Jevons' Boon: Towards Modelling the Expanding Viability of Institutional Forms under Synthetic Cognitive Labor

By Tushita Jha, Rory Svarc, Ameya Prabhu, and Matthias Bethge*

Working Draft: 2026-03-27: 03:57:19 GMT

This paper provides a framework for thinking about cognitive scarcity as constraint on institutional viability, focusing on the economic impacts of large language models (LLMs). We interpret LLMs as generators of 'synthetic cognitive work', and develop a framework for reasoning about organization of cognitive production. We observe how increasing productivity of synthetic cognitive labor, via Jevons' paradox, can lead to a rebound in the demand for cognitive work, by turning previously cost-prohibitive cognitive goods and services into cost-viability. We call this Jevons' Boon: The marginal expansion (real or potential) of viable cognitive production.

This has important relationship to institution design and possibility, via a cognitive institutionalist lens: with decreasing cognitive scarcity comes an expansion of the production possibility frontier for diverse forms of institutional arrangements. Institutional arrangements and organization structures rely upon the provision of appropriate cognitive work to sustain them. Historically, binding constraints on institutional design the legibility traps analyzed by James C Scott, the transaction costs emphasized by Coase and Williamson, and the difficulty of collective action studied by Olson can be understood as manifestations of cognitive scarcity. Hence we argue that as AI systems become feasible sources of cognitive work, unprecedented forms of institutional arrangements can become viable. Concretely, we note expansions along qualitative dimensions: granularity (across contexts), plurality (of stakes), and fidelity (or stringency). Any institutional theory adequate to ongoing technological transitions must recognize how qualitative improvements along these three axes expand the viability of novel forms of fine-grained administration, cooperation across diverse differences, and availability of information commons.

And therefore the aim of this paper is to characterize Boon, by providing a formal model for the real and potential viability expansions, respectively.

JEL: O33, D02, D23, D83, H41, L22, L86

Keywords: cognitive production, cognitive scarcity, institutional design, cognitive institutionalism, jevons paradox, synthetic cognitive labor, artificial intelligence, large language models, bounded rationality

I. Introduction

This is a working draft for a paper aimed at answering the following question: **As the synthetic cognitive work that is extracted from AI systems expands, how do we characterize the corresponding expansion of cognitive production in: a) available realized cognitive goods and services (Real Jevons' Boon), and b) the space of viable aggregate cognitive productions (Potential Jevons' Boon).**

In this working draft, we briefly discuss the motivations, and connection to literature below; provide a elaboration of the formal setup in Section 2, with current results including shadow prices derivation for endogenous demand over unrealized cognitive goods/services; and then provide a concrete roadmap towards an analytic characterization of the viability expansions.

Reader guide.

This working draft is modular and may be read non-linearly. Section I introduces the motivating question and situates the project within the economics of AI and institutional analysis. Section II sets out the papers aims and the perspective of cognitive production systems. Section III introduces the basic conceptual objects, while Section IV provides the formal setup. Readers interested chiefly in the current theoretical payoff may wish to move directly to Section V, where the paper develops the layered account of Jevons Boon and presents the main result on demand cascades. Section VI clarifies which parts of the argument remain provisional by outlining the roadmap ahead, and Section VII offers informal interpretation of the framework, including its implications for institutional forms and possible applications.

A. Motivation

Motivation 1: LLMs (and Generative AI, more generally, can be understood as generators of synthetic cognitive work. [1] Consider Alice. Alice is a typical remote knowledge worker, who begins her workday by opening her laptop. Her interface with the world is, we'll suppose, remarkably simple—a computer window displaying text, images and videos. A request appears:"*Draft a market analysis for renewable energy in Southeast Asia.*" Alice reads the prompt, absorbing its context and requirements. She processes this against her background

* Jha (primary author): Mimir Center for Long Term Futures Research (particle.mania@gmail.com); Svarc: Arb Research; Prabhu and Bethge: Tbingen AI Centre, University of Tbingen.

[1] We are intentionally agnostic about questions of phenomenology or subjectivity. Our concern is functional utility, and for our purposes it matters not whether the Chinese Room thinks (Searle, 1980; Preston and Bishop, 2002; Cole, 2024). Instead, we focus on what becomes possible when scalable 'Chinese Room' style systems can be deployed as a service. Thus, the term 'synthetic cognitive work' carries no more metaphysical commitment than "horsepower implies the presence of horses.

knowledge, draws upon her professional training, refers to online resources, and begins typing her response, one line at a time. Throughout the day, various requests arrive through this interface: code to debug, emails to draft, documents to summarize, and strategic questions to answer. For each, Alice maintains focus on the specific task, produces appropriate outputs, and awaits the next prompt. This familiar scenario of remote cognitive work provides a surprisingly apt metaphor for understanding large language models (LLMs). Like Alice, an LLM interfaces with the world primarily through language (and digitized audio-visual) data. It receives a prompt, processes it against its pretrained knowledge, can access the Internet and digital tools, and generates a response one token at a time. Likewise, the LLM's cognitive work is metered in its own way (in this case, by the number of tokens consumed and generated rather than Alice who charges by the hour or by the project). What makes this metaphor particularly illuminating is not that LLMs perfectly replicate human cognition, but rather that they simulate the functional process of language-mediated cognitive work. As processual approximations of how cognitive workers like Alice produce meaningful and economically useful work, LLMs offer something Alice cannot: cheap *duplicability* and abundant *scalability*[2]

Motivation 2: *Institutions* can be characterized in terms of cognitive work required to sustain them. The cognitive scaffolding view (Denzau and North, 2000; Clark, 2006; Petracca and Gallagher, 2020) interprets institutions as enablers of systematic cognitive offloading. Analyzing cognitive production thus allows us to attend to how offloaded cognitive work is itself organized and maintained, treating the "scaffolding" not as given infrastructure but as a production system with its own internal structure. This also allows characterizing institutions by their communicative purposes: the set of expectations about cognitive outputs (deliberations, syntheses, coordinations) that an institution is selected for or designed to meet. A parliament's communicative purpose includes producing legitimate collective decisions through deliberative aggregation; a firm's includes sustaining coordinated production through monitoring, planning, and managerial interpretation. Furthermore, and most consequentially for our analysis, the space of viable institutions, especially in terms of the richness of communicative purposes that can be viably served, is constrained by the cost structure of the cognitive work required to construct and maintain them.

Motivation 2a: Cognitive scarcity can be seen as a major source of constraint on the space of viable institutional arrangements. Herbert Simon's work on bounded rationality emphasized how organizational forms emerge

[2]To understand this scalability, we must distinguish between two concepts often conflated in discussions of LLMs : a) The *Generator* of language-mediated work (i.e., the LLM itself). This is a statistical compression of processual approximations drawn from many workers like Alice—a static information artifact encoded in model weights that can be cheaply copied and deployed wherever compute is available. b) *LLM-generated Cognitive Work* (i.e., the tokens produced, and the work done through them). In contrast to the generator, LLM-generated cognitive work emerges dynamically, through the time-evolution of model inference in interaction with the digital environment. As the LLM generates tokens in response to specific prompt, it thereby simulates the processes of cognitive workers like Alice.

as adaptations to cognitive constraints (Simon, 1971, 1972, 1991, 1997). This lends an interpretative lens for ideas that reverberate throughout the history of institutional theory. James C. Scott's *'legibility'*, or how states simplify complex realities into administratively tractable categories (Scott, 1998; Coyne, Goodman and Quintas, 2025), can be understood in terms of standardization as an economy of abstraction when fine-grained comprehension is prohibitive. Friedrich Hayek's contributions to the economic calculation debate emphasize how market arrangements economize on cognitive work by dispersing it across the society (Hayek, 1945). Ronald Coase and Oliver Williamson's theory of the firm identifies transaction costs, many of them cognitive (negotiation, monitoring, enforcement), as determining the boundary between market and hierarchy (Coase, 1937, 1992; Williamson, 1975, 1985, 1995). The issue of social cost and the failure of prices to internalize all externalities (Coase, 1960; Stigler, 1966) in markets, as well as the limitations on the size and scope of coordination in collective action (Olson, 1965), can be understood similarly through the scarcity of relevant cognitive work. Several theories of firm explain them in terms of ability to utilize information (Grant, 1996) or cognitive work required to monitor economic activities (Alchian and Demsetz, 1972). Even institutional isomorphism (DiMaggio and Powell, 1983; Powell and DiMaggio, 1991), as well as path dependencies and lock-ins (Arthur, 1989, 1994), can be seen as responses to the cost of cognitive work required for adaptively customizable and revisable arrangements.

Motivation 3: The Characterization of the Expansion of Viable Institutional Arrangements must occupy an important role in Economics of AI. Prior discussions both in economics and an AI itself have often read AI as primarily as a story of substitution or displacement. While we believe that such framings are often important and intellectually productive, solely focusing on the degree to which AI system may automate existing human functions prematurely narrows the space of possible outcomes through focusing on current institutional arrangements rather than how institutional arrangements may endogeneously evolve consequent to the introduction of capable AI systems. In this sense, economics should focus on what AI systems newly enable in addition to what they displace (Brynjolfsson, 2022; Korinek and Brynjolfsson, 2022).

Although we provide a framework which (in principle) affords sensible interpretation sans the introduction of newly capable AI systems, we believe that large language models (LLMs) make this analytic perspective particularly urgent. As progress in AI systems shifts the cost structure of cognitive work, the space of viable cognitive productions expands we term *Jevons' Boon* becomes more salient. Among the many institutional arrangements that may be included in this newly expanded viable set include institutions allowing for finer-grained administration with cheaper customization and context-sensitivity (Foster, 2023), pluralistic and broader collaboration with falling barriers to continuous deliberation and preference aggregation (Weyl, Tang and the Plurality Community, 2023), and a richer information commons with falling costs of faithful translation and

synthesis across heterogeneous sources (Heylighen, 2011). Still, we note that the expanded space of viable institutional arrangements is while it may be of interest to normative theorists itself shorn of prescriptive claims. In what follows we aim to understand the contours of this potentially expanded frontier, and ask which institutional configurations become viable along which dimensions, and at what costs.

II. Methods and Aims.

Our paper hence views LLMs as Cognitive Production Systems, and considers what becomes economically viable when the cost structure of cognitive work undergoes transformation. Accordingly, we claim that as a previously scarce factor cognitive work becomes more abundant, a binding economic constraint loosens, and consequently new institutional arrangements become viable. We call the consequence of this dynamic *Jevon's Boon*: a newly expanded viability set of institutional arrangements.

In this paper we develop a theoretical perspective centered on *cognitive production systems (CPS)* – organized workflows that consume cognitive labor and representational stocks to produce cognitive outputs (decisions, analyses, coordinations, etc.). Our analysis focuses on the demand structure that CPS serve what we call *communicative purpose* and on three dimensions determining the quality of that service: *Granularity* (context-sensitivity), *Plurality* (stakeholder reconciliation), and *Fidelity* (informational integration). In broad terms we introduce a model which decomposes signals that Cognitive Production Systems receive from the world, and treat LLMs as policies which map directly perceived inputs from the world (via information channels) to distributions over actions. From this we discuss a mechanism whereby newly entrant cognitive production systems induces a marginal demand for unrealized cognitive good/services in the economy which are below the level of cost-viability. As the marginal cost of cognitive labor falls (due to the increasing capabilities of AI), goods or services which were previously below the level of cost-viability become cost-viable, and their realization expands the quality of the aggregate cognitive output (ie. decisions) of the overall system, institution, or economy.

- We aim to understand how the institutional viability envelope expands once the intermediate supply–demand interactions generated by cognitive production are treated endogenously. The central question is not only which existing tasks become cheaper, but how reductions in the cost of cognitive labor induce new demand for previously unrealized cognitive goods and services, thereby enlarging the set of institutionally viable arrangements.
- We aim to provide a framework in which many AI-enabled cognitive workflows are understood less as "agents" in the strong sense of persistent, self-directed entities, and more as production organizations oriented toward demand-satisfaction. On this view, the relevant economic question concerns

how such systems organize cognitive tasks and how their effective boundaries shift with the relative costs of coordination, contracting, monitoring, and externalization.

- We aim to use the general framework of cognitive production systems to think about institutions more broadly as arrangements for the production of cognitive outputs. This perspective may help identify the relevant margins of institutional choice—including where activities are centralized or decentralized, which tradeoffs are technologically or organizationally salient, and which elasticities matter as the cost of cognitive labor changes. It also provides a way to ask what synthetic infrastructure would be required for alternative institutional forms to replicate or improve upon existing functions.

- We aim to develop a public-economics perspective on the allocation of synthetic cognitive labor. If synthetic cognitive capacity becomes an economically important input, then its deployment raises familiar questions about efficiency, access, externalities, and social value. The issue is not only how much synthetic cognitive labor is produced, but how it is allocated across uses when private incentives and social returns may diverge.

III. Setup: Modeling the Economics of Organization of Cognitive Work

We will focus our discussion on 'cognitive work'. This term admits some imprecision, but we can begin by listing some examples: the production of codebases, industry reports, policy analyses, strategic recommendations, or textbooks. We can think of cognitive work as the use of perceptual inputs and (previously) accumulated representations in the production of decisions and further representations. Our use of the term 'cognitive work' is somewhat imperfect; for instance, we do not mean to suggest that forms of work such as plumbing, construction, or retail services somehow involve 'less human cognition'. Nonetheless, we will henceforth use the term 'cognitive work' to distinguish between work which centrally involves the production of representations from those that don't.[3]

We will say that the primary factor input to cognitive work is **cognitive labor**: the capacity to perform cognitive work, whether supplied by human reasoning or extracted from AI systems. A **cognitive production system** (CPS), generally speaking, is any organized workflow that transforms cognitive labor into cognitive outputs in order to meet some demand. Thus, the *process* by which a CPS produces output is 'cognitive work', its *input* its output can be understood as any output which can (in principle, at least) serve as representational capital to be used in the future production of cognitive outputs.

[3]Philosophical footnote on representations

A. Supplying Cognitive Work

IV. Formal Model

This section develops the formal apparatus. Section IV.A introduces the primitive spaces. Section IV.B provides a running illustration. Section IV.C defines the three quality channels. Section IV.D formalizes communicative purpose and ensurance. Sections IV.E–IV.F construct the economy of cognitive productions and the planner's problem. Section V presents the layered account of Jevons' Boon. Section V.A states and proves the main result: the demand cascade under cross-channel complementarity. Section VII.B interprets institutional archetypes as channel specializations. Section VI outlines the project roadmap.

A. Primitive Spaces and Causal Structure

The World State and Its Functional Projections

Let W denote the world state, a random element over a measurable space $(\Omega, \mathcal{W})$. Two projections are generated from W via conditionally independent Markov channels,

$$W \to X, \qquad W \to Y,$$

so that $p(X, Y \mid W) = p(X \mid W)\, p(Y \mid W)$. Each projection serves a distinct functional role.

- $(X, \mathcal{X})$ is the **inferential scope**: the ecology of representational commitments—sources, frameworks, logical dependencies, evidential relations—to which cognitive outputs bear referential obligation. X is the ground over which epistemic competence resolves uncertainty.
- $(Y, \mathcal{Y})$ is the **operating context**: the space of demand contingencies across which cognitive outputs must be produced. Y admits a product structure $Y = \prod_{i \in I} Y_i$ over contextual features indexed by a finite set I. The product structure is imposed on the state space; the components may exhibit arbitrary stochastic dependence under the joint measure.

A probability measure μ on $(Y, \mathcal{Y})$ governs the frequency distribution of demand contingencies. The decomposition is functional, not ontological: it separates the roles that world-features play in cognitive production—epistemically and operationally—without requiring a partition of the state space itself.

Outputs, Effects, and the Consequence Structure

Let $(\mathcal{O}, \Sigma_{\mathcal{O}})$ denote the **output space** of cognitive productions. The mapping from output-in-context to downstream effects is governed by a collection of Markov kernels:

$$\lambda = \big(\lambda_i : \mathcal{O} \times Y_i \to \Delta(E_i)\big)_{i \in I}.$$

Each λ_i determines $p(E_i \mid O, Y_i)$: the i-th effect depends on the output and the i-th contextual feature only. The effect profile $E = \langle E_i \rangle_{i \in I}$ satisfies:

$$p(E \mid O, Y) = \prod_{i \in I} \lambda_i(E_i \mid O, Y_i).$$

All covariation among effects, conditional on context, is induced through the single common cause O.

Standing technical assumptions.

Throughout, we assume that $\mathcal{O}$ is compact, that the kernels $\mathcal{P}$ and λ are continuous (in the weak topology on $\Delta(\cdot)$), and that $\mathrm{supp}(\pi_K)$ is compact. These conditions ensure that the maximin value introduced in Section IV.D is well-defined and upper semicontinuous in the channel parameters.

B. Running Illustration: Municipal Housing Policy Advisory

We introduce a two-tier example to ground the formal objects. The downstream production—a municipal housing policy advisory—consumes the output of an upstream production—a legal compliance review—as an intermediate cognitive good.

The Downstream Production

Consider a cognitive production system that generates housing policy recommendations for a municipal government overseeing n neighborhoods.

The *operating context* is $Y = (Y_1, \ldots, Y_n)$, where each Y_i encodes the conditions in neighborhood i: housing stock composition, vacancy rates, demographic profile, zoning status, construction market conditions, prevailing rents, and environmental constraints.

The *inferential scope* X is the structured body of knowledge that adequate recommendations must cohere with: urban economic theory (rent dynamics, filtering models, supply elasticity estimates), housing law (state preemption statutes, fair housing requirements, zoning enabling acts), environmental regulation, and demographic analysis. The inferential scope is not the state of any particular city; it is the ecology of frameworks, sources, and logical dependencies to which outputs bear referential obligation.

The *output space* $\mathcal{O}$ consists of policy briefs specifying, for each neighborhood or class of neighborhoods, a recommended regulatory action (upzoning, rent stabilization, inclusionary requirements, density bonuses, etc.) together with supporting analysis.

The *consequence structure* λ encodes, for each neighborhood i, how the recommendation propagates to effects $E_i = (E_{i,1}, \ldots, E_{i,m})$—including changes in rents, new construction activity, displacement of existing renters, fiscal impact,

environmental consequences, and changes in property values—given local conditions Y_i.

An Intermediate Cognitive Good: The Legal Compliance Review

The housing advisory does not operate in isolation. Among the cognitive productions it may consume is an upstream *legal compliance review*: a specialized analysis of the state-level statutory landscape identifying binding legal constraints on municipal housing policy.

The upstream producer examines state preemption statutes, enabling acts, fair housing obligations, and recent judicial interpretations. It produces a compliance memorandum mapping the space of possible municipal regulatory actions onto their legal permissibility.

The downstream advisory receives this memorandum. Prior to receiving it, the advisory's epistemic state K regarding the legal subspace of X may be impoverished—knowing that state preemption exists as a general concept but lacking the resolution to determine which specific actions would survive legal challenge. After integrating the memorandum, the advisory's effective epistemic channel $\mathcal{K}$ is enriched: it can form sharper posteriors over the legal aspects of X, ruling out impermissible recommendations and exploiting legal safe harbors.

Let $\mathcal{K}_{\text{pre}}$ denote the advisory's epistemic channel before consuming the review and $\mathcal{K}_{\text{post}}$ the enriched channel afterward. The legal review's contribution is a Blackwell improvement:

$$\mathcal{K}_{\text{post}} \succeq_B \mathcal{K}_{\text{pre}}.$$

By Proposition IV.D, the advisory's ensurable class expands: $\text{Ens}_\delta(\mathcal{K}_{\text{post}}, \mathcal{P}, \lambda) \supseteq \text{Ens}_\delta(\mathcal{K}_{\text{pre}}, \mathcal{P}, \lambda)$. Communicative purposes requiring legal coherence that the prior epistemic state could not verify now enter the ensurable class.

The enrichment is not free. The advisory incurs *translation friction* $f_{\mathcal{K}}$: the cognitive labor of interpreting the compliance memorandum—bridging the vocabulary and analytical conventions of statutory interpretation and urban economics, and extracting actionable content from legal terms-of-art whose policy implications are not self-evident. The friction cost $f_{\mathcal{K}}$ is the *interface cost* borne by the downstream consumer: the price of converting an upstream output into a channel enrichment.

Other Intermediate Goods in the Network

Two further intermediate goods complete the picture. A *neighborhood conditions survey* (customization adapter, enriching $\mathcal{P}$) conducts site-specific assessments, enabling the advisory to distinguish between neighborhoods that a coarse-grained data source would treat identically. The associated friction cost $f_{\mathcal{P}}$ covers interpreting heterogeneous site-level data and converting it into the advisory's analytical categories. A *displacement impact assessment* (evaluative mediator,

enriching λ) models how specific policy actions affect vulnerable tenant populations. Its output brings displacement effects within the advisory's anticipatory horizon. The friction cost f_λ covers integrating the impact model's predictions into the advisory's consequence structure.

The Production Network

The housing advisory sits at the center of a production network in which each upstream production enriches a specific channel:

$$\tau_{\text{legal}} \xrightarrow{\text{enriches } \mathcal{K}} \tau_{\text{advisory}} \xleftarrow{\text{enriches } \mathcal{P}} \tau_{\text{survey}}, \qquad \tau_{\text{impact}} \xrightarrow{\text{enriches } \lambda} \tau_{\text{advisory}} \to \text{policy recommendation.}$$

The advisory's aggregate quality depends on the combined channel profile $(\mathcal{K}_{\text{post}}, \mathcal{P}_{\text{post}}, \lambda_{\text{post}})$ achieved through these enrichments. Whether each upstream production is actually realized depends on whether its shadow value exceeds its cost at prevailing cognitive labor prices—the viability question formalized in Sections IV.F–V.

C. The Three Quality Channels

A cognitive production technology transforms operating contexts into outputs. Its quality is characterized by three structural channels—two of the standard Blackwell type, one of the dual Blackwell type—each governing a distinct dimension.

Channel (a): Epistemic Channel—Governing Fidelity

Let K denote the production's epistemic state, satisfying the sufficiency property $\pi_k(\cdot) \equiv p(X \mid K = k)$. The epistemic state yields well-defined posteriors over the inferential scope X, characterizing the production's informational position regarding the representational ecology.

Blackwell ordering. $\mathcal{K}' \succeq_B \mathcal{K}$ iff $\mathcal{K}$ is a garbling of $\mathcal{K}'$: there exists a Markov kernel G such that $p(K \mid X) = \int G(K \mid K')\, p(K' \mid X)\, dK'$.

Quality dimension: Fidelity. The Blackwell rank of $\mathcal{K}$ governs the depth of inferential integration achievable in output. Translational losses—misrepresented sources, inconsistently applied frameworks, degraded logical coherence—arise from information about X lost in the epistemic channel.

In the illustration, a production with rich $\mathcal{K}$ identifies that a proposed rent stabilization measure conflicts with a state preemption statute, or that a displacement risk model is inapplicable to neighborhood 4's demographic composition. A production with impoverished $\mathcal{K}$ may recommend a policy that violates state law or rely on elasticity estimates from inapplicable markets.

Channel (b): Perceptual Channel—Governing Granularity

The operating context Y is not observed directly. A perceptual channel generates observables $Y \to P$ via $p(P \mid Y)$, with $W \to Y \to P$ ensuring $P \perp W \mid Y$.

Blackwell ordering. $\mathcal{P}' \succeq_B \mathcal{P}$ iff $\mathcal{P}$ is a garbling of $\mathcal{P}'$ conditional on Y.

Quality dimension: Granularity. The Blackwell rank of $\mathcal{P}$ governs the fineness of contextual discrimination. Standardization loss—the gap between context-bespoke and categorical treatment—arises from information about Y lost in the perceptual channel. This is the formal structure underlying Scott's **?** analysis of state legibility: administrative categories compress heterogeneous local conditions into coarse types, and the standardized policy response is the inevitable consequence.

In the illustration, a production with rich $\mathcal{P}$ distinguishes that neighborhood 3 has a tight rental market with displacement pressure while neighborhood 7 has abundant vacant lots. A production with coarse $\mathcal{P}$ perceives the city as an undifferentiated unit and issues a uniform citywide policy.

Channel (c): Consequence Channel—Governing Plurality

The production chooses output O via a K-parameterized policy $\sigma_k : P \to \Delta(\mathcal{O})$. The consequence structure λ maps output-in-context to effects.

Dual-Blackwell characterization. Fix $Y_i = y_i$. The forward kernel $\lambda_i(\cdot \mid o, y_i)$ maps outputs to effect distributions. Read in reverse, $\lambda_i(e_i \mid o, y_i)$ defines a likelihood for o given observation e_i, yielding a statistical experiment $\Lambda_i^{y_i} = (E_i, \{\lambda_i(\cdot \mid o, y_i)\}_{o \in \mathcal{O}})$ in which o is the unknown state and e_i the signal.

Quality dimension: Plurality via anticipation-accountability. The dual-Blackwell rank of λ governs the **anticipation-accountability scope**: the degree to which the consequence structure renders effects across the full profile $(E_i)_{i \in I}$ *anticipatable* (the production can foresee how outputs propagate) and *attributable* (effects can be traced to the output that caused them). An externality is formally a consequence kernel λ_j absent from the production's consequence model, or sufficiently noisy that the reverse experiment Λ_j is uninformative about O. The effect on party j exists but falls outside the production's anticipatory horizon.

Ordering. $\lambda' \succeq_C \lambda$ iff the reverse experiment Λ' is Blackwell-more-informative about O than Λ.

In the illustration, a production with rich λ anticipates that upzoning in neighborhood 3 will raise tax revenue but displace existing renters and strain infrastructure. It can couple the upzoning with a relocation fund and impact fee. A production with impoverished λ recommends the upzoning on supply-side grounds alone; displacement and infrastructure strain materialize as externalities.

The Quality Profile

DEFINITION 1 (Quality Profile): *The quality profile of a cognitive production technology $\tau = (\mathcal{K}, \mathcal{P}, \lambda, \sigma)$ is the channel triple $(\mathcal{K}, \mathcal{P}, \lambda)$, where $\mathcal{K}$ and $\mathcal{P}$ are*

ordered by standard Blackwell dominance and λ by dual Blackwell dominance. The composite partial order is $\tau' \succeq \tau$ iff $\mathcal{K}' \succeq_B \mathcal{K}$, $\mathcal{P}' \succeq_B \mathcal{P}$, and $\lambda' \succeq_C \lambda$.

The asymmetry is structural: fidelity and granularity are *observation problems* (the production cannot honor commitments it does not access, nor adapt to contexts it does not perceive), while plurality is an *anticipation-accountability problem* (the production cannot serve evaluative criteria whose associated effects fall outside the scope of foreseeability and attributability). The asymmetry reflects the economic fact that observing the world and anticipating one's effects upon it are structurally distinct activities.

Transmission Between Producer and Consumer

The producer holds epistemic state K and observes percept P. It selects output $O = \sigma_K(P)$. The consumer at position i receives O in the context of local conditions Y_i, experiencing effect $E_i \sim \lambda_i(\cdot \mid O, Y_i)$. The quality of this transmission decomposes along the three channels.

Fidelity of transmission. Output O bears referential obligations to the inferential scope X. Whether O fulfills these depends on the resolution of $\mathcal{K}$: the producer can only honor commitments in X that its epistemic state distinguishes.

Granularity of transmission. Output O should adapt to Y. The degree of adaptation depends on the resolution of $\mathcal{P}$: the producer can only tailor O to aspects of Y that the percept P transmits. For any pair of contexts y, y' requiring distinct outputs, the production distinguishes them only if $\mathcal{P}(y) \neq \mathcal{P}(y')$ with sufficient probability.

Plurality of transmission. Output O induces effects across the full profile $(E_i)_{i \in I}$. Whether these effects are jointly adequate depends on the anticipation-accountability scope of λ: the producer can design O to satisfy joint constraints on the effect profile only if the consequence structure renders those effects foreseeable.

D. Communicative Purpose and Ensurable Properties

Communicative Purpose

DEFINITION 2 (Communicative Purpose): *A communicative purpose is a measurable set*

$$\Phi \in \mathcal{Y} \otimes \Sigma_{\mathcal{O}} \otimes \mathcal{X} \otimes \mathcal{E},$$

where $\mathcal{E} = \bigotimes_{i \in I} \mathcal{E}_i$ is the product σ-algebra over the effect space. The tuple $(y, o, x, e) \in \Phi$ is interpreted: "output o is adequate for context y, faithful to inferential commitment x, and satisfactory across effect profile e."

There is no separate evaluative ground: the evaluation *is* the demand specification. The three aspects of communicative purpose are structural features of Φ: the *situation coverage* $\mathrm{proj}_{\mathcal{Y}}(\Phi)$ specifies across which contingencies adequacy must

hold; the *inferential constraints* (the $\mathcal{X}$-fibers of Φ) specify which representational commitments the output must honor; and the *evaluative constraints* (the joint structure of Φ over $\mathcal{E} = \prod_i \mathcal{E}_i$) specify which cross-effect configurations count as adequate.

In the illustration, the communicative purpose Φ^* for the housing advisory specifies: for each neighborhood y_i, the recommendation o must be adapted to local conditions (granularity fiber), cohere with applicable legal and economic frameworks (inferential fiber), and produce an effect profile across rents, construction, displacement, and fiscal impact satisfying joint adequacy criteria encoding the municipality's obligations to all affected parties (evaluative fiber).

Separability and Entanglement

DEFINITION 3 (Channel-Separable Purpose): *A communicative purpose Φ is **channel-separable** if it admits a decomposition $\Phi = \Phi_{\mathcal{K}} \cap \Phi_{\mathcal{P}} \cap \Phi_{\lambda}$, where $\Phi_{\mathcal{K}}$ constrains only the (o, x) margin, $\Phi_{\mathcal{P}}$ constrains only the (y, o) margin, and Φ_{λ} constrains only the (o, e) margin.*

DEFINITION 4 (Channel-Entangled Purpose): *A communicative purpose is **channel-entangled** if no such decomposition exists.*

In the illustration, the housing advisory's purpose is channel-entangled. The requirement "the recommendation must not produce displacement effects that violate fair housing obligations" jointly constrains adaptation to local demographics ($\mathcal{P}$-dependent), coherence with fair housing law ($\mathcal{K}$-dependent), and anticipation of displacement consequences (λ-dependent). Adequacy cannot be checked channel by channel.

Ensurance

DEFINITION 5 (Ensurance—Robust Formulation): *A production technology τ **ensures** communicative purpose Φ at level $(1-\delta)$ if there exists a policy $\sigma \in \Sigma(\tau)$ such that*

$$\forall\, p \in \mathrm{supp}(\mathcal{P}),\ \ \forall\, x \in \mathrm{supp}(\pi_K) : \qquad \Pr_{(Y,E)|P=p}\big((Y,\, \sigma_K(p),\, x,\, E) \in \Phi\big)\ \geq\ 1 - \delta.$$

The universal quantifiers enforce robustness: the production must perform adequately for every *percept it may receive and* every *inferential commitment consistent with its epistemic state. The residual tolerance δ absorbs exactly two sources of irreducible stochasticity: residual uncertainty over Y given $P = p$ (since $\mathcal{P}$ is generically a noisy channel), and consequence stochasticity in $E \sim \lambda(\cdot \mid O, Y)$.*

DEFINITION 6 (Ensurable Class): *The ensurable class of τ at level $(1-\delta)$ is*

$$\mathrm{Ens}_{\delta}(\tau) = \big\{\Phi : \tau \text{ ensures } \Phi \text{ at level } (1-\delta)\big\}.$$

[Channel Monotonicity] Channel enrichment weakly expands the ensurable class:

$$\mathcal{K}' \succeq_B \mathcal{K} \implies \mathrm{Ens}_\delta(\mathcal{K}', \mathcal{P}, \lambda) \supseteq \mathrm{Ens}_\delta(\mathcal{K}, \mathcal{P}, \lambda),$$

and analogously for $\mathcal{P}$ (under $\succeq_B$) and λ (under $\succeq_C$), each holding the other two channels fixed.

PROOF:

A richer channel admits all policies available under the coarser channel: any $\mathcal{K}$-measurable policy is *a fortiori* $\mathcal{K}'$-measurable for $\mathcal{K}' \succeq_B \mathcal{K}$. The achievable set of joint distributions weakly expands. Any property ensured under the coarser channel remains ensurable under the richer one, since the witnessing policy remains admissible and the ensurance probability is preserved.

E. The Economy of Cognitive Productions

Production Configurations and Composition

A **production configuration** $\mathcal{T} = \{\tau_1, \ldots, \tau_m\}$ is a finite collection of realized cognitive productions organized in a directed acyclic network. Productions compose: the output of an upstream production enriches a specific channel of a downstream production. Three types of inter-production contribution correspond to the three channels: *customization adapters* enrich $\mathcal{P}$ downstream, *representational endowments* enrich $\mathcal{K}$ downstream, and *evaluative mediators* enrich λ downstream. Terminal nodes are end-point decisions consumed directly.

Aggregate Quality Functional

The network, considered as a composite production, induces an aggregate ensurable class $\mathrm{Ens}_\delta^{\mathrm{agg}}(\mathcal{T})$.

DEFINITION 7 (Aggregate Quality): *The aggregate quality of a production configuration $\mathcal{T}$ serving target communicative purpose Φ^* is:*

$$V(\mathcal{T} \mid \Phi^*) = \sup_{\sigma \in \Sigma(\mathcal{T})} \inf_{\substack{p \in \mathrm{supp}(\mathcal{P}) \\ x \in \mathrm{supp}(\pi_K)}} \Pr_{(Y,E)|P=p} \big((Y,\, \sigma_K(p),\, x,\, E) \in \Phi^*\big).$$

This is a maximin formulation: the planner selects the best policy, evaluated at the worst-case percept and inferential commitment. Under the standing compactness assumptions (Section IV.A), the infimum is attained and V is upper semicontinuous in the channel parameters. Enriching any channel of any node weakly improves V, since a richer channel weakly expands the policy space (raising the supremum) and can only tighten—never loosen—the worst case (the infimum is taken over a superset or is evaluated on a finer-grained probability).

Cost Structure with Friction Decomposition

Each production technology τ has a cost in cognitive labor units decomposing as:

$$c(\tau) = c_0(\tau) + f_{\mathcal{K}}(\tau) + f_{\mathcal{P}}(\tau) + f_{\lambda}(\tau), \tag{1}$$

where $c_0(\tau)$ is the *base production cost* (cognitive labor for the core transformation given full access to inputs), $f_{\mathcal{K}}(\tau)$ is the *translation friction* (search, extraction, and synthesis of upstream representational resources), $f_{\mathcal{P}}(\tau)$ is the *adaptation friction* (situational assessment, case analysis, and output customization), and $f_{\lambda}(\tau)$ is the *coordination friction* (impact assessment, deliberation across evaluative criteria, and attributability monitoring).

In the illustration: c_0 is the labor of writing the policy brief given all inputs; $f_{\mathcal{K}}$ includes interpreting the legal compliance review; $f_{\mathcal{P}}$ covers analyzing neighborhood-level survey data; f_{λ} covers integrating the displacement impact model.

The friction costs are currently exogenous parameters. Their endogenization—as functions of interpretive distances in the production network—is a principal target of the ongoing project (Section VI).

Monotonicity of cost. Channel enrichment is weakly costly: $\tau' \succeq \tau \implies c(\tau') \geq c(\tau)$.

F. The Planner's Problem and the Demand Functional

Formulation

A social planner chooses a production configuration to maximize aggregate quality subject to a cognitive labor budget. At effective price w per unit of cognitive labor, the effective budget is $B(w) = L/w$ for resource base L:

$$(\mathcal{P}\text{-}1) \qquad \max_{\mathcal{T} \subseteq T^{\text{pot}}} V(\mathcal{T} \mid \Phi^*) \qquad \text{s.t.} \qquad \sum_{\tau \in \mathcal{T}} c(\tau) \leq B(w).$$

Denote the optimal configuration by $\mathcal{T}^*(w)$ and the optimal value by $V^*(w)$. The shadow price of cognitive labor in the economy is:

$$\eta^*(w) = \left. \frac{\partial V^*}{\partial B} \right|_{B=B(w)},$$

measuring the marginal improvement in aggregate quality per additional unit of cognitive labor.

Shadow Value and Reservation Price

DEFINITION 8 (Shadow Value): *For any unrealized production $\tau \notin \mathcal{T}^*(w)$:*

$$\pi(\tau \mid \mathcal{T}^*, w) \;=\; V\big(\mathcal{T}^* \cup \{\tau\} \mid \Phi^*\big) - V(\mathcal{T}^* \mid \Phi^*).$$

DEFINITION 9 (Reservation Price): *The reservation price of an unrealized production τ is:*

$$p(\tau \mid \mathcal{T}^*, w) \;=\; \frac{\pi(\tau \mid \mathcal{T}^*, w)}{c(\tau)} \;=\; \frac{\pi(\tau)}{c_0(\tau) + f_{\mathcal{K}}(\tau) + f_{\mathcal{P}}(\tau) + f_{\lambda}(\tau)}.$$

The reservation price is the maximum effective price of cognitive labor at which the economy would find it worthwhile to bring τ into existence. At prevailing price w, the production is viable iff $p(\tau) \geq w$. At the planner's optimum, $p(\tau) = \eta^*$ for marginal productions, $p(\tau) > \eta^*$ for infra-marginal (realized) productions, and $p(\tau) < \eta^*$ for excluded productions.

The reservation price is *decreasing* in each friction cost: $\partial p / \partial f_c = -\pi(\tau)/c(\tau)^2 < 0$ for $c \in \{\mathcal{K}, \mathcal{P}, \lambda\}$. A production with high shadow value but high friction may be excluded—not because the economy does not value the enrichment, but because the cognitive labor required to overcome the interface costs exceeds what the enrichment is worth at prevailing prices.

DEFINITION 10 (Demand Functional): *The demand functional assigns to each unrealized production its reservation price:*

$$D(\cdot \mid \mathcal{T}^*, w) : T^{\text{pot}} \setminus \mathcal{T}^* \to R_+, \qquad D(\tau) = p(\tau \mid \mathcal{T}^*, w).$$

DEFINITION 11 (Viable and Latent Demand Sets):

$$\mathcal{V}(w) = \big\{\tau \notin \mathcal{T}^* : D(\tau) \geq w\big\}, \qquad \mathcal{L}(w) = \big\{\tau \notin \mathcal{T}^* : 0 < D(\tau) < w\big\}.$$

The latent demand set $\mathcal{L}(w)$ is the economy's unmet demand: productions with positive shadow value whose reservation price falls below the prevailing cost of cognitive labor.

V. Jevons' Boon: A Layered Account

We distinguish four layers of the Boon, in order of increasing theoretical depth and decreasing current formalization.

Layer 0: The Unit-Cost Jevons Effect (Standard)

As the productive efficiency of synthetic cognitive labor improves, the effective price w falls. If aggregate demand for cognitive labor is price-elastic ($|\varepsilon| > 1$), total expenditure on cognitive labor increases despite falling per-unit cost. This

is the standard Jevons mechanism applied to cognitive labor as a factor input; it requires no special structure beyond elastic demand.

Layer 1: The Real Jevons' Boon (Typed Extensive Margin)

[Monotone Expansion of the Viable Set] For $w' < w$: $\mathcal{V}(w') \supseteq \mathcal{V}(w)$.
PROOF:

Immediate from the definition of the viable set.

As w falls, productions cross from the latent demand set into viability. The expansion is typed—each newly viable production enriches a specific channel: customization adapters enrich $\mathcal{P}$ (granularity margin), representational endowments enrich $\mathcal{K}$ (fidelity margin), and evaluative mediators expand λ (plurality margin). The substantive content lies not in the set inclusion (which is definitional) but in the typed composition with the channel structure: the economy does not merely produce "more cognitive work" but extends its capacity along specific quality dimensions.

Layer 2: The Demand Cascade (Cross-Channel Amplification)

The demand functional D depends on the current configuration $\mathcal{T}^*$. When a newly viable production is realized and the configuration changes, D shifts. Under cross-channel complementarity, this shift *raises* the reservation prices of complementary productions, pulling them into viability through a cascade that amplifies the initial cost reduction. Section V.A develops this mechanism formally.

Layer 3: The Virtual Jevons' Boon (Qualitative Institutional Expansion)

As $B(w)$ grows, the set of feasible configurations $C(B) = \{\mathcal{T} \subseteq T^{\text{pot}} : \sum c(\tau) \leq B\}$ expands. The *Virtual Boon* is the entry into feasibility of configurations exhibiting balanced channel richness—institutional arrangements that are simultaneously context-sensitive, inferentially rigorous, and deliberatively inclusive. This is stronger than Layer 1: not merely that new productions become individually viable, but that the *mode of organization* shifts from specialized (corner) to balanced (interior).

Converting the Virtual Boon to a theorem requires structural assumptions on V and c—specifically, supermodularity of V in aggregate channel qualities and a binding budget constraint at high w—under which the comparative static follows from Topkis **?** and Milgrom–Roberts **?**. This is a principal target of the ongoing project (Section VI).

Layer 3+: The Cognitive Accumulation Cycle

Cognitive productions may enrich the channels of the productions that produce other productions—the cognitive infrastructure itself. A representational endowment deepening the epistemic channel of a planning system also improves the

planning system's capacity to design future endowments. This recursive structure suggests the possibility of self-reinforcing cognitive accumulation, whose conditions for convergence, divergence, or bounded growth require an explicitly dynamic model and are entirely open.

A. Main Result: Demand Cascade Under Cross-Channel Complementarity

The Economic Question

Return to the housing advisory. Suppose cognitive labor costs fall enough that the legal compliance review τ_{legal} (a fidelity enrichment) becomes viable and is realized. The advisory now has a richer epistemic channel.

Does this change what the economy is willing to pay for the neighborhood survey τ_{survey} (a granularity enrichment) or the displacement assessment τ_{impact} (a plurality enrichment)?

The answer is yes, if quality channels are complementary. Once the advisory can ensure legal coherence ($\mathcal{K}$ enriched), the marginal value of also being able to tailor recommendations to local conditions ($\mathcal{P}$) increases—because a legally coherent *and* context-specific recommendation is worth more than the sum of its parts. A recommendation that correctly identifies the legal safe harbor *and* adapts to neighborhood 3's displacement risk is substantially more valuable than one that does either alone. The advisory can *exploit* the legal knowledge only if it can also *perceive* the local conditions to which the legal safe harbors are relevant.

This is the demand cascade: realizing one channel enrichment endogenously raises the economy's willingness to pay for complementary enrichments, potentially pulling them into viability without any further decrease in w.

Formalizing Complementarity

DEFINITION 12 (Cross-Channel Complementarity): *The aggregate quality functional V exhibits **cross-channel complementarity** at configuration $\mathcal{T}$ and node n if, for any two productions τ_a, τ_b enriching distinct channels of n:*

$$\pi(\tau_a \mid \mathcal{T} \cup \{\tau_b\}) \;\geq\; \pi(\tau_a \mid \mathcal{T}).$$

Equivalently, in the standard supermodularity form:

$$V(\mathcal{T} \cup \{\tau_a, \tau_b\}) - V(\mathcal{T} \cup \{\tau_b\}) \;\geq\; V(\mathcal{T} \cup \{\tau_a\}) - V(\mathcal{T}).$$

Cross-channel complementarity holds generically when the communicative purpose Φ^* is channel-entangled (Definition 4): if adequacy jointly constrains outputs across multiple channels, then enriching one channel in isolation yields limited improvement (the entangled constraints still bind through the other channels), while enriching multiple channels jointly releases the entangled constraint.

Channel-separable purposes, by contrast, yield additive shadow values and no complementarity.

In the illustration, the advisory's purpose—"produce legally sound, locally adapted recommendations that account for displacement effects"—is channel-entangled. Enriching only $\mathcal{K}$ (legal compliance) yields modest improvement because the entangled constraint still binds through $\mathcal{P}$ and λ. But once $\mathcal{K}$ is enriched, the marginal value of enriching $\mathcal{P}$ rises: a legally coherent advisory that can also perceive local conditions can exploit legal safe harbors *specifically appropriate* to each neighborhood, yielding quality improvements exceeding the sum of the independent enrichments.

The Cascade Proposition

[Demand Cascade] Let V exhibit strict cross-channel complementarity at $\mathcal{T}^*$ and node n. Let $\tau_1, \tau_2 \in \mathcal{L}(w)$ enrich distinct channels of n. Suppose a reduction in effective price from w to $w' = w - \epsilon$ makes τ_1 viable:

$$D(\tau_1 \mid \mathcal{T}^*) < w, \qquad D(\tau_1 \mid \mathcal{T}^*) \geq w'.$$

Then:

(i) Realizing τ_1 strictly raises the reservation price of τ_2:

$$D(\tau_2 \mid \mathcal{T}^* \cup \{\tau_1\}) \;>\; D(\tau_2 \mid \mathcal{T}^*).$$

(ii) If the complementarity-induced increase exceeds the remaining viability gap of τ_2,

$$\underbrace{D(\tau_2 \mid \mathcal{T}^* \cup \{\tau_1\}) - D(\tau_2 \mid \mathcal{T}^*)}_{\text{complementarity boost}} \;\geq\; \underbrace{w' - D(\tau_2 \mid \mathcal{T}^*)}_{\text{viability gap at } w'},$$

then τ_2 becomes viable at w' under the updated configuration: $D(\tau_2 \mid \mathcal{T}^* \cup \{\tau_1\}) \geq w'$.

PROOF:

Part (i). By strict cross-channel complementarity, $\pi(\tau_2 \mid \mathcal{T}^* \cup \{\tau_1\}) > \pi(\tau_2 \mid \mathcal{T}^*)$. Since $c(\tau_2) > 0$, the reservation price $D(\tau_2) = \pi(\tau_2)/c(\tau_2)$ is strictly higher under the updated configuration.

Part (ii). Rearranging the stated condition yields $D(\tau_2 \mid \mathcal{T}^* \cup \{\tau_1\}) \geq w'$, which is the viability condition at w'.

In the illustration: at price w, neither the legal review, the neighborhood survey, nor the displacement assessment is viable. Cognitive labor costs fall to w', making the legal review viable. With $\mathcal{K}$ enriched, the advisory can use neighborhood-level information more effectively—legal safe harbors can be matched to local conditions. The shadow value of the survey rises, and if the complementarity boost exceeds the viability gap, the survey enters the viable set. With both

$\mathcal{K}$ and $\mathcal{P}$ enriched, the marginal value of the displacement assessment may rise further, triggering a second round of the cascade.

The Cascade Fixed Point

[Cascade Fixed Point] Define the cascade operator on subsets of T^{pot}:

$$\Gamma_w(\mathcal{S}) = \left\{\tau \in T^{\text{pot}} \setminus \mathcal{T}^* : D(\tau \mid \mathcal{T}^* \cup \mathcal{S}) \geq w\right\}.$$

Under cross-channel complementarity and with T^{pot} finite, the iteration $\mathcal{S}_0 = \emptyset$, $\mathcal{S}_{k+1} = \Gamma_w(\mathcal{S}_k)$ converges in finitely many steps to a fixed point $\mathcal{V}^*(w) = \Gamma_w(\mathcal{V}^*(w))$, with $\mathcal{V}^*(w) \supseteq \mathcal{V}(w)$.

PROOF:

Γ_w is monotone on the lattice of subsets of T^{pot} under set inclusion: adding productions to $\mathcal{S}$ weakly increases shadow values (by complementarity) and hence weakly expands the set of productions meeting the viability threshold. With T^{pot} finite, the monotone sequence converges in at most $|T^{\text{pot}}|$ steps.

The **cascade-adjusted viable set** $\mathcal{V}^*(w)$ includes all productions whose viability is sustained either directly (reservation price exceeds w) or indirectly (reservation price raised above w by complementarity with other viable productions). The inclusion $\mathcal{V}^*(w) \supseteq \mathcal{V}(w)$ may be strict.

Summary of the Economic Content

Proposition V.A and Corollary V.A establish that cost reductions in cognitive labor have *amplified* effects on the viable set when cross-channel complementarities are present. A marginal decrease in w triggers a direct expansion (Layer 1: productions whose individual reservation prices cross the new threshold) and a cascade expansion (Layer 2: productions pulled into viability by complementarity-induced increases in their reservation prices, propagating until a fixed point is reached). The total expansion of the viable set—and consequently of the aggregate ensurable class—exceeds what a production-by-production accounting of viability thresholds would predict.

This mechanism gives formal content to the claim that the Jevons' Boon operates *in kind*: the realization of a fidelity enrichment endogenously creates demand for granularity and plurality enrichments, because the channels are jointly productive in serving entangled communicative purposes. The economy's demand for cognitive goods and services is not merely scaled up but structurally transformed by complementarity.

VI. Roadmap Ahead

A. Endogenization of Friction Costs (Primary Target)

The friction costs $(f_{\mathcal{K}}, f_{\mathcal{P}}, f_{\lambda})$ are currently exogenous parameters. Their endogenization—expressing each as a function of the *interpretive distance* between the production

and the upstream/downstream structures it must interface with—is the primary remaining theoretical task. This would close the model: the demand functional D would be fully determined by the economy's configuration, the communicative purposes it serves, and structural distances in the production network, with w entering only as the price of cognitive labor.

Specific AI capabilities map to specific friction reductions: large language models reduce $f_{\mathcal{K}}$ (cheaper search, translation, and synthesis of textual sources); multimodal and contextual AI reduces $f_{\mathcal{P}}$ (cheaper situational observation and interpretation); agent-based and simulation-capable AI reduces f_{λ} (cheaper impact modeling and consequence forecasting). The pattern of capability progress thus determines which margin of the Boon expands first.

B. Virtual Boon Theorem (Secondary Target)

Formalizing the conditions under which the optimal configuration shifts from specialized (corner) to balanced (interior) as $B(w)$ grows. The approach applies Topkis's monotone comparative statics **?** to the planner's problem ($\mathcal{P}$-1), treating B as the parameter. Sufficient conditions: supermodularity of V in aggregate channel qualities and a binding budget constraint at high w.

C. Cost Structure Microfoundation (Tertiary Target)

Derivation of the cost function $c(\mathcal{K}, \mathcal{P}, \lambda)$ from primitive assumptions about cognitive labor allocation. The conjectured sign pattern—$\partial_i c > 0$ (enrichment is costly), $\partial_{ij} c < 0$ for $i \neq j$ (pairwise economies of scope), $\partial_{ijk} c > 0$ (the pairwise economies erode as the third channel is also enriched)—would generate a richer viability landscape than either pure supermodularity or pure submodularity of cost.

D. Dynamic Extension (Long-Term)

Extension to an explicitly dynamic framework capturing the cognitive accumulation cycle (Layer 3+, Section V): cognitive productions that enrich the channels of the production system itself, generating self-reinforcing cognitive investment. Conditions for convergence, divergence, or bounded growth are entirely open.

VII. Informal Commentary

A. Cognitive Production in the Economy

Cognitive work is systematically organized rather than performed in isolation. We analyze this organization through the lens of factor inputs, productive assets, and the environment within which production occurs. The primary factor input is **cognitive labor**: the capacity to perform cognitive work, whether supplied by human reasoning or extracted from AI systems. In formal treatments,

cognitive labor can be modeled as a budget allocated across available sources, each characterized by cost and capability profile. Cognitive labor operates upon **representational capital**: accumulated stocks of documents, models, schemas, and codified knowledge that augment labor productivity. Representational capital bears the standard features of productive capital – it results from past investment, enhances current productivity, and persists across production cycles – while exhibiting the distinctive property that it may appreciate through refinement and connection rather than depreciate through use.

Cognitive production may be performed internally or outsourced to markets for cognitive work, to commons of accumulated representations, or to other specialized production systems. This environmental structure shapes what resources are available and at what cost.

DEFINITION 13 (Granularity): *The fineness of discrimination across the situation space – how much situational heterogeneity is preserved versus compressed in the mapping from situations to outputs.*

High granularity obtains when the CPS preserves situational heterogeneity: distinct situations receive distinctly appropriate responses, and decision-relevant contextual information is retained. Low granularity obtains when diverse situations are mapped to coarse categories, each receiving standardized treatment that erases within-category variation. The asymptotic ideal is pointwise context-specificity: every relevantly distinct situation receives a distinctly appropriate response, with no decision-relevant information compressed away. Failure to achieve this ideal constitutes *standardization loss*: the gap between what context-bespoke responses would achieve and what categorical responses actually deliver. Achieving high granularity requires cognitive labor for discrimination (recognizing relevant distinctions), customization (tailoring responses appropriately), and consistency (maintaining coherent treatment across fine partitions).

DEFINITION 14 (Plurality): *The extent of reconciliation across the evaluative profile – how much of the profile's complexity is genuinely incorporated rather than collapsed to simpler objectives or externalized.*

High plurality obtains when the CPS genuinely engages the evaluative profile's complexity: diverse criteria are surfaced, commensurated where possible, deliberated where necessary, and the production explores deeper into positive-sum frontiers. Low plurality obtains when the CPS collapses to simpler objectives, leaving some evaluative dimensions unaddressed or treating them as outside the decision frame. The asymptotic ideal is Pareto welfare-optimality: no criterion can be better served without thereby worsening performance on another. Achieving high plurality requires cognitive labor for elicitation (surfacing what different parties actually value), commensuration (rendering diverse criteria comparable where possible), deliberative synthesis (navigating genuine trade-offs where criteria conflict), and creative negotiation (disclosing otherwise obscure positive-sum

opportunities). This can be understood as the cognitive dimension of transaction costs and coordination costs (Coase, 1937, 1960, 1992; Williamson, 1975, 1979, 1985, 1995): the overhead of serving plural constituencies, reframed as cognitive labor required for evaluative reconciliation.

DEFINITION 15 (Fidelity): *The degree of coherence with the inferential scope – how deeply are logical, semantic and evidential commitments – inferential commitements qua representations – faithfully honored.*

High fidelity obtains when the CPS honors both external and internal commitments: outputs accurately reflect referenced sources, correctly implement presupposed frameworks, maintain the logical rigor their scope demands, and *synoptically* cohere with established knowledge. Low fidelity obtains when inferential commitments are nominally invoked but imperfectly honored – sources misrepresented, frameworks inconsistently applied, logical standards relaxed in practice. The asymptotic ideal is universal coherence: full integration with the representational ecology such that outputs would survive rational scrutiny from any position within it. Failure to achieve this ideal constitutes *translational losses* and cost barriers for *search and retrieval*. Achieving high fidelity requires cognitive labor for search (locating relevant representations), extraction (deriving usable content), translation (bridging interpretive distances across vocabularies and frameworks), and verification (ensuring internal coherence and external correspondence).

These three dimensions shape how the quality of cognitive work and its' output is assessed: how situationally appropriate it is across varying contexts; how adequate it is towards the downstream needs; and how faithfully integrated is it with distal expertise.

B. Institutional Archetypes as Channel Specializations

The three channels define a space of possible production profiles. Three canonical institutional forms occupy distinct positions—not by essential nature, but as rational responses to binding cognitive labor budgets.

Markets: High Granularity, Bounded Fidelity, Slack Plurality

Markets achieve fine-grained situational discrimination through the price mechanism; the perceptual channel $\mathcal{P}$ is highly resolved. Market outcomes cohere with the inferential scope only along efficiency criteria—prices reflect revealed preferences and opportunity costs, but distributional equity, ecological sustainability, and procedural fairness are structurally excluded from what the market's epistemic channel resolves. Fidelity is *bounded*: the price mechanism constitutes a specific and limited form of inferential integration. The consequence channel λ has narrow anticipation-accountability scope: effects mediated by price are anticipatable and attributable, while effects on non-transacting parties fall outside the market's anticipatory horizon. An externality is formally a consequence kernel

λ_j absent from the market's consequence model. Plurality is *slack*: the market could in principle be supplemented with mechanisms that bring non-priced effects within scope (Pigouvian taxes, cap-and-trade), but under cognitive scarcity, the labor cost of designing and maintaining such mechanisms is prohibitive.

Administration: High Fidelity, Bounded Plurality, Slack Granularity

Administrative institutions achieve deep inferential integration; the epistemic channel $\mathcal{K}$ is richly resolved, with outputs cohering with legal precedent, technical standards, procedural requirements, and accumulated expertise. Administrative decisions account for effects insofar as existing monitoring and compliance mechanisms can detect and attribute them, but effects that are diffuse, delayed, or observable only through non-standard channels remain outside the anticipatory horizon. Plurality is *bounded*: the administrative apparatus imposes a structural ceiling on which effects enter the decision frame. Administrative categories compress diverse local conditions into tractable types; the perceptual channel $\mathcal{P}$ is coarse—not because fine discrimination is structurally precluded, but because the cognitive labor cost of case-by-case contextual assessment is prohibitive under scarcity. Granularity is *slack*: this is Scott's **?** legibility, standardization as an economy of abstraction.

Electoral/Aggregation: High Plurality, Bounded Granularity, Slack Fidelity

Electoral and aggregation mechanisms achieve broad anticipation-accountability scope; the consequence channel λ brings effects on a wide range of constituencies within the anticipatory horizon of collective decision-making. Votes aggregate preferences over bundles; election cycles are temporally sparse; the perceptual channel maps the continuous space of citizen circumstances to a small number of options. Granularity is *bounded*: the aggregation requirement precludes fine-grained contextual discrimination by design. Aggregation procedures can produce outputs violating logical compositionality—the doctrinal paradox, in which individually coherent judgments aggregate to an incoherent collective judgment, is the canonical demonstration. Fidelity is *slack*: the inferential quality of collective outputs degrades as a consequence of aggregation, and the cognitive labor of restoring coherence is under-invested.

Institutional Specialization as Scarcity Response

These corner configurations reflect optimal channel investment under scarcity. When the budget permits heavy investment in at most one channel per institutional function, rational design concentrates resources at the institutional form's comparative advantage. The Jevons' Boon thesis reinterprets institutional history through this lens: the current institutional landscape reflects channel specialization under historically binding cognitive labor constraints. As synthetic cognitive

labor drives w downward, balanced configurations—context-sensitive *and* inferentially rigorous *and* deliberatively inclusive—shift from the latent demand set into the viable set.

C. Expanding the Frontier of Institutional Futures Enabled by Jevons' Boon

Modern institutions trade richness for manageability. They compress messy local knowledge into uniform categories so large systems can function at all. However, we argue that this compromise is not fundamental; it reflects the scarcity of cognitive labour.

[TODO: Add a deeper survey literature]

Fine-Grained Administration: Bespoke Provision Across Contexts

We interpret Scott's analysis of state legibility (Scott, 1998; Coyne, Goodman and Quintas, 2025) as one instance of a broader phenomenon: a practical tension between 'scalability' and 'contextual fit'. While (for example) standardized curricula and corporate policies, and medical protocols serve important roles in maintaining legibility, each suffers from potential ill-fit to novel contexts. Students with idiosyncratic or otherwise non-traditional learning needs may well learn best from more heterodox curricula, nor can corporate policies or medical protocols *ex ante* anticipate every local condition. These are failures inherent to the use of coarse categories, necessitated when the costs of fine-grained cognitive judgment are too expensive. With increasing abundance of cognitive labor comes the possibility of ameliorating some of these insalubrious dynamics. In short, decreased cognitive scarcity allows for provision which can increasingly be tailored without sacrificing reach. While there are clear benefits to be gained here in the form of better administration, we further believe that a loosening of the tradeoff may help mitigate dynamics of hierarchical provision which follow from this tradeoff.

Pluralisitic Collaboration: Cooperation Across (Increasingly Diverse) Differences

Intimate deliberation that involves surfacing diverse stakes, negotiating genuine disagreements, and discovering mutually beneficial arrangements has historically been feasible only at small scales. To coordinate larger groups, institutions thin out what is communicated: elections, for instance, collapse multidimensional preferences into binary choices, and prices can ignore hidden externalities. Weyl, Tang and the Plurality Community (2023) articulates a related vision for pushing this frontier outward, where deeper collaboration across differences manages to sustain diversity rather than collapsing it into coarsely aggregated summaries. When the cognitive work of surfacing perspectives, clustering arguments, identifying shared ground, and articulating remaining disagreements becomes cheap, the tradeoff loosens. In turn, engagement can be broader (more voices heard) as well as deeper (more dimensions of preference registered, and more opportunities

discovered for arrangements that serve multiple parties without forcing premature convergence).

Richer Information Commons: Making World's Information Increasingly Accessible and Useful

The present world contains an abundance of information. However, any aspiration to makes the world's information both useful and broadly accessible has been bottlenecked by the availability of cognitive labor. Converting the world's store of information into useful and accessible form [4] requires familiar forms of cognitive labor; information must be retrieved, interpreted, translated across different vocabularies, and then integrated into understandable forms that serve specific purposes. Cognitive scarcity thus fetters the usefulness of this work; both search and synthesis remain shallow, and in turn institutional memory becomes static or at least insufficiently adaptive. With decreased costs to cognitive labor institutions benefit alongside improvements in the *efficiency* of retrieval from an increased ability to integrate sources of information with sacrificing *fidelity*. Appropriate markers of provenance, uncertainty, and contextual dependency can be preserved across heterogeneous sources and integrated into useful forms.

[4] For example, the mission statement of Google – a major internet services firm – presents an idealized goal *"to organize the world's information and make it universally accessible and useful"*.

REFERENCES

Alchian, Armen A., and Harold Demsetz. 1972. "Production, Information Costs, and Economic Organization." *American Economic Review*, 62(5): 777–795.

Arthur, W. Brian. 1989. "Competing Technologies, Increasing Returns, and Lock-In by Historical Events." *The Economic Journal*, 99(394): 116–131.

Arthur, W. Brian. 1994. *Increasing Returns and Path Dependence in the Economy.* Ann Arbor:University of Michigan Press.

Brynjolfsson, Erik. 2022. "The Turing Trap: The Promise & Peril of Human-Like Artificial Intelligence." *Daedalus*, 151(2): 272–287.

Clark, Andy. 2006. "Language, Embodiment, and the Cognitive Niche." *Trends in Cognitive Sciences*, 10(8): 370–374.

Coase, Ronald H. 1937. "The Nature of the Firm." *Economica*, 4(16): 386–405.

Coase, Ronald H. 1960. "The Problem of Social Cost." *Journal of Law and Economics*, 3: 1–44.

Coase, Ronald H. 1992. "The Institutional Structure of Production." *The American Economic Review*, 82(4): 713–719. 1991 Alfred Nobel Memorial Prize Lecture in Economic Sciences.

Cole, David. 2024. "The Chinese Room Argument." In *The Stanford Encyclopedia of Philosophy.* . Winter 2024 ed., , ed. Edward N. Zalta and Uri Nodelman. Metaphysics Research Lab, Stanford University.

Coyne, Christopher J., Nathan Goodman, and André Quintas. 2025. "Ways of Seeing the World: Legibility in Alternative Institutional Settings." *European Economic Review*, 178: 105116.

Denzau, Arthur T, and Douglass C North. 2000. "Shared mental models: Ideologies and institutions." *Elements of reason: Cognition, choice, and the bounds of rationality*, 23–46.

DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review*, 48(2): 147–160.

Foster, Jacob G. 2023. "From Thin to Thick: Toward a Politics of Human-Compatible AI." *Public Culture*, 35(3): 417–430.

Grant, Robert M. 1996. "Toward a Knowledge-Based Theory of the Firm." *Strategic Management Journal*, 17(S2): 109–122.

Hayek, Friedrich A. 1945. "The Use of Knowledge in Society." *The American Economic Review*, 35(4): 519–530.

Heylighen, Francis. 2011. "Conceptions of a Global Brain: an historical review." In *Evolution: Cosmic, Biological, and Social.* , ed. Leonid E. Grinin, Robert L. Carneiro, Andrey V. Korotayev and Fred Spier, 274–289. Uchitel Publishing.

Korinek, Anton, and Erik Brynjolfsson. 2022. "The Turing Trap: A Conversation with Erik Brynjolfsson on the Promise and Peril of Human-Like AI." The Brookings Institution Webinar transcript. Webinar held November 2, 2022.

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups.* Vol. 124 of *Harvard Economic Studies*, Cambridge, MA:Harvard University Press.

Petracca, Enrico, and Shaun Gallagher. 2020. "Economic Cognitive Institutions." *Journal of Institutional Economics*, 16(6): 747–765.

Powell, Walter W., and Paul J. DiMaggio, ed. 1991. *The New Institutionalism in Organizational Analysis.* Chicago:University of Chicago Press.

Preston, John, and Mark Bishop, ed. 2002. *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence.* Oxford:Oxford University Press.

Scott, James C. 1998. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed.* New Haven, CT:Yale University Press.

Searle, John R. 1980. "Minds, Brains, and Programs." *Behavioral and Brain Sciences*, 3(3): 417–424.

Simon, Herbert A. 1971. "Designing Organizations for an Information-Rich World." In *Computers, Communications, and the Public Interest.* , ed. Martin Greenberger, 37–72. Baltimore:The Johns Hopkins Press.

Simon, Herbert A. 1972. "Theories of Bounded Rationality." In *Decision and Organization.* , ed. C. B. McGuire and Roy Radner, 161–176. Amsterdam:North-Holland Publishing Company.

Simon, Herbert A. 1991. "Organizations and Markets." *Journal of Economic Perspectives*, 5(2): 25–44. Spring.

Simon, Herbert A. 1997. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations.* . 4 ed., New York:Free Press.

Stigler, George J. 1966. *The Theory of Price.* . 3 ed., New York:Macmillan.

Weyl, E. Glen, Audrey Tang, and the Plurality Community. 2023. "Chapter 5: Collaborative Technology and Democracy." In *Plurality: The Future of Collaborative Technology and Democracy.* GitHub / pluralitybook (online edition). Accessed [date].

Williamson, Oliver E. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications: A Study in the Economics of Internal Organization.* New York:Free Press.

Williamson, Oliver E. 1979. "Transaction-Cost Economics: The Governance of Contractual Relations." *Journal of Law and Economics*, 22(2): 233–261.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting.* New York:Free Press.

Williamson, Oliver E., ed. 1995. *Organization Theory: From Chester Barnard to the Present and Beyond.* New York, NY:Oxford University Press.