proof of Proposition 1 and the regular-simplex value

Every $R \in \mathcal{C}_N$ is the Gram matrix of unit vectors $u_1, \ldots, u_N$ in a Euclidean space of dimension at most N. If G is a standard Gaussian vector in that space, then

$$(G \cdot u_1, \ldots, G \cdot u_N) \sim \mathsf{N}(0, R),$$

and therefore

$$\mathbb{E} \max_i (G \cdot u_i) = \mathbb{E} h_K(G), \qquad K = \operatorname{conv}\{u_1, \ldots, u_N\},$$

where h_K is the support function of K. Conversely, any collection of unit vectors $u_1, \ldots, u_N$ generates a matrix $R \in \mathcal{C}_N$ by $R_{ij} = u_i^\top u_j$. Hence the two suprema in Proposition 1 are equal. Since G is symmetric,

$$\mathbb{E}\big[h_K(G) + h_K(-G)\big] = 2\mathbb{E} h_K(G),$$

so the same optimisation can be read as the Gaussian mean-width problem for the convex hull of N unit vectors. This proves the reduction. No claim that the regular simplex solves the unrestricted mean-width problem for every N is used in the paper; the subsequent analysis evaluates the regular-simplex benchmark.

For that benchmark, let $X = (X_1, \ldots, X_N)$ have independent standard normal coordinates and let $\bar{X} = N^{-1} \sum_i X_i$. The regular-simplex correlation matrix satisfies

$$R_N^{\mathrm{simp}} = \frac{N}{N-1} \left(I_N - \frac{1}{N} \mathbf{1}\mathbf{1}^\top \right).$$

Thus

$$Z^{(N)} \equiv \sqrt{\frac{N}{N-1}} (X - \bar{X}\mathbf{1}) \sim \mathsf{N}(0, R_N^{\mathrm{simp}}).$$

Consequently,

$$\mathbb{E}\left[\max_i Z_i^{(N)}\right] = \sqrt{\frac{N}{N-1}}\, \mathbb{E}\left[\max_i (X_i - \bar{X})\right] = \sqrt{\frac{N}{N-1}}\, \mathbb{E}\left[\max_i X_i\right],$$

because $\max_i (X_i - \bar{X}) = \max_i X_i - \bar{X}$ and $\mathbb{E}[\bar{X}] = 0$. Multiplying by the posterior learning factor $\sqrt{T/(T+N)}$ gives

$$\overline{V}_N(T) = m_N \sqrt{\frac{T}{T+N}} \sqrt{\frac{N}{N-1}}, \qquad N \geq 2,$$

which is the baseline regular-simplex value used in the main text.

A.5 Proof of Proposition 2

Write

$$a_N = m_N \sqrt{\frac{N}{N-1}}, \qquad \overline{V}_N(T) = a_N \sqrt{\frac{T}{T+N}}.$$

The maximiser exists. Indeed, for any $\lambda > 0$,

$$m_N = \mathbb{E} \max_i Z_i \le \frac{1}{\lambda} \log \mathbb{E} \sum_{i=1}^{N} e^{\lambda Z_i} = \frac{\log N}{\lambda} + \frac{\lambda}{2}.$$

Choosing $\lambda = \sqrt{2 \log N}$ gives $m_N \le \sqrt{2 \log N}$. Since $\sqrt{N/(N-1)} \le \sqrt{2}$ for $N \ge 2$,

$$\overline{V}_N(T) \le 2\sqrt{\frac{T \log N}{T+N}} \to 0 \qquad \text{as } N \to \infty.$$

Because $\overline{V}_2(T) > 0$, only finitely many N can be optimal, so $\mathcal{A}(T)$ is nonempty and finite.

For any $K > N$,

$$\frac{\overline{V}_K(T)}{\overline{V}_N(T)} = \frac{a_K}{a_N}\sqrt{\frac{T+N}{T+K}}.$$

The first factor is independent of T. The second factor is strictly increasing in T, because

$$\frac{d}{dT} \log \sqrt{\frac{T+N}{T+K}} = \frac{1}{2}\left(\frac{1}{T+N} - \frac{1}{T+K}\right) > 0.$$

Thus, whenever a larger breadth K weakly beats N at some budget, K strictly beats N at any higher budget.

Let $T' > T$. If $\underline{N}(T') < \underline{N}(T)$, set $K = \underline{N}(T')$ and $N = \underline{N}(T)$. Since $N \in \mathcal{A}(T)$, $\overline{V}_N(T) \ge \overline{V}_K(T)$. By the single-crossing property, $\overline{V}_N(T') > \overline{V}_K(T')$, contradicting $K \in \mathcal{A}(T')$. Hence $\underline{N}(T') \ge \underline{N}(T)$. The same argument with $K = \overline{N}(T')$ and $N = \overline{N}(T)$ proves $\overline{N}(T') \ge \overline{N}(T)$.

A.6 Proof of Corollary 1

By the regular-simplex value formula,

$$\frac{\overline{V}_3(T)}{\overline{V}_2(T)} = \frac{m_3}{m_2}\sqrt{\frac{T+2}{T+3}}\sqrt{\frac{3/2}{2}}.$$

For two independent standard normal random variables,

$$m_2 = \mathbb{E}[\max\{Z_1, Z_2\}] = \frac{1}{\sqrt{\pi}}.$$

For three independent standard normal random variables,

$$m_3 = 3\int_{\mathbb{R}} x\varphi(x)\Phi(x)^2 \, dx = \frac{3}{2\sqrt{\pi}}.$$

Therefore
$$\frac{\overline{V}_3(T)}{\overline{V}_2(T)} = \frac{3\sqrt{3}}{4}\sqrt{\frac{T+2}{T+3}}.$$
This ratio is greater than one if and only if $27(T+2) > 16(T+3)$, which is equivalent to $11T + 6 > 0$. Thus $\overline{V}_3(T) > \overline{V}_2(T)$ for every $T > 0$.

A.7 Common-factor residual reduction

Fix $\alpha \in (0,1)$ and $N \geq 2$. A feasible correlation matrix R is consistent with the equal-loading common-factor representation if and only if
$$R = \alpha^2 \mathbf{1}\mathbf{1}^\top + (1-\alpha^2)Q$$
for some $Q \in \mathcal{C}_N$. Indeed, if $R - \alpha^2\mathbf{1}\mathbf{1}^\top \succeq 0$, then
$$Q = \frac{R - \alpha^2\mathbf{1}\mathbf{1}^\top}{1-\alpha^2}$$
is positive semidefinite and has unit diagonal. Conversely, every such Q produces a feasible R. If $Z_0 \sim \mathsf{N}(0,1)$ and $Y \sim \mathsf{N}(0,Q)$ are independent, then
$$X_i = \alpha Z_0 + \sqrt{1-\alpha^2}Y_i$$
has covariance matrix R.

A.8 Proof of Proposition 3

Write $s = \sqrt{1-\alpha^2}$. By the residual reduction, the three-path problem is equivalent, up to the scalar learning factor, to choosing $Q \in \mathcal{C}_3$ to maximise
$$\Phi_\alpha(Q) = \mathbb{E}\left[\left(\alpha Z_0 + s\max_{i\leq 3} Y_i\right)^+\right], \qquad Y \sim \mathsf{N}(0,Q).$$
For $m \in \mathbb{R}$, define
$$\psi_\alpha(m) = \mathbb{E}[(\alpha Z_0 + sm)^+].$$
Then
$$\psi_\alpha'(m) = s\,\Phi\Big(\frac{sm}{\alpha}\Big) > 0,$$
so the objective is a strictly increasing function of the residual maximum.

No positive definite residual matrix can be optimal. If $Q \succ 0$, reduce every off-diagonal element by a small common $\varepsilon > 0$. The resulting matrix Q^ε remains feasible for small ε, has the same marginal variances, and has lower pairwise covariances. Slepian's comparison theorem implies that $\max_i Y_i^\varepsilon$ first-order stochastically dominates $\max_i Y_i$, where $Y^\varepsilon \sim \mathsf{N}(0,Q^\varepsilon)$. The dominance is strict for some threshold because at least one covariance is reduced and both Gaussian vectors are nondegenerate. Since ψ_α is strictly increasing, $\Phi_\alpha(Q^\varepsilon) > \Phi_\alpha(Q)$. Hence the optimum lies on the boundary $\det Q = 0$.

Every boundary 3×3 correlation matrix can be represented by three unit vectors on the circle. Write $Q_{ij} = u_i^\top u_j$, with $u_i = (\cos\theta_i, \sin\theta_i)$, and let $G \sim \mathsf{N}(0, I_2)$. Then $Y_i = u_i^\top G$. This representation also covers rank-one boundary matrices by embedding the line in the plane. In polar form, $G = \Lambda(\cos\Theta, \sin\Theta)$, where Θ is uniform on $[0, 2\pi]$, Λ has Rayleigh density $re^{-r^2/2}$, and Λ and Θ are independent. Let

$$m(\theta) = \max_{i \le 3} \cos(\theta - \theta_i), \qquad H_\alpha(x) = \mathbb{E}[(\alpha Z_0 + s\Lambda x)^+].$$

The objective becomes

$$\Phi_\alpha(Q) = \frac{1}{2\pi} \int_0^{2\pi} H_\alpha(m(\theta))\, d\theta.$$

Moreover,

$$H_\alpha'(x) = s\,\mathbb{E}\left[\Lambda \Phi\left(\frac{s\Lambda x}{\alpha}\right)\right] > 0.$$

Order the three angles around the circle and let the circular gaps be $\delta_1, \delta_2, \delta_3$, with $\delta_1 + \delta_2 + \delta_3 = 2\pi$. On a gap of length δ, parameterise a direction by its circular distance $t \in [0, \delta]$ from the left endpoint. Since the gap contains no design angle, one of the two endpoints attains the largest inner product on that gap; the contribution is therefore

$$\int_0^\delta H_\alpha(\max\{\cos t, \cos(\delta - t)\})\ dt = 2\int_0^{\delta/2} H_\alpha(\cos t)\, dt.$$

Defining

$$\mathcal{G}_\alpha(x) = \int_0^x H_\alpha(\cos t)\, dt,$$

we have

$$\mathcal{G}_\alpha''(x) = -\sin x\, H_\alpha'(\cos x) < 0 \qquad \text{for } x \in (0, \pi).$$

Thus $\mathcal{G}_\alpha$ is strictly concave on $[0, \pi]$. Since $\sum_j \delta_j/2 = \pi$, Jensen's inequality implies that the sum of the three gap contributions is uniquely maximised by

$$\delta_1 = \delta_2 = \delta_3 = \frac{2\pi}{3}.$$

Hence the residual optimum is the regular simplex, $Q_{ij} = -1/2$ for $i \neq j$, giving

$$R_3^\star(\alpha) = \alpha^2 \mathbf{1}\mathbf{1}^\top + (1 - \alpha^2) R_3^{\mathrm{simp}}.$$

It remains to compute the value. At equal gaps,

$$F_3(\alpha) = \frac{3}{\pi} \int_0^{\pi/3} H_\alpha(\cos t)\, dt.$$

Conditional on $\Lambda = r$,

$$\mathbb{E}_{Z_0}[(\alpha Z_0 + srx)^+] = \alpha\varphi\Big(\frac{srx}{\alpha}\Big) + srx\Phi\Big(\frac{srx}{\alpha}\Big).$$

Integrating against the Rayleigh density gives

$$H_\alpha(x) = \frac{\alpha}{\sqrt{2\pi}} + \frac{sx}{\sqrt{2\pi}}\left[\frac{\pi}{2} + \arctan\Big(\frac{sx}{\alpha}\Big)\right].$$

Using

$$\int_0^{\pi/3}\left[\alpha + s\cos t \arctan\left(\frac{s\cos t}{\alpha}\right)\right]dt = \arctan(\sqrt{3}\,\alpha) + \frac{\sqrt{3}s}{2}\arctan\Big(\frac{s}{2\alpha}\Big),$$

which follows by integration by parts using $d[\arctan(s\cos t/\alpha)] = -\alpha s \sin t(\alpha^2 + s^2\cos^2 t)^{-1}dt$ and then the substitution $u = \tan t$. Since $\int_0^{\pi/3}\cos t\, dt = \sqrt{3}/2$, we obtain

$$F_3(\alpha) = \frac{3}{\pi\sqrt{2\pi}}\left[\arctan(\sqrt{3}\,\alpha) + \sqrt{3}\,s\left(\frac{\pi}{4} + \frac{1}{2}\arctan\frac{s}{2\alpha}\right)\right],$$

which is the stated expression.

A.9 Pair value and proof of Proposition 4

For a centred bivariate Gaussian vector (X_1, X_2) with unit variances and correlation ρ,

$$\max\{0, x, y\} + \max\{0, -x, -y\} = \frac{|x| + |y| + |x-y|}{2}.$$

By symmetry,

$$\mathbb{E}[\max\{0, X_1, X_2\}] = \frac{1}{2\sqrt{2\pi}}\left(2 + \sqrt{2(1-\rho)}\right).$$

Pair feasibility under the common-factor model requires $\rho \geq 2\alpha^2 - 1$, and the bivariate value is decreasing in ρ. The optimal pair therefore uses $\rho = 2\alpha^2 - 1$, which gives

$$F_2(\alpha) = \frac{1 + \sqrt{1-\alpha^2}}{\sqrt{2\pi}}.$$

Let $s = \sqrt{1-\alpha^2}$. From Proposition 3,

$$\Gamma(\alpha) = \frac{F_3(\alpha)}{F_2(\alpha)} = \frac{3}{\pi(1+s)}\left[\arctan(\sqrt{3}\,\alpha) + \sqrt{3}s\left(\frac{\pi}{4} + \frac{1}{2}\arctan\frac{s}{2\alpha}\right)\right].$$

Define

$$L(\alpha) = \frac{\pi}{4} + \frac{1}{2}\arctan\frac{s}{2\alpha}, \qquad B(\alpha) = \arctan(\sqrt{3}\alpha) + \sqrt{3}sL(\alpha).$$

A direct derivative calculation gives

$$B'(\alpha) = -\frac{\sqrt{3}\alpha}{s} L(\alpha).$$

Since $s'(\alpha) = -\alpha/s$,

$$\Gamma'(\alpha) = \frac{3\alpha}{\pi s(1+s)^2}\left[\arctan(\sqrt{3}\alpha) - \sqrt{3}L(\alpha)\right].$$

For $\alpha \in (0,1)$, $\arctan(\sqrt{3}\alpha) \le \pi/3$ and $L(\alpha) \ge \pi/4$. Since $\sqrt{3}\pi/4 > \pi/3$, we have $\Gamma'(\alpha) < 0$. The endpoint values, obtained by continuity, are $\Gamma(0) = 3\sqrt{3}/4$ and $\Gamma(1) = 1$.

The triple-pair value ratio is

$$\frac{V_3(T,\alpha)}{V_2(T,\alpha)} = \sqrt{\frac{T+2}{T+3}}\Gamma(\alpha).$$

Thus $V_3(T,\alpha) > V_2(T,\alpha)$ if and only if

$$\Gamma(\alpha) > \sqrt{\frac{T+3}{T+2}}.$$

Because

$$1 < \sqrt{\frac{T+3}{T+2}} < \sqrt{\frac{3}{2}} < \frac{3\sqrt{3}}{4},$$

continuity and strict monotonicity give a unique threshold $\alpha^\star_{23}(T) \in (0,1)$, with equality at the threshold.

For the pair-singleton comparison,

$$V_2(T,\alpha) > V_1(T,\alpha)$$

is equivalent to

$$1 + s > \sqrt{\frac{T+2}{T+1}}.$$

Solving for α gives

$$\alpha < \alpha^\star_{12}(T) = \sqrt{1 - \left(\sqrt{\frac{T+2}{T+1}} - 1\right)^2},$$

with equality at $\alpha = \alpha^\star_{12}(T)$.

It remains to show the pair interval is nonempty. Let

$$s_{12} = \sqrt{\frac{T+2}{T+1}} - 1.$$

Then $s_{12} \in (0, \sqrt{2}-1)$ and $\alpha^\star_{12}(T) = \sqrt{1-s_{12}^2}$. At this value, writing $s = s_{12}$,

$$\sqrt{\frac{T+3}{T+2}} = \frac{\sqrt{1+4s+2s^2}}{1+s}.$$

Define

$$D(s) = \sqrt{1+4s+2s^2} - \frac{3}{\pi} B\left(\sqrt{1-s^2}\right).$$

Since $D(0) = 0$, it is enough to show $D'(s) > 0$ on $[0, \sqrt{2}-1]$. Using $dB/ds = \sqrt{3}L$,

$$D'(s) = \frac{2(1+s)}{\sqrt{1+4s+2s^2}} - \frac{3\sqrt{3}}{\pi}\left(\frac{\pi}{4} + \frac{1}{2}\arctan\frac{s}{2\sqrt{1-s^2}}\right).$$

The first term is bounded below by $\sqrt{8/3}$ on this interval, and

$$\frac{s}{2\sqrt{1-s^2}} \le \frac{1}{4}.$$

Therefore

$$D'(s) \ge \sqrt{\frac{8}{3}} - \frac{3\sqrt{3}}{\pi}\left(\frac{\pi}{4} + \frac{1}{8}\right) > 0.$$

Hence $D(s) > 0$, which implies

$$\Gamma(\alpha^\star_{12}(T)) < \sqrt{\frac{T+3}{T+2}}.$$

Since Γ is strictly decreasing and

$$\Gamma(\alpha^\star_{23}(T)) = \sqrt{\frac{T+3}{T+2}},$$

we obtain $\alpha^\star_{23}(T) < \alpha^\star_{12}(T)$.

A.10 Proof of Proposition 5

The proof uses two auxiliary facts about maxima of regular-simplex Gaussian vectors. For $N \ge 2$, let

$$M_N = \max_{1 \le i \le N} S_i^N,$$

where S^N is centred Gaussian with unit variances and pairwise correlations $-1/(N-1)$. Let F_N and f_N denote the distribution and density of M_N on $(0, \infty)$. For $N \ge 2$, $M_N \ge 0$ almost surely and $F_N(0) = 0$.

Lemma 3 (Likelihood-ratio ordering of simplex maxima). *If $K > N \ge 2$, then M_K dominates M_N in the likelihood-ratio order.*

Proof. For $N = 2$, $M_2 = |Z|$, so $F_2(x) = 2\Phi(x) - 1$ and $f_2(x) = 2\varphi(x)$ for $x > 0$. For $N \geq 3$, exchangeability and conditioning on one coordinate give

$$f_N(x) = N\varphi(x)F_{N-1}(a_N x), \qquad a_N = \sqrt{\frac{N}{N-2}}, \qquad x > 0. \tag{3}$$

Indeed, conditional on $S_1^N = x$, the remaining $N-1$ coordinates have mean $-x/(N-1)$, variance $N(N-2)/(N-1)^2$, and pairwise correlation $-1/(N-2)$. Hence

$$\Pr(S_2^N \leq x, \ldots, S_N^N \leq x \mid S_1^N = x) = F_{N-1}\left(\sqrt{\frac{N}{N-2}}x\right),$$

which proves (3).

We first prove the following induction claim: for every $k \geq 2$ and every $0 < p < q$, the function

$$x \mapsto \frac{F_k(px)}{F_{k-1}(qx)}$$

is increasing on $(0, \infty)$, with the convention $F_1 \equiv 1$. The case $k = 2$ is immediate. Suppose the claim holds for $k - 1$. By the monotone l'Hôpital rule, since the numerator and denominator both vanish at zero, it is enough to show that

$$x \mapsto \frac{pf_k(px)}{qf_{k-1}(qx)}$$

is increasing. For $k = 3$, (3) gives this ratio as a positive constant times

$$\exp\left(\frac{q^2 - p^2}{2}x^2\right) F_2(a_3 px),$$

which is increasing. For $k \geq 4$, (3) gives

$$\frac{pf_k(px)}{qf_{k-1}(qx)} = C_k \exp\left(\frac{q^2 - p^2}{2}x^2\right) \frac{F_{k-1}(a_k px)}{F_{k-2}(a_{k-1}qx)},$$

where $C_k > 0$. Since $a_k < a_{k-1}$ and $p < q$, we have $a_k p < a_{k-1}q$, so the final ratio is increasing by the induction hypothesis. The exponential factor is also increasing. This proves the induction claim.

For $N = 2$,

$$\frac{f_3(x)}{f_2(x)} = \frac{3}{2}F_2(\sqrt{3}x),$$

which is increasing. For $N \geq 3$, (3) gives

$$\frac{f_{N+1}(x)}{f_N(x)} = \frac{N+1}{N}\frac{F_N(a_{N+1}x)}{F_{N-1}(a_N x)}.$$

Because $a_{N+1} < a_N$, the induction claim shows that this ratio is increasing. Therefore M_{N+1} dominates M_N in the likelihood-ratio order; transitivity gives the result for every $K > N$. $\square$

Lemma 4 (A low-budget bound)**.** *Let $\beta_0 = \sqrt{2} - 1$ and $c = 1/\sqrt{2\pi}$. For every $N \geq 3$,*

$$\frac{1}{\sqrt{N}} \mathbb{E}\left[\left(\sqrt{1-\beta_0^2}\, Z_0 + \beta_0 M_N\right)^+\right] < c,$$

where $Z_0 \sim \mathsf{N}(0,1)$ is independent of M_N.

Proof. Let $a_0 = \sqrt{1-\beta_0^2}$ and

$$\psi(x) = \mathbb{E}[(Z+x)^+] = \varphi(x) + x\Phi(x).$$

Then

$$h_N(\beta_0) \equiv \mathbb{E}\left[(a_0 Z_0 + \beta_0 M_N)^+\right] = a_0 \mathbb{E}\left[\psi\left(\frac{\beta_0}{a_0} M_N\right)\right].$$

Since $\psi(0) = c$, $\psi'(0) = 1/2$, and $\psi''(x) = \varphi(x) \leq c$, Taylor's theorem gives, for $x \geq 0$,

$$\psi(x) \leq c + \frac{x}{2} + \frac{cx^2}{2}.$$

Consequently,

$$h_N(\beta_0) \leq a_0 c + \frac{\beta_0}{2}\mu_N + \frac{c\beta_0^2}{2a_0}\eta_N, \tag{4}$$

where $\mu_N = \mathbb{E}[M_N]$ and $\eta_N = \mathbb{E}[M_N^2]$.

Represent the simplex vector as

$$S^N = \sqrt{\frac{N}{N-1}}(Z - \bar{Z}\mathbf{1}), \qquad \bar{Z} = \frac{1}{N}\sum_{i=1}^{N} Z_i,$$

with $Z_1, \ldots, Z_N$ independent standard normals. If $Z_{(N)} = \max_i Z_i$, then

$$\mu_N = \sqrt{\frac{N}{N-1}}\mathbb{E}[Z_{(N)}], \qquad \eta_N = \frac{N}{N-1}\mathbb{E}[Z_{(N)}^2] - \frac{1}{N-1}.$$

The second identity uses $\mathbb{E}[Z_{(N)}\bar{Z}] = 1/N$, which follows from Stein's lemma because $\partial Z_{(N)}/\partial Z_i = 1$ on the event that i is the unique maximiser, an event of probability $1/N$.

For $N = 3, 4, 5, 6$, the order-statistic moments are the one-dimensional integrals

$$\mathbb{E}[Z_{(N)}] = N\int_{-\infty}^{\infty} z\varphi(z)\Phi(z)^{N-1}\, dz, \qquad \mathbb{E}[Z_{(N)}^2] = N\int_{-\infty}^{\infty} z^2\varphi(z)\Phi(z)^{N-1}\, dz.$$

Deterministic interval integration of these four one-dimensional integrals gives the following upper bounds. The integration is over $[-8, 8]$ with interval enclosures for φ and Φ;

the omitted normal tails are bounded by Mills' ratio and contribute less than 10^{-12} to the displayed moments. Substituting the moment bounds into (4) gives:

N	$\mathbb{E}[Z_{(N)}]$	$\mathbb{E}[Z_{(N)}^2]$	$h_N(\beta_0)/\sqrt{N}$
3	≤ 0.84629	≤ 1.27567	≤ 0.365
4	≤ 1.02938	≤ 1.55133	≤ 0.338
5	≤ 1.16297	≤ 1.80003	≤ 0.317
6	≤ 1.26721	≤ 2.02174	≤ 0.300

All four upper bounds are strictly below $c \approx 0.399$.

For $N \ge 7$, use the Gaussian-maxima bounds

$$\mathbb{E}[Z_{(N)}] \le \sqrt{2\log N}, \qquad \mathbb{E}[Z_{(N)}^2] \le 2\log N + 1.$$

The first follows from the log-sum-exp bound and optimisation over the exponential tilting parameter. The second follows because $z \mapsto \max_i z_i$ is one-Lipschitz, so the Gaussian Poincaré inequality gives $\operatorname{Var}(Z_{(N)}) \le 1$. Together with (4), these imply

$$\frac{h_N(\beta_0)}{\sqrt{N}} \le B_N \equiv \frac{a_0 c + \frac{\beta_0}{2}\sqrt{\frac{N}{N-1}}\sqrt{2\log N} + \frac{c\beta_0^2}{2a_0}\left[\frac{N}{N-1}(2\log N + 1) - \frac{1}{N-1}\right]}{\sqrt{N}}.$$

The three components of B_N are decreasing for $N \ge 7$. For the last component, the derivative of its continuous extension has numerator

$$3x^2 - 2x - 1 - 2x(x+1)\log x,$$

which is negative for $x \ge 7$ because $\log x > 3/2$ implies $2x(x+1)\log x > 3x(x+1) > 3x^2 - 2x - 1$. Therefore $B_N \le B_7$ for all $N \ge 7$, and direct substitution gives $B_7 < 0.383 < c$. This proves the lemma. $\square$

We now prove Proposition 5. Write

$$\beta = \sqrt{1-\alpha^2}, \qquad \lambda = \frac{\beta}{\alpha}, \qquad g(y) = \mathbb{E}[(Z+y)^+] = \varphi(y) + y\Phi(y).$$

For $N \ge 2$ and $\alpha \in (0,1)$,

$$V_N^{\mathrm{simp}}(T,\alpha) = \sqrt{\frac{T}{T+N}}\,\alpha A_N(\lambda), \qquad A_N(\lambda) = \mathbb{E}[g(\lambda M_N)].$$

The kernel $(\lambda, m) \mapsto g(\lambda m)$ is log-supermodular on $(0,\infty)^2$, since

$$\frac{\partial^2}{\partial\lambda\,\partial m}\log g(\lambda m) = \frac{d}{dy}\left(\frac{y\Phi(y)}{\varphi(y)+y\Phi(y)}\right)\bigg|_{y=\lambda m} = \frac{\varphi(y)\left((1+y^2)\Phi(y) + y\varphi(y)\right)}{(\varphi(y)+y\Phi(y))^2} > 0.$$

By Lemma 3, M_K dominates M_N in the likelihood-ratio order whenever $K > N$. Logsupermodularity and likelihood-ratio order imply that $A_K(\lambda)/A_N(\lambda)$ is increasing in λ. To see this directly, for $\lambda_2 > \lambda_1$, the ratio $g(\lambda_2 m)/g(\lambda_1 m)$ is increasing in m, while the density ratio of M_K to M_N is increasing; symmetrising the resulting double integral gives

$$A_K(\lambda_2)A_N(\lambda_1) - A_K(\lambda_1)A_N(\lambda_2) \geq 0.$$

Hence, for $K > N \geq 2$,

$$\frac{V_K^{\text{simp}}(T,\alpha)}{V_N^{\text{simp}}(T,\alpha)} = \sqrt{\frac{T+N}{T+K}} \frac{A_K(\lambda)}{A_N(\lambda)}$$

is increasing in λ. Since $\lambda = \sqrt{1-\alpha^2}/\alpha$ is strictly decreasing in α, this ratio is decreasing in α.

It remains to include $N = 1$. We have

$$V_1(T) = \sqrt{\frac{T}{T+1}}c, \qquad V_2^{\text{simp}}(T,\alpha) = \sqrt{\frac{T}{T+2}}c(1+\beta).$$

Thus V_2^{simp}/V_1 is increasing in λ. For $K > 2$,

$$\frac{V_K^{\text{simp}}}{V_1} = \frac{V_K^{\text{simp}}}{V_2^{\text{simp}}} \frac{V_2^{\text{simp}}}{V_1},$$

so V_K^{simp}/V_1 is also increasing in λ, hence decreasing in α. Therefore, for every $K > N \geq 1$, the value ratio $V_K^{\text{simp}}(T,\alpha)/V_N^{\text{simp}}(T,\alpha)$ is decreasing in α. Pairwise single crossing on the finite ordered set $\{1, \ldots, |\mathcal{N}|\}$ implies that the smallest and largest elements of $\mathcal{A}_{\text{RS}}(T,\alpha)$ are weakly decreasing in α. This proves part 1; the endpoint cases $\alpha = 0, 1$ follow by continuity.

For part 2, define, for $N \geq 2$,

$$h_N(\beta) = \mathbb{E}\left[\left(\sqrt{1-\beta^2}\, Z_0 + \beta M_N\right)^+\right], \qquad h_1(\beta) = c.$$

The common factor $\sqrt{T}$ does not affect the argmax, so the relevant objective is

$$\frac{h_N(\beta)}{\sqrt{T+N}}.$$

We first consider large T. Let $\mu_N = \mathbb{E}[M_N]$ for $N \geq 2$, and set $\mu_1 = 0$. At $\beta = 0$, $h_N(0) = c$ for every N, and, for $N \geq 2$,

$$h_N'(0) = \mathbb{E}[M_N \mathbf{1}\{Z_0 > 0\}] = \frac{\mu_N}{2}.$$

Hence, uniformly over the finite menu and over b in any compact set,

$$h_N(b/T) = c + \frac{b\mu_N}{2T} + O(T^{-2}).$$

Since

$$\frac{1}{\sqrt{T+N}} = \frac{1}{\sqrt{T}}\left(1 - \frac{N}{2T} + O(T^{-2})\right),$$

we have

$$\frac{h_N(b/T)}{\sqrt{T+N}} = \frac{1}{\sqrt{T}}\left[c + \frac{1}{2T}(b\mu_N - cN) + O(T^{-2})\right].$$

The local ranking is therefore determined by

$$\Gamma_N(b) = b\mu_N - cN.$$

For $N = 2$, $M_2 = |Z|$, so $\mu_2 = 2c$ and $\Gamma_2(b) = 2cb - 2c$. Thus $N = 2$ beats $N = 1$ in the local ranking if and only if $b > 1/2$. For $N \geq 3$, $N = 2$ beats N whenever

$$b(\mu_N - 2c) < c(N-2).$$

Define

$$B_2 = \min_{\substack{3 \leq N \leq |\mathcal{N}| \\ \mu_N > 2c}} \frac{c(N-2)}{\mu_N - 2c},$$

with $B_2 = \infty$ if the set is empty. We claim that $B_2 > 1/2$. For any centred vector $y \in \mathbb{R}^N$,

$$\max_i y_i \leq \sqrt{\frac{N-1}{N}}\|y\|.$$

Using $M_N = \sqrt{N/(N-1)}(Z_{(N)} - \bar{Z})$, this gives

$$M_N \leq \|Z - \bar{Z}\mathbf{1}\|, \qquad \mu_N \leq \mathbb{E}\|Z - \bar{Z}\mathbf{1}\| \leq \sqrt{\mathbb{E}\|Z - \bar{Z}\mathbf{1}\|^2} = \sqrt{N-1}.$$

For $N \geq 3$, $\sqrt{N-1} < 2c(N-1)$. Hence, whenever $\mu_N > 2c$,

$$\frac{c(N-2)}{\mu_N - 2c} > \frac{1}{2}.$$

Thus $B_2 > 1/2$. Choose a compact interval $J \subset (1/2, B_2)$ with nonempty interior. Uniformity of the expansion over the finite menu implies that, for all sufficiently large T, $N = 2$ uniquely maximises the exact objective whenever $\beta = b/T$ with $b \in J$. Equivalently,

$$\alpha = \sqrt{1 - \frac{b^2}{T^2}}, \qquad b \in J.$$

The corresponding set I_T^∞ is a nonempty interval in $(0, 1)$, and both of its endpoints converge to one as $T \to \infty$.

Now consider small T. As $T \downarrow 0$, the objective $h_N(\beta)/\sqrt{T+N}$ converges uniformly over the finite menu and over compact β-intervals to $h_N(\beta)/\sqrt{N}$. Let

$$\beta_0 = \sqrt{2} - 1.$$

Then

$$\frac{h_1(\beta_0)}{\sqrt{1}} = c, \qquad \frac{h_2(\beta_0)}{\sqrt{2}} = \frac{c(1+\beta_0)}{\sqrt{2}} = c.$$

By Lemma 4, every $N \geq 3$ lies strictly below this value at β_0. Since the menu is finite, there is a uniform gap

$$\Delta = \min_{3 \leq N \leq |\mathcal{N}|} \left\{ \frac{h_2(\beta_0)}{\sqrt{2}} - \frac{h_N(\beta_0)}{\sqrt{N}} \right\} > 0.$$

By continuity, choose $\varepsilon \in (0, 1-\beta_0)$ such that for every $\beta \in [\beta_0, \beta_0 + \varepsilon]$ and every $N \geq 3$,

$$\frac{h_2(\beta)}{\sqrt{2}} - \frac{h_N(\beta)}{\sqrt{N}} \geq \frac{\Delta}{2}.$$

Uniform convergence then implies that, for all sufficiently small $T > 0$, $N = 2$ beats every $N \geq 3$ throughout this interval.

The comparison with the singleton is exact. Since $h_2(\beta) = c(1+\beta)$ and $h_1(\beta) = c$, $N = 2$ beats $N = 1$ if and only if

$$\frac{c(1+\beta)}{\sqrt{T+2}} > \frac{c}{\sqrt{T+1}},$$

or

$$\beta > \beta_{12}(T) \equiv \sqrt{\frac{T+2}{T+1}} - 1.$$

For every $T > 0$, $\beta_{12}(T) < \sqrt{2} - 1 = \beta_0$. Therefore, for all sufficiently small $T > 0$, $N = 2$ uniquely beats every other breadth whenever $\beta \in (\beta_0, \beta_0 + \varepsilon)$. Returning to α, define

$$I^0 = \left(\sqrt{1 - (\beta_0 + \varepsilon)^2}, \sqrt{1 - \beta_0^2} \right).$$

For every $\alpha \in I^0$ and every sufficiently small $T > 0$, $\mathcal{A}_{\mathrm{RS}}(T, \alpha) = \{2\}$. Finally,

$$\sqrt{1 - \beta_0^2} = \sqrt{1 - (\sqrt{2} - 1)^2} = \sqrt{2\sqrt{2} - 2} \approx 0.910,$$

so the low-budget interval lies near α_0. This completes the proof.

References

AGARWAL, R., F. BACCO, A. CAMUFFO, A. COALI, A. GAMBARDELLA, H. MSANGI, S. SONKA, A. TEMU, B. WAIZED, AND A. WORMALD (2025): "Does a Theory-of-Value Add Value? Evidence from a Randomized Control Trial with Tanzanian Entrepreneurs," *Organization Science*, 36, 601–625.

AGRAWAL, A., A. CAMUFFO, A. GAMBARDELLA, J. S. GANS, E. L. SCOTT, AND S. STERN (2026): *Bayesian Entrepreneurship*, Cambridge, MA: The MIT Press.

ARTS, S., B. CASSIMAN, AND J. HOU (2023): "Position and Differentiation of Firms in Technology Space," *Management Science*, 69, 7253–7265.

BERNARDO, A. E. AND I. WELCH (2001): "On the Evolution of Overconfidence and Entrepreneurs," *Journal of Economics & Management Strategy*, 10, 301–330.

BLANK, S. (2026): "Your Startup Is Probably Dead On Arrival," Blog post, accessed May 6, 2026.

BOUDREAU, K. J., N. LACETERA, AND K. R. LAKHANI (2011): "Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis," *Management Science*, 57, 843–863.

CAMUFFO, A., A. CORDOVA, A. GAMBARDELLA, AND C. SPINA (2020): "A Scientific Approach to Entrepreneurial Decision Making: Evidence from a Randomized Control Trial," *Management Science*, 66, 564–586.

CHATTOPADHYAY, S., F. HONORÉ, AND S. WON (2025): "Free Range Startups? Market Scope, Academic Founders, and the Role of General Knowledge in AI," *Strategic Management Journal*, 46, 1027–1079.

CSASZAR, F. A. AND D. A. LEVINTHAL (2016): "Mental Representation and the Discovery of New Strategies," *Strategic Management Journal*, 37, 2031–2049.

DENRELL, J., C. FANG, AND D. A. LEVINTHAL (2004): "From T-Mazes to Labyrinths: Learning from Model-Based Feedback," *Management Science*, 50, 1366–1378.

DENRELL, J. AND J. G. MARCH (2001): "Adaptation as Information Restriction: The Hot Stove Effect," *Organization Science*, 12, 523–538.

DUTT, N. AND M. LAWRENCE (2022): "Learning to Manage Breadth: Experience as Repetition and Adaptation," *Organization Science*, 33, 1300–1318.

FELIN, T., A. GAMBARDELLA, S. STERN, AND T. R. ZENGER (2020): "Lean Startup and the Business Model: Experimentation Revisited," *Long Range Planning*, 53, 101889.

GANS, J. S., E. L. SCOTT, AND S. STERN (2024): *Entrepreneurship: Choice and Strategy*, Norton.

GANS, J. S., S. STERN, AND J. WU (2019): "Foundations of Entrepreneurial Strategy," *Strategic Management Journal*, 40, 736–756.

GIUS, L. (2025): "Commercializing Contrarian Ideas: Evidence from AI Contests," Working paper.

GOLDREICH, D. AND H. HAŁABURDA (2013): "When Smaller Menus Are Better: Variability in Menu-Setting Ability," *Management Science*, 59, 2518–2535.

GOLDSMITH, A. (2023): "A Proof of the Simplex Mean Width Conjecture," ArXiv preprint.

GRUBER, M., I. C. MACMILLAN, AND J. D. THOMPSON (2008): "Look Before You Leap: Market Opportunity Identification in Emerging Technology Firms," *Management Science*, 54, 1652–1665.

GUZMAN, J. AND A. LI (2023): "Measuring Founding Strategy," *Management Science*, 69, 101–118.

HARRISON, D. A. AND K. J. KLEIN (2007): "What's the Difference? Diversity Constructs as Separation, Variety, or Disparity in Organizations," *Academy of Management Review*, 32, 1199–1228.

KABLUCHKO, Z., A. E. LITVAK, AND D. ZAPOROZHETS (2017): "Mean Width of Regular Polytopes and Expected Maxima of Correlated Gaussian Variables," *Journal of Mathematical Sciences*, 225, 770–787.

KLINGEBIEL, R. AND C. RAMMER (2014): "Resource Allocation Strategy for Innovation Portfolio Management," *Strategic Management Journal*, 35, 246–268.

KLUENDER, R., J. L. KRIEGER, AND M. B. WEISS (2021): "Biobot Analytics," Case 821-045, Harvard Business School, revised June 12, 2023.

KONING, R., S. HASAN, AND A. CHATTERJI (2022): "Experimentation and Start-up Performance: Evidence from A/B Testing," *Management Science*, 68, 6434–6453.

LEIPONEN, A. AND C. E. HELFAT (2010): "Innovation Objectives, Knowledge Sources, and the Benefits of Breadth," *Strategic Management Journal*, 31, 224–236.

LEVINTHAL, D. A. (2017): "Mendel in the C-Suite: Design and the Evolution of Strategies," *Strategy Science*, 2, 282–287.

LITVAK, A. E. (2018): "Around the Simplex Mean Width Conjecture," in *Analytic Aspects of Convexity*, ed. by G. Bianchi, A. Colesanti, and P. Gronchi, Cham: Springer, vol. 25 of *Springer INdAM Series*, 73–84.

LOCH, C. H., C. TERWIESCH, AND S. THOMKE (2001): "Parallel and Sequential Testing of Design Alternatives," *Management Science*, 47, 663–678.

MANSO, G. (2011): "Motivating Innovation," *The Journal of Finance*, 66, 1823–1860.

NELSON, R. R. (1961): "Uncertainty, Learning, and the Economics of Parallel Research and Development Efforts," *Review of Economics and Statistics*, 43, 351–364.

PRAHALAD, C. K. AND R. A. BETTIS (1986): "The Dominant Logic: A New Linkage Between Diversity and Performance," *Strategic Management Journal*, 7, 485–501.

RAO, H. (2009): *Market Rebels: How Activists Make or Break Radical Innovations*, Princeton, NJ: Princeton University Press.

SHERMON, A. AND M. MOEEN (2022): "Zooming In or Zooming Out: Entrants' Product Portfolios in the Nascent Drone Industry," *Strategic Management Journal*, 43, 2217–2252.

SOMMER, S. C. AND C. H. LOCH (2004): "Selectionism and Learning in Projects with Complexity and Unforeseeable Uncertainty," *Management Science*, 50, 1334–1347.

SRINIVASAN, R., A. S. CHOO, S. NARAYANAN, S. SARKAR, AND A. TENHIÄLÄ (2021): "Knowledge Sources, Innovation Objectives, and Their Impact on Innovation Performance: Quasi-Replication of Leiponen and Helfat (2010)," *Strategic Management Journal*, 42, 2104–2136.

SUN, W., Z.-C. HU, AND G. LAN (2020): "The 4-D Gaussian Random Vector Maximum Conjecture and the 3-D Simplex Mean Width Conjecture," ArXiv preprint.

SUN, Y. (1997): "Stochastic Iterative Algorithms for Signal Set Design for Gaussian Channels and Optimality of the L_2 Signal Set," *IEEE Transactions on Information Theory*, 43, 1574–1587.

THOMKE, S., E. VON HIPPEL, AND R. FRANKE (1998): "Modes of experimentation: an innovation process—and competitive—variable," *Research Policy*, 27, 315–332.

VAN DEN STEEN, E. (2010): "Culture Clash: The Costs and Benefits of Homogeneity," *Management Science*, 56, 1718–1738.