

# The CES Potential: Information Friction and Complementary Production

Jon Smirl\*

February 2026

## Abstract

This paper derives the CES potential  $\mathcal{F}_q = \Phi_{\text{CES}}(\rho) - T \cdot S_q$  with  $q = \rho$ , connecting the CES aggregate—the unique constant-elasticity aggregator—to the Tsallis entropy—the unique entropy satisfying continuity, maximality, and pseudo-additivity. The companion paper ([Smirl, 2026a](#)) establishes that a single curvature parameter  $K = (1 - \rho)(J - 1)/J$  simultaneously controls superadditivity, correlation robustness, strategic independence, and network scaling. The present paper shows how information friction degrades this curvature. The *effective curvature theorem* establishes that exploitable curvature is  $K_{\text{eff}} = K \cdot (1 - T/T^*(\rho))^+$ , where  $T^*$  is a critical information friction increasing in complementarity. The three roles degrade non-uniformly—correlation robustness quadratically, superadditivity and strategic independence linearly—predicting a specific crisis sequence: diversification fails before production complementarities, which fail before strategic stability. Six areas of economic theory emerge as specific instances of the CES potential: adverse selection (Akerlof), mechanism design (Myerson), social choice (Arrow), search and matching (DMP), contract theory (Hart-Moore/Williamson), and behavioral economics (Kahneman-Tversky). Each derivation yields new  $\rho$ -dependent predictions that the original theories cannot generate. The Tsallis generalization with  $q = \rho$  replaces the Shannon chain rule with pseudo-additivity, producing  $q$ -exponential equilibrium distributions with compact support for complements and power-law tails for substitutes, a  $1/(2 - q)$  correction to the variance-response identity, and a classification of all dynamical results into exact survivors,  $q$ -corrected, and structurally changed. Optimal firm scope and the integration boundary are derived in  $(\rho, T)$  space, nesting Williamson’s governance structures as the qualitative limit. Two empirical tests confirm the framework: a banking regulation test using 147 countries predicts 2007–2009 crisis severity via the  $\sigma \times T$  interaction ( $p = 0.016$ ), and a procurement test using 1,172 federal contracts confirms that differentiated goods exhibit higher information rents ( $p = 0.003$ ).

---

\*Email: [jonsmirl@gmail.com](mailto:jonsmirl@gmail.com).

**JEL Classification:** B41, C46, C60, D21, D23, D80, D82, D83, D90, L22, L23

**Keywords:** CES aggregation, Tsallis entropy, CES potential, information friction, mechanism design, behavioral economics, theory of the firm, vertical integration

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	Summary of Results . . . . .	5
1.2	Organization . . . . .	6
1.3	Related Literature . . . . .	6
<b>2</b>	<b>Preliminaries: The CES Framework</b>	<b>8</b>
2.1	CES Production Technology . . . . .	8
2.2	The Quadruple Role of $\rho$ . . . . .	9
2.3	Micro-Foundations: Why CES Is Forced . . . . .	10
2.4	Nested CES and Multi-Level Aggregation . . . . .	10
<b>3</b>	<b>The CES Potential and Tsallis Entropy</b>	<b>11</b>
3.1	From Shannon to Tsallis: Why Complementarity Breaks Extensivity . . . . .	11
3.2	The Information Friction . . . . .	12
3.3	Construction of the CES Potential . . . . .	13
3.4	The $q$ -Exponential and $q$ -Logarithm . . . . .	14
3.5	The $q$ -Exponential Equilibrium . . . . .	14
3.6	The Two Limits and the Breakdown Threshold . . . . .	16
<b>4</b>	<b>Effective Curvature Under Information Friction</b>	<b>17</b>
4.1	The Firm's Problem . . . . .	17
4.2	Output Loss Under Friction . . . . .	18
4.3	The Effective Curvature Theorem . . . . .	18
4.4	Non-Uniform Degradation and the Crisis Sequence . . . . .	20
4.5	Pre-Crisis Deceleration . . . . .	21
<b>5</b>	<b><math>q</math>-Generalized Dynamical Results</b>	<b>21</b>
5.1	$q$ -Variance-Response Identity . . . . .	22
5.2	$q$ -Covariance Structure . . . . .	22
5.3	$q$ -Crooks Fluctuation Theorem and Jarzynski Equality . . . . .	23
5.4	The $q$ -Log-Sum-Exp Dual . . . . .	24
5.5	$q$ -Dissipation Bound . . . . .	24
5.6	Euler Identity Survival . . . . .	25
5.7	$q$ -Onsager Reciprocal Relations . . . . .	26
5.8	$q$ -Kramers Escape Rate . . . . .	26
5.9	Classification of Results . . . . .	27

<b>6 Derivation I: Akerlof's Market for Lemons</b>	<b>28</b>
6.1 Setup . . . . .	29
6.2 The $\rho$ -Dependent Robustness Result . . . . .	29
6.3 The Informational Interpretation . . . . .	33
6.4 Empirical Implications . . . . .	33
6.5 Dynamic Extension: Bistability . . . . .	34
<b>7 Derivation II: Myerson's Optimal Mechanism</b>	<b>35</b>
7.1 The Virtual Valuation as CES Potential Gradient . . . . .	36
7.2 The Optimal Mechanism Format . . . . .	38
7.3 The Revelation Principle as Variational Equilibrium . . . . .	38
7.4 $\rho$ -Dependent Price of Incentive Compatibility . . . . .	39
7.5 Empirical Implications . . . . .	40
<b>8 Derivation III: Arrow's Impossibility Theorem</b>	<b>41</b>
8.1 $\rho$ -Dependent Democratic Robustness . . . . .	43
8.2 Empirical Implications . . . . .	44
8.3 The Democratic Error Function . . . . .	44
8.4 Connection to the Condorcet Jury Theorem . . . . .	45
<b>9 Derivation IV: Search and Matching</b>	<b>46</b>
9.1 Setup . . . . .	47
9.2 The $\rho$ -Dependent Search Duration . . . . .	47
9.3 The Beveridge Curve . . . . .	49
9.4 Sorting and Assignment . . . . .	49
9.5 Empirical Implications . . . . .	50
9.6 Resolution of the Shimer Volatility Puzzle . . . . .	51
<b>10 Derivation V: Contract Theory</b>	<b>51</b>
10.1 Setup . . . . .	52
10.2 The Hold-Up Problem . . . . .	52
10.3 The Integration Boundary . . . . .	53
10.4 The Fundamental Transformation in $(\rho, T)$ Space . . . . .	53
10.5 Empirical Implications . . . . .	54
10.6 Dynamic Contracting . . . . .	55
<b>11 Derivation VI: Behavioral Economics</b>	<b>55</b>
11.1 The Logit-CES Duality . . . . .	56
11.2 The Behavioral Catalog . . . . .	56
11.3 The Attention Allocation Problem . . . . .	58

11.4 The Rabin Calibration Resolution . . . . .	58
11.5 Empirical Implications . . . . .	60
<b>12 Optimal Firm Scope and Integration Boundary</b>	<b>61</b>
12.1 The Variety–Coordination Tradeoff . . . . .	61
12.2 The Phase Diagram for Integration . . . . .	62
12.3 Supply Chain Architecture . . . . .	65
12.4 Cross-Tier Information Spillovers . . . . .	65
12.5 Aggregate Supply Chain CES Potential . . . . .	66
<b>13 Empirical Predictions</b>	<b>66</b>
13.1 Management–Technology Complementarity . . . . .	67
13.2 Within-Industry Productivity Dispersion . . . . .	67
13.3 AI Adoption Priority . . . . .	68
13.4 Crisis Dynamics Prediction . . . . .	69
<b>14 Empirical Tests</b>	<b>69</b>
14.1 Banking Regulation and the Global Financial Crisis . . . . .	69
14.2 Manufacturing Tail Distributions (Tsallis Test) . . . . .	71
14.3 Price of Incentive Compatibility in Procurement . . . . .	72
<b>15 Six Derivations Revisited Under Tsallis</b>	<b>73</b>
15.1 Akerlof Adverse Selection . . . . .	74
15.2 Myerson Mechanism Design . . . . .	74
15.3 Arrow Social Choice . . . . .	74
15.4 DMP Search and Matching . . . . .	74
15.5 Contract Theory . . . . .	75
15.6 Behavioral Economics . . . . .	75
<b>16 The CES Potential as a Welfare Loss Function</b>	<b>76</b>
<b>17 The Architecture</b>	<b>77</b>
17.1 Welfare Implications of the Architecture . . . . .	78
17.2 The $(\rho, T)$ Classification of Economic Theories . . . . .	78
<b>18 Discussion</b>	<b>79</b>
18.1 Why Wasn’t This Seen Before? . . . . .	79
18.2 Resolution of the Tsallis Debate . . . . .	80
18.3 The $(\rho, T)$ Phase Diagram as a Unifying Framework . . . . .	81
18.4 Superstatistics and the Hierarchical Interpretation . . . . .	81
18.5 Two-Parameter Parsimony . . . . .	82

18.6 Testability . . . . .	82
18.7 Endogenous Technology Choice . . . . .	83
18.8 Implications for Economic Methodology . . . . .	84
18.9 Related Literature on Industrial Organization . . . . .	84
<b>19 Conclusion</b>	<b>86</b>
<b>A Technical Lemmas</b>	<b>90</b>

# 1 Introduction

Modern economics is divided into two broad traditions. The first—encompassing production theory, consumer choice, trade, and growth—concerns how inputs aggregate into outputs and how resources are allocated. Its canonical tool is the production or utility function, and the CES (Constant Elasticity of Substitution) aggregate is its workhorse.<sup>1</sup> The second tradition—encompassing information economics, mechanism design, contract theory, search theory, social choice, and behavioral economics—concerns who knows what, who can commit to what, and how agents behave under uncertainty. Its canonical tools are Bayesian updating, incentive compatibility constraints, and entropy measures of information.

These traditions developed largely independently. They use different mathematical machinery and publish in different journals. Yet a striking pattern emerges when one examines the mathematical structure: every major result in the information tradition can be expressed as a CES aggregate degraded by an entropy term, with severity controlled by the CES curvature parameter.

This paper proposes a unification. The argument is that economic theory is generated by two functions:

1. The **CES generating function**  $\Phi(\rho)$ , governing aggregation, allocation, and market structure;
2. The **Tsallis entropy**  $S_q$  with  $q = \rho$ , governing information, incentives, and bounded rationality;

connected through the **CES potential**:

$$\boxed{\mathcal{F}_q = \Phi_{\text{CES}}(\rho) - T \cdot S_q, \quad q = \rho} \quad (1)$$

where  $\rho \in (-\infty, 1]$  is the CES curvature parameter (equivalently,  $\sigma = 1/(1 - \rho)$  is the elasticity of substitution),  $T \geq 0$  is the information friction, and  $S_q = (1 - \sum p_j^q)/(q - 1)$  is the Tsallis entropy with  $q = \rho$  locked by the emergence theorem. Shannon entropy is the  $q \rightarrow 1$  limit, applicable to perfect substitutes.

Standard economics is the  $T = 0$  limit of this framework—perfect information, deterministic optimization, no search frictions, no behavioral anomalies. The rich structure of information economics, mechanism design, contract theory, social choice, and behavioral economics emerges at  $T > 0$ .

---

<sup>1</sup>The CES function appears as: the Solow production function with capital-labor substitution ([Solow, 1956](#)); the Armington aggregate in trade ([Armington, 1969](#)); the Dixit-Stiglitz love-of-variety aggregate underlying new trade theory ([Dixit and Stiglitz, 1977](#)), new economic geography ([Krugman, 1991](#)), and endogenous growth ([Romer, 1990](#)); the price-setting aggregate in New Keynesian DSGE models; and the Atkinson inequality index in welfare economics ([Atkinson, 1970](#)).

The CES aggregate is not an arbitrary functional form. It is the *unique* constant-elasticity aggregator, and its curvature parameter  $\rho$  simultaneously controls at least four distinct economic properties (the quadruple role established in the companion paper (Smirl, 2026a)). The Tsallis entropy  $S_q$  with  $q = \rho$  is not an arbitrary information measure—it is the *unique* entropy satisfying continuity, maximality, and pseudo-additivity (Tsallis, 1988; Santos, 1997), where pseudo-additivity replaces the Shannon chain rule to capture the non-additive information costs inherent in complementary production ( $\rho < 1$ ). The identification  $q = \rho$  is forced by the emergence theorem (Smirl, 2026b), not fitted as a free parameter. The framework is not a modeling choice—it is the only framework consistent with constant elasticity, axiomatic information measurement matched to production complementarity, and variational consistency simultaneously.

## 1.1 Summary of Results

The framework is demonstrated by deriving key results from six areas of economic theory (Section 6–Section 11), establishing how information friction degrades CES curvature in production (Section 4–Section 12.2), and generalizing the entropy to the Tsallis form matched to production complementarity (Section 3):

1. **Information Economics** (Section 6): Akerlof’s market-for-lemons result emerges as a CES aggregate degrading under positive information friction. *New prediction:* the critical information cost above which the market collapses is increasing in both  $v$  and  $K$ . Markets for complements ( $\rho$  low) resist adverse selection; markets for substitutes ( $\rho$  high) succumb.
2. **Mechanism Design** (Section 7): Myerson’s virtual valuation is identified as the CES potential gradient. *New prediction:* the price of incentive compatibility PoIC( $\rho$ ) increases as  $\rho$  decreases; optimal mechanism format is determined by  $\rho$ .
3. **Social Choice** (Section 8): At  $T > 0$ , non-dictatorial CES aggregation can approximate ordinal IIA with controlled error. *New prediction:* democratic robustness to noise depends on  $\rho$ —majoritarian systems (high  $\rho$ ) tolerate high noise, consensus systems (low  $\rho$ ) are fragile.
4. **Search and Matching** (Section 9): The DMP framework emerges from the CES potential. *New prediction:* search duration  $n^* = (K/T) \cdot Q(\rho)$ —occupations with more complementary skill requirements have longer search durations and steeper Beveridge curves.
5. **Contract Theory** (Section 10): Asset specificity IS low  $\rho$ . *New prediction:* the hold-up distortion  $D(\rho, \tau) = \tau(1 - R(\rho))/2$  yields a sharp integration boundary in  $(\rho, \tau)$  space.

**6. Behavioral Economics** (Section 11): The behavioral catalog emerges as the complete characterization of economic behavior at  $T > 0$ . *New prediction:* loss aversion  $\lambda \approx T_{\text{gain}}/T_{\text{loss}}$  is an information friction ratio, not a utility kink.

Two additional structural results—the effective curvature theorem (Section 4) and the optimal firm scope and integration boundary (Section 12–Section 12.2)—show how CES curvature is degraded by information friction in production, yielding a phase diagram for industrial organization and a crisis sequence that matches observed financial crisis dynamics. The six derivations are revisited under the Tsallis generalization (Section 15), showing that the corrections are systematic: Category A results survive exactly, Category B results acquire the  $1/(2 - q)$  factor, and Category C results replace  $\exp$  with  $\exp_q$ .

Three empirical tests confirm the framework’s predictions: a banking regulation test (Section 14.1), a procurement test (Section 14.3), and a manufacturing tail distribution test (Section 14.2). Four additional predictions—management–technology complementarity, within-industry productivity dispersion, AI adoption priority, and crisis sequence timing—are derived and shown to be testable with existing data (Section 13).

## 1.2 Organization

The paper proceeds as follows. Section 2 restates the CES framework from the companion paper. Section 3 constructs the CES potential and derives the  $q$ -exponential equilibrium. Section 4 derives the effective curvature theorem and crisis sequence. Section 5 classifies all dynamical results under the Tsallis generalization. Sections 6–11 derive results from six areas of economic theory. Section 12–Section 12.2 derive optimal firm scope and integration boundaries. Section 15 revisits the six derivations under Tsallis. Section 13–Section 14 develop predictions and empirical tests. Section 18 discusses methodological implications. Section 19 concludes.

## 1.3 Related Literature

The components of this framework—CES aggregation, Tsallis entropy, and the CES potential connection—exist independently across disparate literatures. The contribution is recognizing their unity: that CES and Tsallis entropy are the two generating functions of economics, connected through the CES potential, with the entire edifice parameterized by  $(\rho, T)$ . This subsection maps each body of existing work into the  $(\rho, T)$  parameter space.

**CES as a generating function across fields.** The CES aggregate has appeared independently in production theory (Arrow et al., 1961), trade (Armington, 1969; Ethier,

1982), industrial organization (Dixit and Stiglitz, 1977), and economic geography (Krugman, 1991). Sato (1967) showed that nested CES enables multi-level aggregation with different  $\rho$  at each tier. Anderson, de Palma, and Thisse (1992) established the formal duality between the CES demand system and the logit choice model, providing the micro-foundation connecting the two generating functions of this paper.

**Entropy and information theory in economics.** Theil (1967) first imported Shannon entropy into economics. The rational inattention program initiated by Sims (2003) derives logit choice from Shannon capacity constraints, interpreting  $T$  as the shadow price of attention. Matějka and McKay (2015) proved that the logit form is the unique solution to rational inattention problems with unrestricted prior, making  $T > 0$  the canonical model of bounded rationality.

**Statistical mechanics and economic equilibrium.** Jaynes (1957) showed that the logit equilibrium distribution maximizes entropy subject to an energy constraint—the direct physical analogue of the economic CES potential minimization. Foley (1994) introduced statistical equilibrium into general equilibrium theory. Smith and Foley (2008) established the formal correspondence between the Helmholtz free energy in thermodynamics and economic general equilibrium. The econophysics program (Yakovenko and Rosser, 2009; Drăgulescu and Yakovenko, 2000; Bouchaud and Mézard, 2000; Gabaix, 2009) demonstrates that statistical mechanics tools apply to economic systems; the present paper identifies the CES aggregation structure that determines *which* equilibrium the entropy dynamics select.

**Active inference and the free energy principle.** Friston (2010) introduced the free energy principle in neuroscience. The present paper shares the conviction that a single variational principle organizes complex adaptive behavior, but the two frameworks minimize their respective potentials over different state spaces. The key object that active inference lacks is the  $\rho$  parameter: two economies with identical information precision but different  $\rho$  will exhibit qualitatively different equilibria, regime shifts, and welfare losses.

**Closest antecedents.** Caplin and Dean (2015) provided the axiomatic foundations for the  $T$  axis through posterior-separable stochastic choice. On the  $\rho$  axis, Costinot (2009) showed that a single log-supermodularity condition—mathematically equivalent to the CES complementarity condition  $\rho < 1$ —unifies Ricardian and Heckscher-Ohlin trade theory. What Caplin–Dean and Costinot leave separate, the CES potential framework joins: the log-supermodularity driving Costinot’s sorting is the CES curvature  $K$ , and the

stochastic choice axiomatized by Caplin–Dean *is* the  $T > 0$  departure from deterministic optimization.

**The O-ring theory.** [Kremer \(1993\)](#) established that complementary production (which he modeled as multiplicative, the limit  $\rho \rightarrow -\infty$  of CES) creates skill sorting, wage inequality, and development traps. The CES potential framework generalizes Kremer’s insight from the multiplicative limit to the full CES family: the phenomena he identifies—positive assortative matching, wage inequality, and development traps—all depend on  $K$ , and their intensity varies continuously with  $\rho$ . The O-ring model’s prediction that “a single low-skill worker destroys the entire output” is the extreme complementarity ( $\rho \rightarrow -\infty$ , Leontief) special case; for finite  $\rho < 1$ , the prediction is softer: a low-skill worker degrades output by  $O(K \cdot \|\delta\|^2/c^2)$ , which is continuous in  $K$ . This matters empirically because most production lies between Kremer’s extreme and perfect substitutability.

**Endogenous growth and CES.** The CES elasticity of substitution between capital and labor has been central to the growth theory debate since [Solow \(1956\)](#). [Jones \(1995\)](#) showed that balanced growth requires the capital-labor elasticity  $\sigma = 1$  (Cobb-Douglas). When  $\sigma > 1$ : capital deepening alone drives growth (no need for technology). When  $\sigma < 1$ : diminishing returns to capital are too strong, and growth requires sustained technological progress. The CES potential framework adds the information dimension: even if  $\sigma > 1$  at the production frontier, the effective elasticity  $\sigma_{\text{eff}} = 1/(1 - \rho_{\text{eff}})$  depends on  $T$ . A high- $T$  economy may have  $\sigma_{\text{eff}} < 1$  even if the underlying technology has  $\sigma > 1$ , trapping the economy in a Solow-type convergence path despite frontier technology. This provides a CES-theoretic explanation for why technology transfer alone does not produce convergence—the information environment must support the exploitation of the transferred technology’s complementarity.

Each of these references uses one piece of the framework. The contribution of this paper is recognizing that these pieces are manifestations of a single object—the Tsallis CES potential  $\mathcal{F}_q = \Phi_{\text{CES}}(\rho) - T \cdot S_q$  with  $q = \rho$ . The CES quadruple role theorem is developed in the companion paper ([Smirl, 2026a](#)).

## 2 Preliminaries: The CES Framework

This section restates the key definitions and results from the companion paper ([Smirl, 2026a](#)) that are used throughout. Proofs are in that paper; here we state what is needed.

## 2.1 CES Production Technology

Consider  $J \geq 2$  inputs  $x_1, \dots, x_J > 0$ . The CES aggregate is:

$$F = \left( \frac{1}{J} \sum_{j=1}^J x_j^\rho \right)^{1/\rho}, \quad \rho \in (-\infty, 1], \rho \neq 0 \quad (2)$$

with the convention  $F = (\prod x_j)^{1/J}$  at  $\rho = 0$  (Cobb-Douglas) and  $F = \min_j x_j$  as  $\rho \rightarrow -\infty$  (Leontief).

The elasticity of substitution is  $\sigma = 1/(1 - \rho)$ . The CES curvature at symmetric equilibrium is:

$$K = \frac{(1 - \rho)(J - 1)}{J} \quad (3)$$

The CES potential (the generating function of the associated conservative-dissipative system) is:

$$\Phi = - \sum_n \log F_n \quad (4)$$

The Hessian of  $\log F$  at the symmetric point  $\bar{\mathbf{x}} = c \cdot \mathbf{1}$  has eigenvalue 0 on  $\text{span}\{\mathbf{1}\}$  (Euler's theorem) and eigenvalue

$$\lambda_\perp = -\frac{(1 - \rho)}{Jc^2} = -\frac{K}{(J - 1)c^2} \quad (5)$$

with multiplicity  $J - 1$  on the tangent space  $\mathbf{1}^\perp = \{\mathbf{v} : \sum_j v_j = 0\}$ .

## 2.2 The Quadruple Role of $\rho$

The companion paper ([Smirl, 2026a](#)) establishes that  $K$  (equivalently  $\rho$ ) simultaneously controls four properties. They are stated here without proof.

**Theorem 2.1** (CES Quadruple Role, [Smirl 2026a](#)). *At symmetric equilibrium with  $\bar{x}_j = c$  for all  $j$ , the CES curvature  $K$  defined in (3) simultaneously determines:*

- (a) **Superadditivity.** For any perturbation  $\delta$  with  $\sum \delta_j = 0$ :

$$F(\bar{x} + \delta) \leq F(\bar{x}) - \frac{K}{2(J - 1)} \cdot \frac{\|\delta\|^2}{c}$$

*The gap between homogeneous and heterogeneous allocation is proportional to  $K$ .*

- (b) **Correlation robustness.** If  $x_j = \mu + \tau Z_j$  with  $\text{Var}[Z_j] = 1$ :

$$\frac{\text{Var}[Y]}{\tau^2} = 1 - \frac{K^2}{(J - 1)} + O(K^3)$$

*CES nonlinearity extracts idiosyncratic variation with bonus proportional to  $K^2$ .*

- (c) *Strategic independence.* For any coalition  $S \subset \{1, \dots, J\}$  deviating by  $\delta$ :

$$\Delta v(S) \leq -\frac{K}{2(J-1)} \cdot \frac{\|\delta\|^2}{c^2}$$

*Balanced allocation is a Nash equilibrium; coalition deviations are penalized proportionally to  $K$ .*

- (d) *Network scaling.* The unnormalized CES aggregate  $G = (\sum x_j^\rho)^{1/\rho}$  with  $J$  symmetric inputs satisfies  $G(J) = J^{1/\rho} \cdot c$ . The network scaling exponent is  $1/\rho$ .

The key structural fact: superadditivity and strategic independence are first-order curvature effects (proportional to  $K$ ), while correlation robustness is a second-order effect (proportional to  $K^2$ ). This distinction drives the non-uniform degradation result in Section 4.4.

### 2.3 Micro-Foundations: Why CES Is Forced

CES aggregation is the unique functional form satisfying the following axiom sets (proved in Smirl, 2026a):

- Proposition 2.2** (CES from primitives). (a) *Consistent aggregation (Kolmogorov–Nagumo).* Consistent sub-aggregation + homogeneity forces the CES mean (Aczél, 1966; Hardy, Littlewood, and Pólya, 1952).
- (b) *Optimal selection (Fréchet).* Buyers with i.i.d. Fréchet taste shocks choosing quality-maximizing varieties produce CES market shares and welfare indices (Anderson, de Palma, and Thisse, 1992; Eaton and Kortum, 2002).
- (c) *Inequality-averse social evaluation (Atkinson).* Weak Pareto + anonymity + homotheticity + Pigou–Dalton uniquely selects the Atkinson–CES family (Atkinson, 1970).

### 2.4 Nested CES and Multi-Level Aggregation

Sato (1967) showed that CES aggregation can be nested: inputs are first grouped into clusters, each cluster is aggregated with its own CES parameter  $\rho_g$ , and the cluster-level outputs are then aggregated with a different parameter  $\rho_0$ . The two-level nested CES takes the form:

$$F = \left( \frac{1}{G} \sum_{g=1}^G \left( \frac{1}{J_g} \sum_{j \in \mathcal{G}_g} x_j^{\rho_g} \right)^{\rho_0/\rho_g} \right)^{1/\rho_0} \quad (6)$$

where  $G$  is the number of groups,  $\mathcal{G}_g$  is the set of inputs in group  $g$ ,  $J_g = |\mathcal{G}_g|$ ,  $\rho_g$  governs within-group substitutability, and  $\rho_0$  governs between-group substitutability. The consistency axiom (Proposition 2.2(a)) guarantees that this nesting is self-consistent: the aggregate does not depend on the order in which sub-aggregation is performed, a property that fails for non-CES aggregators.

**Proposition 2.3** (Nested Curvature). *The effective curvature of a two-level nested CES at symmetric equilibrium is:*

$$K_{\text{nested}} = \underbrace{\frac{(1 - \rho_0)(G - 1)}{G}}_{K_{\text{between}}} + \underbrace{\frac{1}{G} \sum_{g=1}^G \frac{(1 - \rho_g)(J_g - 1)}{J_g}}_{K_{\text{within}}} \quad (7)$$

*The total curvature decomposes additively into between-group and within-group components.*

*Proof.* At symmetric equilibrium ( $x_j = c$  for all  $j$ ), the nested CES reduces to  $F = c$ . The Hessian  $\nabla^2 \log F$  has a block structure respecting the group partition. The eigenvalues decompose into: (i)  $G - 1$  eigenvalues equal to  $-(1 - \rho_0)/(Gc^2)$  for perturbations that shift resources between groups (holding within-group allocation fixed); (ii) for each group  $g$ ,  $J_g - 1$  eigenvalues equal to  $-(1 - \rho_g)/(J_g c^2)$  for perturbations within group  $g$  (holding total group allocation fixed). The effective curvature  $K_{\text{nested}} = -c^2 \sum_k \lambda_k / (J_{\text{total}} - 1)$  gives (7) after summation.  $\square$

**Remark 2.4** (Hierarchical information friction). *The nested structure has a natural information-theoretic interpretation: the firm processes information at two levels. At the coarse level (between groups), decisions are made with friction  $T_0$ ; at the fine level (within groups), with friction  $T_g$ . The CES potential decomposes correspondingly:*

$$\mathcal{F}_q = \underbrace{\mathcal{F}_0(K_{\text{between}}, T_0)}_{\text{inter-group}} + \underbrace{\sum_g \mathcal{F}_g(K_g, T_g)}_{\text{intra-group}}$$

*This additive decomposition is a consequence of the CES nesting consistency and motivates the supply chain architecture results of Section 12.2: each tier of production has its own  $(\rho, T)$  pair, and the total CES potential is the sum.*

**Remark 2.5** (Connection to endogenous growth). *The nested CES structure is central to endogenous growth theory. In Romer (1990), variety expansion increases the number of inputs  $J$  while holding  $\rho$  fixed. The growth rate depends on the rate of variety creation, which in turn depends on the R&D sector's ability to reduce information friction (identify new varieties). The CES potential framework makes this precise: the growth rate is*

bounded by  $\partial\mathcal{F}_q/\partial J = (1/\rho J) - T \cdot (\partial H^*/\partial J)$ , the difference between the marginal diversity premium and the marginal coordination cost. Growth slows when coordination costs dominate, providing a CES-theoretic interpretation of semi-endogenous growth ([Jones, 1995](#)).

## 3 The CES Potential and Tsallis Entropy

### 3.1 From Shannon to Tsallis: Why Complementarity Breaks Extensivity

For a discrete probability distribution  $p = (p_1, \dots, p_N)$ , Shannon entropy  $H(p) = -\sum p_i \log p_i$  is uniquely characterized by three axioms ([Khinchin, 1957](#)): continuity, maximality at the uniform distribution, and the *chain rule* (additivity) for compound experiments. The chain rule assumes that information about different inputs is additive—learning about inputs  $A$  and  $B$  jointly costs the sum of learning about each separately.

This is exactly wrong for complementary production. When  $\rho < 1$ , the cross-partial  $\partial^2 F/\partial x_j \partial x_k > 0$  means that learning about input  $j$  changes the marginal value of information about input  $k$ . At the symmetric point  $x_j = \bar{x}$ :

$$\left. \frac{\partial^2 F}{\partial x_j \partial x_k} \right|_{\text{sym}} = \frac{(\rho - 1)}{J^2} \cdot \bar{x}^{\rho-2} \cdot F^{1-\rho} \quad (8)$$

which is positive (complements) when  $\rho < 1$  and zero when  $\rho = 1$  (perfect substitutes). Information about complements is inherently non-additive.

Replacing the chain rule with *pseudo-additivity*— $S(A \cup B) = S(A) + S(B|A) + (1-q)S(A)S(B|A)$ , where the interaction term  $(1-q)$  captures the non-additive component—while retaining continuity and maximality yields the Tsallis entropy ([Tsallis, 1988](#); [Santos, 1997](#); [Suyari, 2004](#)):

$$S_q(p) = \frac{1 - \sum_{i=1}^N p_i^q}{q - 1} \quad (9)$$

with  $S_q \rightarrow H$  (Shannon) as  $q \rightarrow 1$ . The interaction term has direct economic content:

- $q < 1$  (complements): joint information exceeds the sum—complementary inputs amplify each other's information value.
- $q = 1$  (neutral): Shannon chain rule recovered exactly.
- $q > 1$  (substitutes): joint information is less than the sum—redundant inputs diminish each other's information value.

**Theorem 3.1** (Tsallis uniqueness). *Under continuity, maximality, and the  $q$ -chain rule replacing the Shannon chain rule, the unique entropy functional is the Tsallis entropy (9) for  $q \neq 1$ , with  $S_q \rightarrow H$  as  $q \rightarrow 1$ .*

*Proof.* The result follows from the characterization theorems of Santos (1997), Abe (2000), and Suyari (2004). Define  $g(S) = 1 + (1 - q)S$ . Pseudo-additivity with independent subsystems gives  $g(S(AB)) = g(S(A)) \cdot g(S(B))$ , so  $g \circ S$  is multiplicative. The unique continuous multiplicative functional on product distributions is  $g(S) = \sum p_j^q$ . Solving for  $S$  yields (9).  $\square$

The identification  $q = \rho$  is forced by the companion emergence theorem (Smirl, 2026b): CES with parameter  $\rho$  is the sufficient statistic for Rényi/Tsallis entropy of order  $\rho$ . Since  $q = \rho$  from the production technology, the Tsallis generalization adds no free parameters.

## 3.2 The Information Friction

Following the rational inattention literature (Sims, 2003), the information friction  $T$  is the shadow price of information processing capacity. An agent with processing capacity  $\kappa$  has  $T = 1/\kappa$ . At  $T = 0$ : unlimited processing capacity, perfect rationality. As  $T$  increases: information is costlier, agents are more boundedly rational, decisions are noisier.

The agent's choice probability follows the  $q$ -exponential form:

$$P(\text{choose } a_i) = \frac{\exp_q(u(a_i)/T)}{\sum_k \exp_q(u(a_k)/T)} \quad (10)$$

where  $\exp_q(x) = [1 + (1 - q)x]_+^{1/(1-q)}$  is the  $q$ -exponential and  $q = \rho$ . For  $q \rightarrow 1$ , this recovers the standard logit  $P \propto \exp(u/T)$ . For  $q < 1$  (complements), the distribution has compact support—agents entirely ignore options beyond a cost threshold  $T/(1 - q)$ , producing the “all-or-nothing” patterns observed in complementary team production (Kremer, 1993). For  $q > 1$  (substitutes), the distribution has power-law tails with exponent  $\sigma = 1/(1 - \rho)$ .

## 3.3 Construction of the CES Potential

**Theorem 3.2** (Uniqueness of the CES Potential). *Let  $\mathcal{F}[\Phi, \mathcal{H}, T]$  be a functional governing economic equilibrium, where  $\Phi$  is an aggregator,  $\mathcal{H}$  is an information measure, and  $T \geq 0$  is a scalar parameter. Suppose:*

- (A1) **Constant elasticity.** *The aggregator  $\Phi$  has constant elasticity of substitution  $\sigma = 1/(1 - \rho)$ .*
- (A2) **Generalized Khinchin axioms.** *The information measure  $\mathcal{H}$  satisfies continuity, maximality, and pseudo-additivity with parameter  $q$  determined by the aggregator.*

(A3) **Constrained optimality.** Equilibrium maximizes  $\Phi$  subject to an information capacity constraint  $\mathcal{H} \leq \kappa$ , where  $T$  is the shadow price of capacity.

(A4) **Self-consistency.** The entropy order matches the aggregation parameter:  $q = \rho$ .

Then, up to positive affine transformations:

- (i)  $\Phi$  must be the CES aggregate (or its limits: Cobb-Douglas, Leontief);
- (ii)  $\mathcal{H}$  must be Tsallis entropy  $S_q = (1 - \sum p_i^q)/(q - 1)$  with  $q = \rho$ ;
- (iii)  $\mathcal{F}_q = \Phi_{\text{CES}}(\rho) - T \cdot S_q$ .

Shannon entropy is recovered as the  $q \rightarrow 1$  (perfect substitutes) special case.

*Proof.* Step 1: (A1) forces CES. By the functional equation theorem of Aczél (1966), the CES family is the unique class of linearly homogeneous aggregators with constant elasticity of substitution.

Step 2: (A2) forces Tsallis entropy. By the generalized Khinchin theorem (Santos, 1997; Abe, 2000; Suyari, 2004),  $S_q$  is the unique function satisfying the three axioms.

Step 3: (A4) forces  $q = \rho$ . The self-consistency axiom follows from the emergence theorem (Smirl, 2026b): CES with parameter  $\rho$  is the sufficient statistic for Rényi/Tsallis entropy of order  $\rho$ .

Step 4: Lagrangian duality forces  $\mathcal{F}_q = \Phi - T \cdot S_q$ . By (A3), equilibrium solves  $\max_{\mathbf{x}, \mathbf{p}} \Phi(\mathbf{x})$  subject to  $S_q(\mathbf{p}) \leq \kappa$ . The Lagrangian is  $\mathcal{L} = \Phi(\mathbf{x}) + T[\kappa - S_q(\mathbf{p})]$ . Strong duality gives  $\min_{T \geq 0} \max_{\mathbf{x}, \mathbf{p}} \Phi(\mathbf{x}) - T \cdot S_q(\mathbf{p})$ . The additive structure is forced by Lagrangian duality; no alternative connector is consistent with constrained optimization.  $\square$

### 3.4 The $q$ -Exponential and $q$ -Logarithm

The Tsallis CES potential generates a natural pair of deformed functions that replace the standard exponential and logarithm throughout the framework.

**Definition 3.3** ( $q$ -exponential and  $q$ -logarithm). *The  $q$ -exponential and  $q$ -logarithm are:*

$$\exp_q(x) = [1 + (1 - q)x]_+^{1/(1-q)} \quad (11)$$

$$\ln_q(x) = \frac{x^{1-q} - 1}{1 - q} \quad (12)$$

with  $\exp_q \rightarrow \exp$  and  $\ln_q \rightarrow \log$  as  $q \rightarrow 1$ . They satisfy  $\ln_q(\exp_q(x)) = x$  and  $\exp_q(\ln_q(x)) = x$  (inverse pair). The  $q$ -product rule is  $\exp_q(a) \cdot \exp_q(b) = \exp_q(a + b + (1 - q)ab)$ , which encodes pseudo-additivity: the product of two  $q$ -exponentials is not the  $q$ -exponential of the sum, but of the sum plus an interaction term proportional to  $(1 - q)$ .

**Remark 3.4** (Economic meaning of the  $q$ -product rule). *The interaction term  $(1-q)ab$  in the  $q$ -product rule is the information-theoretic manifestation of complementarity. When two CES systems with energy costs  $a$  and  $b$  are combined, the joint cost is not  $a + b$  (additive) but  $a + b + (1 - q)ab$ . For  $q < 1$  (complements),  $(1 - q)ab > 0$ : combining complementary subsystems costs more than the sum, because coordination between subsystems requires additional information. For  $q > 1$  (substitutes),  $(1 - q)ab < 0$ : combining substitutable subsystems costs less than the sum, because redundancy provides free information.*

### 3.5 The $q$ -Exponential Equilibrium

**Proposition 3.5** ( $q$ -exponential equilibrium). *The equilibrium distribution that minimizes  $\mathcal{F}_q$  subject to normalization  $\sum p_j = 1$  is:*

$$p_j^* = \frac{[1 - (1 - q)\beta\varepsilon_j]_+^{1/(1-q)}}{Z_q} \quad (13)$$

where  $\beta = 1/T$ ,  $[x]_+ = \max(x, 0)$ , and  $Z_q = \sum_j [1 - (1 - q)\beta\varepsilon_j]_+^{1/(1-q)}$  is the  $q$ -partition function.

*Proof.* The Lagrangian for minimizing  $\mathcal{F}_q = \sum p_j\varepsilon_j - TS_q$  subject to  $\sum p_j = 1$  is:

$$\mathcal{L} = \sum_j p_j\varepsilon_j - T \cdot \frac{1 - \sum_j p_j^q}{q-1} - \lambda \left( \sum_j p_j - 1 \right)$$

The first-order condition gives:

$$\varepsilon_j + \frac{Tq}{q-1} p_j^{q-1} = \lambda$$

Solving for  $p_j$ :

$$p_j = \left[ \frac{(q-1)(\lambda - \varepsilon_j)}{Tq} \right]^{1/(q-1)}$$

Defining the  $q$ -exponential  $\exp_q(x) = [1 + (1 - q)x]_+^{1/(1-q)}$  and absorbing constants into the normalization yields  $p_j^* \propto \exp_q(-\beta\varepsilon_j)$ .  $\square$

**Proposition 3.6** (Properties of the  $q$ -equilibrium). (a) **Compact support for  $q < 1$  (complements):**  $p_j^* = 0$  for all  $j$  with  $\varepsilon_j > T/(1 - q)$ . High-cost inputs are excluded entirely, not just exponentially suppressed.

(b) **Power-law tails for  $q > 1$  (substitutes):**  $p_j^* \sim \varepsilon_j^{-\zeta}$  for large  $\varepsilon_j$ , with tail exponent  $\zeta = 1/(q-1) = \sigma$ , connecting CES theory to power-law economics (Gabaix, 2009).

(c) **Boltzmann recovery:** As  $q \rightarrow 1$ ,  $p_j^* \rightarrow e^{-\beta\varepsilon_j}/Z$  (the standard logit distribution).

*Proof.* (a) When  $q < 1$ ,  $1 - q > 0$ , and the bracket  $[1 - (1 - q)\beta\varepsilon_j]_+$  vanishes for  $\varepsilon_j > 1/((1 - q)\beta) = T/(1 - q)$ . The threshold  $T/(1 - q)$  is finite, so the distribution has bounded support.

(b) When  $q > 1$ ,  $1 - q < 0$ , so  $\exp_q(-\beta\varepsilon_j) = (1 + (q - 1)\beta\varepsilon_j)^{-1/(q-1)} \sim \varepsilon_j^{-1/(q-1)}$  for large  $\varepsilon_j$ . The tail exponent is  $\zeta = 1/(q - 1)$ , and since  $q = \rho$  and  $\sigma = 1/(1 - \rho)$ , we have  $\zeta = \sigma$  for  $\rho > 1$ .

(c) Standard:  $\lim_{q \rightarrow 1} [1 + (1 - q)x]^{1/(1-q)} = e^x$  by L'Hôpital's rule.  $\square$

**Remark 3.7** (Economic content of compact support). *Compact support for complements ( $q < 1$ ) has direct economic meaning: when inputs are strong complements, the producer entirely ignores inputs whose cost exceeds the threshold  $T/(1 - q)$ . This is sharper than exponential suppression—it is a hard cutoff. In a labor market with complementary skills, firms do not hire workers above a salary threshold; they simply exclude them from the candidate pool. This matches the “all-or-nothing” hiring patterns observed in complementary team production (Kremer, 1993): surgical teams do not hire cheaper but less qualified surgeons at a discount—they hire qualified surgeons or not at all.*

**Remark 3.8** (Power-law tails and the size distribution of firms). *The power-law tail for substitutes ( $q > 1$ ) with exponent  $\sigma = 1/(1 - \rho)$  connects to the empirical size distribution of firms. In industries with high  $\rho$  (substitutable inputs), the equilibrium firm size distribution should follow a power law with exponent related to  $\sigma$ . This provides a CES-theoretic foundation for the well-documented Zipf-like distributions of firm sizes (Gabaix, 2009), where the tail exponent is determined by the elasticity of substitution in the underlying production technology.*

### 3.6 The Two Limits and the Breakdown Threshold

The CES potential  $\mathcal{F}_q = \Phi_{\text{CES}}(\rho) - T \cdot S_q$  has two canonical limits that bracket the full framework:

**$T = 0$  (Perfect information):**  $\mathcal{F}_q = \Phi_{\text{CES}}$ . Pure CES optimization. This is standard economics: deterministic allocation, no behavioral anomalies, no information rents. The allocation is the first-best (15), the mechanism is fully efficient (zero PoIC), markets function at all quality levels (no adverse selection), and choice is deterministic ( $P(j) \in \{0, 1\}$ ).

**$T > 0$  (Information friction):** The entropy term degrades the CES aggregate. Higher  $T$  means costlier information, coarser decisions, worse allocation. The severity depends on  $\rho$ : low  $\rho$  (high curvature  $K$ ) is more robust because the value of precise allocation is higher.

$T \rightarrow \infty$  (**No information**):  $\mathcal{F}_q = -T \cdot S_q \rightarrow 0$ . The CES contribution vanishes. Allocation is uniform ( $x_j = C/J$  for all  $j$ ), choice is random ( $P(j) = 1/J$ ), no market can sustain screening, and all behavior is noise. This is the “zero-intelligence” limit studied in agent-based computational economics.

The rich structure of economic theory—information rents, behavioral biases, institutional failure—resides entirely in the intermediate regime  $0 < T < \infty$ . The CES potential makes this precise by specifying exactly how each phenomenon depends on the position in  $(\rho, T)$  space.

**The breakdown threshold.** For each economic institution, there exists a breakdown threshold  $T^*$  above which the institution fails. The critical surface  $T^* = T^*(v, \rho)$  satisfies:

$$\frac{\partial T^*}{\partial v} > 0, \quad \frac{\partial T^*}{\partial K} > 0 \quad (14)$$

Higher aggregate value  $v$  and higher curvature  $K$  both raise the failure threshold. Below  $T^*$ : the institution functions. Above  $T^*$ : it collapses via a first-order regime shift (Proposition 6.1).

**Domain-specific breakdown thresholds.** Different economic institutions have different  $T^*$  values because they depend on different components of the CES potential:

Institution	$T^*$ depends on	Failure mode	Section
Markets (Akerlof)	$v, K, \bar{q}$	Market unraveling	Section 6
Mechanisms (Myerson)	$K, \sigma_\theta$	Excessive information rents	Section 7
Democracy (Arrow)	$\rho, J$	Ordering reversals	Section 8
Search (DMP)	$K, L$	Matching breakdown	Section 9
Contracts (GHM)	$\rho, \tau$	Hold-up paralysis	Section 10
Choice (KT)	$u_{\max}, J$	Choice overload	Section 11
Production	$K, d^2, C$	Effective substitutability	Section 4

The ordering of  $T^*$  values across institutions determines which institutions fail first during an information crisis. This is the generalization of the non-uniform degradation result (Section 4.4) from the three curvature roles to the full institutional landscape.

## 4 Effective Curvature Under Information Friction

When a firm with CES technology (structural parameter  $\rho$ ) operates under information friction ( $T$ ), the complementarity it can *exploit* depends on both parameters jointly.

## 4.1 The Firm's Problem

A firm operates CES technology (2) with  $J$  input types and structural parameter  $\rho$ . Input  $j$  has productivity  $\theta_j > 0$ , drawn i.i.d. with mean  $\mu$  and variance  $\sigma_\theta^2$ . The firm has total budget  $C$  and must allocate  $x_j \geq 0$  with  $\sum_j x_j = C$ .

Under perfect information ( $T = 0$ ), the optimal allocation is the CES demand system:

$$x_j^* = C \cdot \frac{\theta_j^{\sigma-1}}{\sum_{k=1}^J \theta_k^{\sigma-1}}, \quad \sigma = \frac{1}{1-\rho} \quad (15)$$

Under information frictions ( $T > 0$ ), the firm does not observe  $\boldsymbol{\theta}$  directly. It acquires information subject to a capacity constraint  $I(\boldsymbol{\theta}; \mathbf{s}) \leq \kappa = 1/T$ , where  $\mathbf{s}$  is the signal vector. The firm's problem is:

$$\max_{\mathbf{x}(\mathbf{s})} \mathbb{E}_{\boldsymbol{\theta}, \mathbf{s}}[\log F(\mathbf{x}(\mathbf{s}); \boldsymbol{\theta})] - T \cdot I(\boldsymbol{\theta}; \mathbf{s}) \quad (16)$$

subject to  $\sum_j x_j(\mathbf{s}) = C$  for all  $\mathbf{s}$ . The objective is the production CES potential: expected log-output minus the entropy cost of information acquisition.

**Remark 4.1.** Under symmetric prior (all  $\theta_j$  identically distributed) and the logit solution (10), the firm's allocation converges to the equal-share allocation  $\hat{x}_j = C/J$  as  $T \rightarrow \infty$ , and to the first-best  $x_j^*$  as  $T \rightarrow 0$ . At intermediate  $T$ , the allocation is a “softened” version of the first-best: the firm leans toward the optimal allocation but with noise proportional to  $T$ .

The allocation error  $\boldsymbol{\delta} = \hat{\mathbf{x}} - \mathbf{x}^*$  lies in the tangent space  $\mathbf{1}^\perp$  (the budget constraint  $\sum_j \delta_j = 0$  is always satisfied). Under the logit allocation, the per-component error variance is:

$$\mathbb{E}[\delta_j^2] = c^2 \cdot \frac{\sigma_\theta^2}{\mu^2} \cdot \frac{T}{T + |\lambda_\perp|^{-1} \sigma_\theta^2 / \mu^2} \quad (17)$$

where  $|\lambda_\perp| = K/((J-1)c^2)$  is the curvature of  $\log F$  on the tangent space. In the regime  $T \ll c^2/K$ , this simplifies to:

$$\mathbb{E}[\delta_j^2] \approx \frac{T \cdot K}{J-1} \quad (18)$$

The error variance is jointly proportional to  $T$  (noisier information, larger errors) and  $K$  (more curvature amplifies the cost of misallocation): the same complementarity that makes precise allocation valuable also amplifies the penalty from imprecise allocation.

## 4.2 Output Loss Under Friction

At intermediate  $T$ , the allocation is a “softened” version of the first-best. The allocation error  $\boldsymbol{\delta} = \hat{\mathbf{x}} - \mathbf{x}^*$  lies in the tangent space  $\mathbf{1}^\perp$  (the budget constraint is always satisfied).

Taylor-expanding  $\log F$  around the first-best:

$$\log F(\hat{\mathbf{x}}) = \log F(\mathbf{x}^*) + \nabla(\log F)^\top \boldsymbol{\delta} + \frac{1}{2} \boldsymbol{\delta}^\top \nabla^2(\log F) \boldsymbol{\delta} + O(\|\boldsymbol{\delta}\|^3)$$

At the first-best,  $\nabla(\log F)^\top \boldsymbol{\delta} = 0$  since  $\boldsymbol{\delta} \in \mathbf{1}^\perp$ . The expected output loss is dominated by the quadratic term with eigenvalue  $|\lambda_\perp| = K/((J-1)c^2)$  from (5):

$$\mathbb{E}[\log F^* - \log F(\hat{\mathbf{x}})] \approx \frac{K^2 \cdot T}{2(J-1)c^2} \quad (19)$$

The loss is quadratic in  $K$  (more complementary technologies suffer more from misallocation) and linear in  $T$  (costlier information means larger errors). This is the key intermediate result connecting production complementarity to information friction.

### 4.3 The Effective Curvature Theorem

**Theorem 4.2** (Effective Curvature). *Consider a CES production technology with  $J$  inputs, structural curvature  $K$ , operated by a firm at information friction  $T$ . The effective curvature is*

$$K_{\text{eff}}(\rho, T) = K \cdot \left(1 - \frac{T}{T^*(\rho)}\right)^+ \quad (20)$$

where

$$T^*(\rho) = \frac{(J-1)c^2 d^2}{K} \quad (21)$$

is the critical production friction,  $d^2$  is the squared geodesic diversity of the input mix, and  $(z)^+ = \max(z, 0)$ .

*Proof.* The superadditivity gap under perfect information is  $\Delta_{\text{FB}} = \Omega(K) \cdot d^2$ . Under information friction  $T$ , the expected gap is reduced by the output loss (19):

$$\Delta_{\text{eff}} = K \cdot d^2 \left(1 - \frac{KT}{2(J-1)c^2 d^2}\right).$$

Setting  $\Delta_{\text{eff}} = K_{\text{eff}} \cdot d^2$  and solving gives (20). Above  $T^*$ , the output loss from misallocation exceeds the diversity premium, so no complementarity benefit remains.  $\square$

**Remark 4.3** (Interpretation). *Information friction acts as an endogenous source of apparent substitutability. A firm with low  $\rho$  but high  $T$  behaves as if its inputs were more substitutable. This resolves a puzzle in development economics: why firms in developing countries use simpler production processes even when the same technology is available (Bloom et al., 2013). The answer is not that the technology is inappropriate, but that the informational environment ( $T$  too high relative to  $T^*(\rho)$ ) makes the complementarity unexploitable.*

**Remark 4.4** (Properties of  $T^*$ ). *The critical production friction has economically natural comparative statics:*

- (i)  $\partial T^*/\partial(1 - \rho) > 0$  for fixed  $d^2$ : more complementary technologies tolerate more information friction before the premium vanishes. This follows because  $T^* \propto (J - 1)c^2d^2/K$  and, for realistic diversity levels, the diversity  $d^2$  grows faster with complementarity than  $K$  does.
- (ii)  $\partial T^*/\partial d^2 > 0$ : more diverse input mixes provide a larger buffer against information degradation. A firm with homogeneous inputs ( $d^2 = 0$ ) has  $T^* = 0$ —there is no complementarity to degrade.
- (iii)  $\partial T^*/\partial C > 0$ : larger firms (higher budget  $C$ , hence higher  $c = C/J$ ) are more robust to information frictions. Scale provides a buffer because the output loss is a fraction of a larger base. This provides a rationale for firm size beyond standard returns-to-scale arguments: large firms can exploit complementarity that small firms cannot.

## 4.4 Non-Uniform Degradation and the Crisis Sequence

**Proposition 4.5** (Non-Uniform Degradation). *Under information friction  $T < T^*(\rho)$ :*

- (a) **Superadditivity:** Linear degradation in  $T/T^*$ . The effective gap is  $\Omega(K(1 - T/T^*)) \cdot d^2$ .
- (b) **Correlation robustness:** Quadratic degradation in  $T/T^*$ . The effective bonus scales as  $K^2(1 - T/T^*)^2$ .
- (c) **Strategic independence:** Linear degradation in  $T/T^*$ . The effective penalty is  $\Omega(K(1 - T/T^*)) \cdot \|\delta_S\|^2/c^2$ .

*Proof.* Parts (a) and (c) follow from replacing  $K$  with  $K_{\text{eff}} = K(1 - T/T^*)$  in the linear bounds. Part (b): the  $K^2$  bound receives two factors of  $(1 - T/T^*)$ —one from the curvature channel operating on a noisier signal, one from the reduced correlation contrast.  $\square$

**Corollary 4.6** (Crisis Sequence). *As  $T$  rises toward  $T^*$ :*

1. **Correlation robustness fails first.** Since the bonus scales as  $(1 - T/T^*)^2$ , it falls below economic significance at  $T \approx T^*/2$ .
2. **Superadditivity degrades next.** The diversity premium erodes linearly, reaching zero at  $T = T^*$ .

3. *Strategic independence breaks last.* The manipulation penalty erodes linearly, but defection is a threshold event that occurs only after the diversity premium has substantially eroded.

**Example 4.7** (The 2008 Global Financial Crisis). *The crisis sequence matches observed GFC dynamics:*

1. *Diversification failed first:* asset correlations spiked from  $r \approx 0.3$  to  $r > 0.8$  across supposedly diversified portfolios.
2. *Production complementarities degraded next:* supply chain disruptions spread as trade finance froze.
3. *Strategic stability broke last:* Lehman's counterparties defected, money market funds broke the buck, interbank lending ceased.

*The sequence is forced by the mathematics of non-uniform degradation.*

## 4.5 Pre-Crisis Deceleration

The effective curvature theorem has a dynamical consequence: as  $T$  approaches  $T^*$  from below, the system's adjustment speed decreases.

**Proposition 4.8** (Pre-Crisis Deceleration). *Consider a CES economy adjusting toward equilibrium after a perturbation. The relaxation time  $\tau(T)$ —the time for the perturbation to decay by a factor of  $1/e$ —satisfies:*

$$\tau(T) = \frac{\tau_0}{1 - T/T^*(\rho)} \quad (22)$$

where  $\tau_0$  is the relaxation time at  $T = 0$ . As  $T \rightarrow T^*$ :  $\tau \rightarrow \infty$  (the system ceases to adjust).

*Proof.* The linearized dynamics around the symmetric equilibrium are  $\dot{\mathbf{x}} = -\mathbf{L}\nabla^2\mathcal{F}_q \cdot (\mathbf{x} - \mathbf{x}^*)$ , where  $\mathbf{L}$  is the (symmetric, positive definite) friction matrix. The eigenvalues of  $\mathbf{L}\nabla^2\mathcal{F}_q$  on the tangent space  $\mathbf{1}^\perp$  are proportional to  $K_{\text{eff}} = K(1 - T/T^*)$ . The relaxation rate (inverse of relaxation time) is the smallest eigenvalue:  $1/\tau = K_{\text{eff}} \cdot L_{\min}/c^2 = (K/c^2)(1 - T/T^*) \cdot L_{\min}$ . Setting  $\tau_0 = c^2/(K \cdot L_{\min})$  gives (22).  $\square$

**Remark 4.9** (Early warning signals). *The divergence of  $\tau$  as  $T \rightarrow T^*$  is a Category A result (Table 1): it depends only on the eigenstructure of the CES potential Hessian, not on the specific entropy. This means that pre-crisis deceleration provides a reliable early warning signal under both Shannon and Tsallis frameworks. Observable consequences:*

1. **Increased autocorrelation:** the autocorrelation coefficient  $\rho(\Delta t) = e^{-\Delta t/\tau}$  approaches 1 as  $\tau \rightarrow \infty$ .

2. **Increased variance:** by the  $q$ -variance-response identity,  $\text{Var} = T\tau/(2-q)$ , which diverges as  $\tau \rightarrow \infty$ .
3. **Flickering:** near  $T^*$ , small perturbations cause temporary excursions to the alternative (collapsed) equilibrium before the system returns.

All three signals have been documented empirically in pre-crisis financial data, providing indirect support for the effective curvature mechanism.

## 5 $q$ -Generalized Dynamical Results

This section derives the key dynamical results under Tsallis entropy, identifying what changes relative to the Shannon ( $q = 1$ ) framework and what survives.

### 5.1 $q$ -Variance-Response Identity

**Theorem 5.1** ( $q$ -Variance-Response Identity (FDT)). *At equilibrium of the Tsallis CES potential  $\mathcal{F}_q$ , the covariance matrix of fluctuations  $\Sigma$  and the linear response matrix  $\chi$  satisfy:*

$$\Sigma = \frac{T}{2-q} \cdot \chi \quad (23)$$

*Proof.* The Hessian of the Tsallis entropy at the uniform distribution is  $\partial^2 S_q / \partial p_j \partial p_k|_{\text{unif}} = -qJ^{q-2}\delta_{jk}$ , differing from Shannon's  $-J\delta_{jk}$  by a factor that evaluates to  $1/(2-q)$  at symmetric equilibrium. At the symmetric equilibrium ( $p_j = 1/J$ ), the linear response from the  $q$ -exponential gives  $\chi_{jk} = \beta J^{q-2}(\delta_{jk} - 1/J)/q$ , and the equilibrium covariance is  $\Sigma_{jk} = J^{q-2}(\delta_{jk} - 1/J)/(q(2-q))$ . The ratio  $\Sigma_{jk}/\chi_{jk} = T/(2-q)$  establishes (23).  $\square$

**Corollary 5.2** ( $q$ -information friction measurement). *The information friction can be measured from observables as:*

$$T_n = (2-q) \cdot \frac{\sigma_n^2}{\chi_n} \quad (24)$$

*For  $q < 1$  (complements),  $2-q > 1$ , so  $T$  is higher than the Shannon estimate—complementary sectors appear noisier than Shannon predicts.*

### 5.2 $q$ -Covariance Structure

**Theorem 5.3** ( $q$ -covariance). *At  $q$ -exponential equilibrium of a CES( $\rho$ ) economy with  $J$  symmetric inputs, the covariance matrix has the permutation  $(J-1, 1)$  structure:*

$$\Sigma = \sigma^2 \mathbf{I} + \gamma \mathbf{1} \mathbf{1}^\top \quad (25)$$

with eigenvalues:

$$\lambda_1 = \frac{T}{(2-q)K_{\text{eff}}} \quad (1\text{-dimensional, along } \mathbf{1}) \quad (26)$$

$$\lambda_{2,\dots,J} = \frac{T}{(2-q)(K_{\text{eff}} + J^{-1})} \quad ((J-1)\text{-dimensional}) \quad (27)$$

where  $K_{\text{eff}} = K \cdot (1 - T/T^*)^+$  is the effective curvature.

*Proof.* The permutation symmetry of the CES aggregate at the symmetric equilibrium forces the  $(J-1, 1)$  eigenstructure: any covariance matrix invariant under the symmetric group  $\mathfrak{S}_J$  must have the form  $\Sigma = \sigma^2 \mathbf{I} + \gamma \mathbf{1}\mathbf{1}^\top$ , with eigenvalue  $\sigma^2 + J\gamma$  on  $\text{span}\{\mathbf{1}\}$  (multiplicity 1) and eigenvalue  $\sigma^2$  on  $\mathbf{1}^\perp$  (multiplicity  $J-1$ ). This depends only on the symmetry group, not on the specific entropy.

The eigenvalues are computed from  $\Sigma = (T/(2-q))(\nabla^2 \mathcal{F}_q|_{\text{eq}})^{-1}$ , where the  $q$ -variance-response identity (Theorem 5.1) provides the proportionality constant  $T/(2-q)$ . The Hessian of  $\mathcal{F}_q$  at the symmetric equilibrium inherits the CES Hessian structure: eigenvalue  $K_{\text{eff}}$  with multiplicity  $J-1$  on  $\mathbf{1}^\perp$  (from the production curvature) and eigenvalue  $K_{\text{eff}} + J^{-1}$  with multiplicity 1 on  $\text{span}\{\mathbf{1}\}$  (the additional  $J^{-1}$  arises from the entropy's contribution to the budget constraint direction).  $\square$

**Remark 5.4** (Observational consequence). *The ratio of eigenvalues  $\lambda_1/\lambda_{2,\dots,J} = (K_{\text{eff}} + J^{-1})/K_{\text{eff}} = 1 + 1/(JK_{\text{eff}})$  provides a direct estimate of effective curvature from observed cross-sectional covariances:  $K_{\text{eff}} = 1/(J(\lambda_1/\lambda_2 - 1))$ . The estimate is independent of  $T$  and  $q$ —the correction factor  $1/(2-q)$  cancels in the ratio. This makes  $K_{\text{eff}}$  estimable from principal component analysis of sectoral fluctuations without knowledge of the information friction.*

### 5.3 $q$ -Crooks Fluctuation Theorem and Jarzynski Equality

**Theorem 5.5** ( $q$ -Crooks). *For a transition between states  $A$  and  $B$  of the Tsallis CES potential, the ratio of forward to reverse path probabilities satisfies:*

$$\frac{P_F(W)}{P_R(-W)} = \exp_q \left( \frac{W - \Delta \mathcal{F}_q}{T} \right) \quad (28)$$

where  $W$  is the work performed and  $\Delta \mathcal{F}_q = \mathcal{F}_q(B) - \mathcal{F}_q(A)$ .

*Proof.* The  $q$ -exponential equilibrium distribution  $p^* \propto \exp_q(-\beta \varepsilon)$  satisfies the detailed balance condition with  $q$ -exponential weight ratios rather than exponential ones. For a driven process between states  $A$  and  $B$ , the ratio of path probabilities under forward and

reverse protocols is:

$$\frac{P_F[\gamma]}{P_R[\bar{\gamma}]} = \frac{\exp_q(-\beta\varepsilon_A)}{\exp_q(-\beta\varepsilon_B)} = \exp_q(\beta(\varepsilon_B - \varepsilon_A))$$

where  $\bar{\gamma}$  is the time-reversed path. Averaging over paths with fixed work  $W$  and using the  $q$ -algebra identity  $\exp_q(a)/\exp_q(b) = \exp_q(a - b + (1 - q)ab)$ , which simplifies at the path-integral level: the work  $W$  along the path satisfies  $W = \varepsilon_B - \varepsilon_A + W_{\text{diss}}$ . The CES potential difference absorbs the equilibrium contribution, yielding (28).  $\square$

**Remark 5.6** (Economic interpretation). *The  $q$ -Crooks theorem governs the probability of “favorable” vs. “unfavorable” transitions in economic restructuring. For a firm transitioning from organizational mode  $A$  to mode  $B$ , the work  $W$  is the total resource cost of the transition, and  $\Delta\mathcal{F}_q$  is the CES potential difference (the long-run efficiency gain). The theorem states that favorable transitions ( $W < \Delta\mathcal{F}_q$ ) are exponentially more likely than unfavorable ones—but with  $q$ -exponential rather than exponential odds. For complements ( $q < 1$ ), the  $q$ -exponential has compact support, meaning there is a maximum possible transition cost beyond which the transition simply does not occur. This matches the observation that firms with highly complementary production face hard ceilings on restructuring costs.*

**Corollary 5.7** ( $q$ -Jarzynski equality). *Integrating the  $q$ -Crooks relation over work values:*

$$\langle \exp_q(-\beta W) \rangle = \exp_q(-\beta \Delta\mathcal{F}_q) \quad (29)$$

*The minimum expected work for any transition is  $\langle W \rangle \geq \Delta\mathcal{F}_q$ , with equality for quasi-static processes.*

## 5.4 The $q$ -Log-Sum-Exp Dual

**Proposition 5.8** ( $q$ -log-sum-exp). *The Legendre dual of the Tsallis CES potential is:*

$$\mathcal{F}_q^*(\boldsymbol{\varepsilon}) = -T \cdot \ln_q \left( \sum_{j=1}^J \exp_q(-\varepsilon_j/T) \right) \quad (30)$$

*which reduces to the standard log-sum-exp as  $q \rightarrow 1$ .*

*Proof.* The conjugate of  $\mathcal{F}_q = \sum p_j \varepsilon_j - TS_q$  is computed by optimizing over  $\mathbf{p}$ :

$$\mathcal{F}_q^* = \min_{\mathbf{p} \in \Delta} \left[ \sum p_j \varepsilon_j - T \cdot \frac{1 - \sum p_j^q}{q - 1} \right]$$

Substituting the  $q$ -exponential optimum (13) and using the identity  $\sum_j (\exp_q(-\beta\varepsilon_j))^q = Z_q^q \sum_j (p_j^*)^q$  yields (30) after algebraic simplification.  $\square$

**Remark 5.9** (Escort distributions). *The optimization naturally introduces the escort distribution  $P_j = p_j^q / \sum_k p_k^q$ , which reweights probabilities by their  $q$ -th power. For complements ( $q < 1$ ), the escort distribution is more uniform than the original—emphasizing rare inputs. For substitutes ( $q > 1$ ), it is more concentrated—emphasizing common inputs. The escort distribution appears in the  $q$ -expectation  $\langle \varepsilon \rangle_q = \sum P_j \varepsilon_j$ , which replaces the standard expectation in all  $q$ -thermodynamic identities. The  $q$ -internal energy is  $U_q = \langle \varepsilon \rangle_q$ , and the CES potential satisfies  $\mathcal{F}_q = U_q - TS_q$  with the escort-averaged energy.*

## 5.5 $q$ -Dissipation Bound

**Corollary 5.10** ( $q$ -dissipation bound). *The expected dissipated work  $W_{\text{diss}} = \langle W \rangle - \Delta \mathcal{F}_q$  in any transition between CES potential equilibria satisfies:*

$$W_{\text{diss}} \geq \frac{T}{2(2-q)} \cdot \frac{\text{Var}(W)}{T^2} \quad (31)$$

*For complements ( $q < 1$ ), the bound is tighter: less dissipation is tolerated for a given variance of outcomes, reflecting the rigidity of complementary systems.*

*Proof.* From the  $q$ -Jarzynski equality (Corollary 5.7), Jensen’s inequality applied to the convex function  $\exp_q(-\cdot)$  gives  $\langle W \rangle \geq \Delta \mathcal{F}_q$ . The tightened bound follows from a second-order expansion of  $\exp_q$  around the mean work:  $\langle \exp_q(-\beta W) \rangle \approx \exp_q(-\beta \langle W \rangle) + \frac{1}{2} \exp''_q(-\beta \langle W \rangle) \cdot \text{Var}(W)$ . Since  $\exp''_q(x) = -\beta^2 q \exp_q(x)^{2q-1}$  at the symmetric point, evaluating at equilibrium gives the factor  $1/(2(2-q))$ .  $\square$

**Remark 5.11.** *The  $q$ -dissipation bound has direct economic content. Consider a firm restructuring from governance mode A to mode B (e.g., vertical integration to outsourcing). The dissipated work is the transition cost—resources consumed during restructuring that do not contribute to final output. The bound says that complementary production systems ( $q < 1$ ) tolerate less restructuring waste for a given level of outcome uncertainty. This explains why restructuring complementary organizations (e.g., merging research labs) is more costly than restructuring substitutable operations (e.g., consolidating distribution centers).*

## 5.6 Euler Identity Survival

**Proposition 5.12** (Euler identity survives). *At any equilibrium of any CES economy with any  $q$ -entropy, the Euler-equilibrium identity holds exactly:*

$$\mathbf{x}^* \cdot \nabla S_q = -\frac{1}{T} \quad (32)$$

This follows from homogeneity of the CES aggregate and the first-order condition—not the specific functional form of the entropy.

*Proof.* The Euler identity follows from homogeneity of the CES aggregate and the first-order condition  $\nabla\Phi = T\nabla S_q$  at equilibrium. Since  $\Phi = -\log F$  with  $F$  homogeneous of degree 1, Euler's theorem gives  $\mathbf{x} \cdot \nabla\Phi = -1$ . At equilibrium,  $\nabla\Phi = T\nabla S_q$ , so  $\mathbf{x}^* \cdot T\nabla S_q = -1$ , yielding (32). The argument uses only homogeneity and the first-order condition—not the specific functional form of the entropy.  $\square$

**Remark 5.13** (Winding number and Casimir invariants). *By the same logic, the winding number of economic trajectories around the critical curve  $T^*(\rho)$  is a topological invariant that depends on the homotopy class of the trajectory, not on the entropy functional. Similarly, the Casimir invariants of the conservative-dissipative formulation live in the kernel of the antisymmetric coupling matrix  $\mathbf{J}$ , which is determined by the trade network topology, not by the entropy. Both survive the Tsallis generalization exactly.*

## 5.7 $q$ -Onsager Reciprocal Relations

The Onsager reciprocal relations state that the matrix of transport coefficients relating fluxes to forces is symmetric:  $L_{ij} = L_{ji}$ . In the Shannon framework, this follows from time-reversal symmetry at equilibrium. The Tsallis generalization preserves the symmetry but inflates the coefficients.

**Proposition 5.14** ( $q$ -Onsager relations). *At Tsallis CES potential equilibrium, the transport matrix  $\mathbf{L}^{(q)}$  relating economic fluxes  $\mathbf{J} = (J_1, \dots, J_N)$  to forces  $\mathbf{X} = (X_1, \dots, X_N)$  via  $J_i = \sum_j L_{ij}^{(q)} X_j$  satisfies:*

- (i) **Symmetry:**  $L_{ij}^{(q)} = L_{ji}^{(q)}$  (Onsager reciprocity survives exactly).
- (ii) **Inflation:**  $L_{ij}^{(q)} = L_{ij}^{(1)} / (2 - q)$ , where  $L_{ij}^{(1)}$  is the Shannon transport coefficient.
- (iii) **Positivity:**  $\mathbf{L}^{(q)}$  is positive semi-definite (the second law is preserved).

*Proof.* Symmetry follows from the time-reversal argument applied to the  $q$ -exponential equilibrium: the path probability ratio  $P_F[\gamma]/P_R[\bar{\gamma}] = \exp_q(\beta(W_F - W_R))$  is symmetric under exchange of  $i$  and  $j$  at equilibrium (where  $W_F = W_R$ ). The inflation factor  $1/(2-q)$  enters because the transport coefficients are proportional to the equilibrium covariance ( $L_{ij} \propto \Sigma_{ij}$ ), which acquires the  $1/(2-q)$  correction from the  $q$ -variance-response identity. Positivity follows from the positive semi-definiteness of  $\Sigma$ .  $\square$

**Remark 5.15** (Economic content). *In the economic context, the Onsager relations govern the cross-effects between sectors: how a price shock in sector  $i$  affects output in sector  $j$ . The symmetry condition  $L_{ij}^{(q)} = L_{ji}^{(q)}$  means that the effect of sector  $i$  on sector  $j$  equals the*

reverse effect, at equilibrium. The inflation factor  $1/(2 - q)$  means that complementary economies ( $q < 1$ ) have stronger cross-sector linkages than Shannon predicts: a shock in one sector propagates more strongly to others because of the non-additive information structure.

## 5.8 $q$ -Kramers Escape Rate

The Kramers escape rate governs the speed of first-order regime shifts—the rate at which an economy trapped in a suboptimal equilibrium can tunnel to the efficient one.

**Proposition 5.16** ( $q$ -Kramers escape). *The escape rate from a metastable equilibrium  $A$  over a CES potential barrier  $\Delta\mathcal{F}_q$  is:*

$$k_q = \nu \cdot \exp_q \left( -\frac{\Delta\mathcal{F}_q}{T} \right) \quad (33)$$

where  $\nu$  is the attempt frequency (proportional to the curvature at the metastable minimum).

**Remark 5.17.** For complements ( $q < 1$ ), the  $q$ -exponential has compact support: if  $\Delta\mathcal{F}_q > T/(1-q)$ , the escape rate is exactly zero. The economy is locked in the metastable state permanently—no amount of waiting produces a transition. This is sharper than the exponential case ( $q = 1$ ), where escape is merely exponentially slow. Economically: institutions with strongly complementary components (e.g., a vertically integrated firm with highly specific assets) cannot restructure gradually; they require a discrete intervention (policy change, technology shock) that lowers the barrier below the critical threshold  $T/(1 - q)$ .

## 5.9 Classification of Results

Every result from the Shannon framework falls into exactly one of three categories under the Tsallis generalization.

**Why the three categories exist.** The classification reflects the mathematical origin of each result:

**Category A** results depend only on the Legendre structure of the CES potential—homogeneity, topology, and conservative-dissipative structure are properties of  $\Phi$  and the network coupling, not the entropy. These survive any entropy generalization, not just Shannon→Tsallis.

**Category B** results depend on the curvature of the entropy at equilibrium. The Tsallis entropy has curvature  $qp^{q-2}$  versus Shannon's  $1/p$ , differing by a factor that eval-

Table 1: Classification of Shannon-framework results under Tsallis generalization

Category	Result	Shannon form	Tsallis form
A: Exact	Euler identity	$\mathbf{x}^* \cdot \nabla H = -1/T$	$\mathbf{x}^* \cdot \nabla S_q = -1/T$
	Winding number	$w \in \mathbb{Z}$	$w \in \mathbb{Z}$ (unchanged)
	Casimir invariants	$\ker \mathbf{J}$	$\ker \mathbf{J}$ (unchanged)
	Pre-crisis deceleration	$\tau \sim  T - T^* ^{-1}$	$\tau \sim  T - T^* ^{-1}$ (unchanged)
B: $q$ -corrected	Variance-response	$\Sigma = T\chi$	$\Sigma = (T/(2-q))\chi$
	Covariance eigenvalues	$T/K_{\text{eff}}$	$T/((2-q)K_{\text{eff}})$
	Kramers escape	$k \propto e^{-\Delta\mathcal{F}/T}$	$k \propto \exp_q(-\Delta\mathcal{F}_q/T)$
	Onsager coefficients	$L_{ij} = L_{ji}$	$L_{ij}^{(q)} = L_{ji}^{(q)}$ (inflated)
C: Structural	Crooks theorem	$P_F/P_R = e^{(W-\Delta\mathcal{F})/T}$	$P_F/P_R = \exp_q((W - \Delta\mathcal{F}_q)/T)$
	Jarzynski equality	$\langle e^{-\beta W} \rangle = e^{-\beta\Delta\mathcal{F}}$	$\langle \exp_q(-\beta W) \rangle = \exp_q(-\beta\Delta\mathcal{F}_q)$
	Equilibrium dist.	$p^* \propto e^{-\beta\varepsilon}$	$p^* \propto \exp_q(-\beta\varepsilon)$
	Log-sum-exp	$-T \log \sum e^{-\varepsilon/T}$	$-T \ln_q \sum \exp_q(-\varepsilon/T)$

uates to  $1/(2 - q)$  at symmetric equilibrium. This multiplicative correction propagates to all second-order quantities: fluctuations, response, transport coefficients.

**Category C** results depend on the *functional form* of the equilibrium distribution. Replacing  $\exp$  with  $\exp_q$  changes the shape of the distribution—from exponential tails to power-law tails (or compact support)—affecting all results that reference the distribution directly.

**Remark 5.18** (The  $1/(2 - q)$  correction factor). *The correction factor  $1/(2 - q)$  has a unified interpretation. For complements ( $q < 1$ ),  $1/(2-q) > 1$ : fluctuations are amplified relative to Shannon. Complementary inputs interact, so a shock to one propagates to all others, amplifying the total variance. For substitutes ( $q > 1$ ),  $1/(2 - q) < 1$ : fluctuations are dampened. Substitutable inputs absorb shocks independently, reducing total variance. At  $q = 1$  (Cobb-Douglas),  $1/(2 - q) = 1$  and Shannon is recovered exactly.*

*The magnitude of the correction is:*

$$\frac{1}{2 - q} - 1 = \frac{q - 1}{2 - q} = \frac{\rho - 1}{2 - \rho} \quad (34)$$

*For typical manufacturing values ( $\sigma \approx 3-5$ , i.e.,  $\rho \approx 0.67-0.80$ ), the correction is  $-25\%$  to  $-17\%$ —a modest but systematic effect. For highly complementary production ( $\rho \approx -2$ , e.g., semiconductor fabrication), the correction is  $+75\%$ : Shannon-based models underestimate fluctuations by almost a factor of two in these sectors.*

**Remark 5.19** (Practical implications of the classification). *The classification has immediate practical consequences:*

1. **For estimation:** Category A results can be estimated using Shannon-based methods without correction. Category B results require the  $1/(2 - q)$  adjustment, which is

computable from  $\rho$  alone. Category C results require fitting the full  $q$ -exponential distribution.

2. **For early warning:** Pre-crisis deceleration (Table 1, Category A) holds exactly under Tsallis: the relaxation time  $\tau \sim |T - T^*|^{-1}$  diverges as  $T \rightarrow T^*$ , regardless of  $q$ . This means standard early-warning signals (increased autocorrelation, increased variance near the critical point) remain valid without  $q$ -correction.
3. **For policy:** Institutional design based on the CES potential can use Category A results (topology, conservation laws) confidently. Calibration exercises using variance-based estimates (Category B) must include the  $1/(2-q)$  correction to avoid systematic bias.

## 6 Derivation I: Akerlof's Market for Lemons

The market-for-lemons result (Akerlof, 1970) is derived from the CES potential framework.

### 6.1 Setup

Consider  $J$  sellers with qualities  $q_j$  drawn from a distribution  $F$  on  $[0, \bar{q}]$ . Each seller knows their own quality. A representative buyer values quality at rate  $v > 1$  but cannot observe quality directly. The CES aggregate of market quality is  $\Phi(S) = (|S|^{-1} \sum_{j \in S} q_j^\rho)^{1/\rho}$ . By Proposition 2.2(b), this CES form is derived from Fréchet taste heterogeneity.

Under rational inattention, the buyer acquires a binary signal (accept/reject) at information cost  $T \cdot h(q^*/\bar{q})$ , where  $h(\cdot)$  is binary entropy. Define  $\psi(\rho) = (\rho + 1)^{1/\rho}$  and  $A(\rho) = v/\psi(\rho) - 1$  (the CES net value rate). The buyer's CES potential is:

$$\max_{q^* \in [0, \bar{q}]} \mathcal{F}(q^*) = \frac{A(\rho)}{\bar{q}} (q^*)^2 - T \cdot h\left(\frac{q^*}{\bar{q}}\right) \quad (35)$$

### 6.2 The $\rho$ -Dependent Robustness Result

**Proposition 6.1** (Market Robustness). *There exists a unique critical information friction*

$$T^* = \frac{A(\rho) \bar{q}}{\gamma^*} = \frac{(v - \psi(\rho)) \bar{q}}{\gamma^* \psi(\rho)} \quad (36)$$

where  $\gamma^* \approx 2.456$  is a universal constant, such that for  $T < T^*$  the CES potential is positive at all interior participation rates; for  $T > T^*$  the CES potential develops a negative barrier, forcing a first-order regime shift.

Since  $\psi(\rho)$  is strictly decreasing in  $\rho$ ,  $T^*$  is strictly increasing in  $(1-\rho)$ : markets with lower  $\rho$  (higher complementarity) are more robust to adverse selection.

*Proof.* The proof proceeds in six steps: (1) state the Lagrangian; (2) derive first-order conditions; (3) verify the second-order condition; (4) solve for the breakdown threshold; (5) connect to CES curvature  $K$ ; (6) prove monotonicity.

*Step 1: The buyer's Lagrangian.* The buyer solves the rational inattention problem in its Lagrangian dual form. Following Sims (2003), an agent who maximizes expected payoff subject to an information constraint  $I \leq \kappa$  has the equivalent unconstrained (penalized) formulation:

$$\max_{q^* \in [0, \bar{q}]} \mathcal{F}(q^*) = \frac{A(\rho)}{\bar{q}}(q^*)^2 - T \cdot h\left(\frac{q^*}{\bar{q}}\right) \quad (37)$$

where  $A(\rho) = v/\psi(\rho) - 1 > 0$  is the CES net value rate,  $T \geq 0$  is the Lagrange multiplier on the information constraint (equivalently, the unit cost of information), and  $h(\pi) = -\pi \log \pi - (1 - \pi) \log(1 - \pi)$  is the binary entropy function with  $\pi = q^*/\bar{q}$ .

The first term is the expected gains from trade: the participation rate  $\pi$  times the per-trade net surplus  $A(\rho)q^*$ . The second term is the information cost of the binary screening signal. The objective  $\mathcal{F}$  is the CES potential: expected surplus minus information cost.

*Step 2: First-order condition.* Substituting  $\pi = q^*/\bar{q}$ , write  $\mathcal{F}$  in terms of  $\pi$ :

$$\mathcal{F}(\pi) = A(\rho)\bar{q}\pi^2 - Th(\pi)$$

The derivative with respect to  $\pi$  is:

$$\frac{d\mathcal{F}}{d\pi} = 2A(\rho)\bar{q}\pi - T \log \frac{1 - \pi}{\pi}$$

using  $dh/d\pi = \log[(1 - \pi)/\pi]$ . Setting  $d\mathcal{F}/d\pi = 0$ , the first-order condition is:

$$2A(\rho)\bar{q}\pi = T \log \frac{1 - \pi}{\pi} \quad (38)$$

This equates the marginal expected surplus from expanding the participating set (left side) with the marginal information cost (right side).

*Step 3: Second-order condition.* The second derivative is:

$$\frac{d^2\mathcal{F}}{d\pi^2} = 2A(\rho)\bar{q} + \frac{T}{\pi(1 - \pi)}$$

Since both terms are positive,  $d^2\mathcal{F}/d\pi^2 > 0$  for all  $\pi \in (0, 1)$ : the CES potential is strictly convex. Any interior critical point satisfying (38) is therefore a *minimum*, not a maximum. The boundary values  $\mathcal{F}(0) = 0$  and  $\mathcal{F}(1) = A(\rho)\bar{q} > 0$  show that  $\pi = 1$  (full participation without screening) always yields positive CES potential.

The economic content emerges from the *value* of  $\mathcal{F}$  at the interior minimum  $\pi_0$ . For small  $T$ ,  $\mathcal{F}(\pi_0) > 0$ : the CES potential is positive everywhere on  $(0, 1)$ , so the buyer

can smoothly adjust the screening threshold—information acquisition is valuable and the market functions efficiently. For large  $T$ ,  $\mathcal{F}(\pi_0) < 0$ : the CES potential dips below zero near  $\pi_0$ , creating a barrier between  $\pi = 0$  (no market) and  $\pi = 1$  (unscreened pooling). The market undergoes a *first-order regime shift*: screening intensity jumps discontinuously from the efficient interior level to a boundary solution.

*Step 4: Solving for the breakdown threshold.* Define the dimensionless ratio  $\tau \equiv T/(A(\rho)\bar{q})$ . In terms of  $\tau$ , the CES potential becomes:

$$\frac{\mathcal{F}(\pi)}{A(\rho)\bar{q}} = \pi^2 - \tau h(\pi) \equiv \Psi(\pi, \tau)$$

The first-order condition is  $2\pi = \tau \log[(1 - \pi)/\pi]$ , and the breakdown threshold  $\tau^*$  is the value of  $\tau$  at which the interior minimum of  $\Psi$  first touches zero.

At the critical point, two conditions hold simultaneously:

$$(\text{FOC}): \quad 2\pi_0 = \tau^* \log \frac{1 - \pi_0}{\pi_0} \tag{39}$$

$$(\text{Zero surplus}): \quad \pi_0^2 = \tau^* h(\pi_0) \tag{40}$$

Dividing (39) by (40):

$$\frac{2}{\pi_0} = \frac{\log \frac{1 - \pi_0}{\pi_0}}{h(\pi_0)}$$

This is a single transcendental equation in  $\pi_0$  alone, independent of all model parameters ( $v, \rho, \bar{q}, J$ ). By intermediate value theorem arguments (analyzing boundary behavior as  $\pi \rightarrow 0^+$  and  $\pi \rightarrow 1/2$ ), a unique root exists at  $\pi_0 \approx 0.217$ .

Once  $\pi_0$  is determined, the critical dimensionless information friction is:

$$\tau^* = \frac{\pi_0^2}{h(\pi_0)} \equiv \frac{1}{\gamma^*}$$

defining the universal constant  $\gamma^* = h(\pi_0)/\pi_0^2 \approx 2.456$  (using natural logarithms). The breakdown threshold in original units is:

$$T^* = \tau^* \cdot A(\rho) \bar{q} = \frac{A(\rho) \bar{q}}{\gamma^*} = \frac{\bar{q}}{\gamma^*} \left( \frac{v}{\psi(\rho)} - 1 \right)$$

This is the closed-form expression (36). Crucially,  $\gamma^*$  is a universal constant—it depends on neither  $\rho$ ,  $v$ ,  $\bar{q}$ , nor  $J$ . All dependence of  $T^*$  on the CES parameter enters through  $A(\rho) = v/\psi(\rho) - 1$ .

*Step 5: Connection to CES curvature  $K$ .* Since  $T^* = (v/\psi(\rho) - 1)\bar{q}/\gamma^*$  and  $\psi(\rho) = (\rho + 1)^{1/\rho}$ , the dependence on  $\rho$  enters through  $1/\psi(\rho)$ . To connect with  $K = (1 - \rho)(J - 1)/J$ ,

expand  $\log \psi$  around  $\rho = 1$  (the substitutes limit,  $K = 0$ ):

$$\log \psi(1) = \log 2, \quad \frac{d}{d\rho} \log \psi \Big|_{\rho=1} = \frac{1}{2} - \log 2 \approx -0.193$$

Therefore  $\psi(\rho) = 2 \exp[(1/2 - \log 2)(1 - \rho) + O((1 - \rho)^2)]$  and since  $1 - \rho = KJ/(J - 1)$ :

$$T^* = \frac{\bar{q}}{\gamma^*} \left[ \left( \frac{v}{2} - 1 \right) + \frac{v}{2} \left( \log 2 - \frac{1}{2} \right) \frac{KJ}{J-1} + O(K^2) \right]$$

The leading term  $(v/2 - 1)$  is the gains from trade at  $\rho = 1$  (linear aggregation, the standard Akerlof case). The correction proportional to  $K$  shows that CES curvature provides additional robustness to adverse selection.

*Step 6: Monotonicity of  $T^*$  in  $(1 - \rho)$ .* It suffices to show that  $\psi(\rho) = (\rho + 1)^{1/\rho}$  is strictly decreasing in  $\rho$  on  $(-1, 1)$ , since  $T^* = (v/\psi(\rho) - 1)\bar{q}/\gamma^*$  and  $v/\psi - 1$  is strictly increasing in  $1/\psi$  for  $v > \psi > 0$ .

The sign of  $d \log \psi / d\rho$  equals the sign of  $g(\rho) \equiv \rho/(\rho + 1) - \log(\rho + 1)$ . Observe  $g(0) = 0$ , and:

$$g'(\rho) = \frac{1}{(\rho + 1)^2} - \frac{1}{\rho + 1} = \frac{-\rho}{(\rho + 1)^2}$$

For  $\rho \in (-1, 0)$ :  $g'(\rho) > 0$ , so  $g$  is strictly increasing toward  $g(0) = 0$ , hence  $g(\rho) < 0$ . For  $\rho \in (0, 1)$ :  $g'(\rho) < 0$ , so  $g$  is strictly decreasing from  $g(0) = 0$ , hence  $g(\rho) < 0$ .

Therefore  $d \log \psi / d\rho < 0$  for all  $\rho \in (-1, 1) \setminus \{0\}$ , which means  $\psi(\rho)$  is strictly decreasing in  $\rho$ , so  $1/\psi(\rho)$  is strictly increasing. Since  $T^* \propto v/\psi(\rho) - 1$  and  $v/\psi$  is increasing as  $\rho$  decreases,  $T^*$  is strictly increasing in  $(1 - \rho)$ . Lower  $\rho$  (higher complementarity, higher CES curvature  $K$ ) implies higher breakdown threshold, and hence greater robustness to adverse selection.  $\square$

**Remark 6.2.** *The standard Akerlof result is the special case  $T \rightarrow \infty$  (or equivalently, the case where information acquisition is impossible). The  $\rho$ -dependent extension reveals that adverse selection severity is not solely a property of information asymmetry—it depends on the complementarity structure of the market.*

**Remark 6.3** (Connection to market microstructure). *The CES potential framework connects adverse selection to market microstructure. The bid-ask spread in financial markets reflects the market maker's information disadvantage—a standard adverse selection cost. In the CES potential framework, the spread is:*

$$\text{Spread} \propto \frac{T}{T^*(\rho)} \cdot \frac{1}{J_{\text{eff}}}$$

where  $J_{\text{eff}}$  is the effective number of informed traders. The spread narrows as  $T/T^*$  falls (better information quality) and as  $J_{\text{eff}}$  rises (more informed traders competing away

rents). For commodity-like assets ( $\rho \approx 1$ , low  $T^*$ ): spreads are high relative to the information advantage because the breakdown threshold is low. For unique, complementary assets (low  $\rho$ , high  $T^*$ ): the same information advantage produces a smaller spread because the market's higher complementarity value absorbs the information cost. This predicts that bid-ask spreads should be positively correlated with asset substitutability after controlling for trading volume—a testable prediction using cross-sectional data from equity and commodity markets.

**Example 6.4** (Numerical illustration). Let  $v = 3$  (buyers value quality at 3x cost),  $\bar{q} = 1$ , and compare three markets:

Market	$\rho$	$\psi(\rho)$	$A(\rho)$	$T^* = A\bar{q}/\gamma^*$
Commodity ( $\rho = 0.8$ )	0.8	1.81	0.66	0.27
Differentiated ( $\rho = 0$ )	0	1.00	2.00	0.81
Complementary ( $\rho = -2$ )	-2	0.33	8.09	3.30

The complementary market tolerates information friction 12 times higher than the commodity market before collapsing. At  $T = 0.5$  (moderate information friction): the commodity market has collapsed ( $T > T^* = 0.27$ ), the differentiated market is stressed ( $T/T^* = 0.62$ ), and the complementary market functions smoothly ( $T/T^* = 0.15$ ). The prediction: at the same level of information asymmetry, commodity markets unravel while complementary markets continue to function.

### 6.3 The Informational Interpretation

The CES potential  $\mathcal{F}(q^*) = A(\rho)\bar{q}\pi^2 - T \cdot h(\pi)$  has a clean information-theoretic decomposition. The first term is the *expected gains from trade*: the CES value of the participating set, quadratic in the participation rate because both the number of trades and the average quality in the participating set increase with  $\pi$ . The second term is the *information cost*: the minimum number of bits (times  $T$ ) needed to implement a binary screening signal with acceptance rate  $\pi$ .

The ratio  $A(\rho)\bar{q}/T$  is the *signal-to-noise ratio* of the market: how much CES value each bit of screening information generates. When this ratio is high (complementary goods, low information costs), the market sustains a precise screening equilibrium. When it is low (substitutable goods, high information costs), the market cannot justify the information expenditure and collapses to a pooling equilibrium.

The universal constant  $\gamma^* \approx 2.456$  has a natural interpretation: it is the maximum *information leverage*—the maximum ratio of surplus destroyed to information cost saved—at the critical participation rate. This maximum is a property of the binary entropy function and does not depend on the economic parameters.

## 6.4 Empirical Implications

Proposition 6.1 yields testable predictions:

1. **Specialized labor markets** (low  $\rho$ ): should function despite severe information asymmetry about talent. *Observed:* elaborate screening (headhunters, multi-round interviews, trial periods) is sustained because the CES value of identifying the right specialist justifies the information cost.
2. **Commodity markets** (high  $\rho$ ): should be susceptible to lemons problems, resolved by institutions that reduce  $H$  directly (grading, standardization, certification) rather than by costly individual evaluation. *Observed:* USDA grading, commodity exchange standards, ISO certification.
3. **Art and collectibles** (very low  $\rho$ ): each piece is unique, near-maximum complementarity. Should sustain very expensive information acquisition. *Observed:* expert appraisals, provenance research, and authentication services costing significant fractions of the item's value.
4. **Used cars** (moderate  $\rho$ ): Akerlof's original domain. Markets partially function through intermediate information institutions (CarFax, dealer certification, warranties).

The cross-sectional prediction is sharp: regress measures of adverse selection severity (bid-ask spreads, return rates, warranty costs) against estimated  $\rho$  for each market. The framework predicts a positive relationship.

**Remark 6.5** (Institutional responses to adverse selection). *The CES potential framework classifies institutional responses to adverse selection by where they operate in  $(\rho, T)$  space:*

- **Reducing  $T$  (information institutions):** certification, grading, disclosure requirements, warranties. These shift the economy leftward in  $(\rho, T)$  space, increasing the distance from the breakdown threshold. Examples: Moody's ratings (reduce  $T$  for bond markets), USDA grading (reduce  $T$  for agricultural markets), Carfax (reduce  $T$  for used car markets).
- **Increasing effective  $\rho$  (market design):** standardization, commoditization, index products. These shift the economy upward in  $(\rho, T)$  space, reducing  $K$  and hence the sensitivity to information friction. Example: commodity futures contracts that standardize quality, delivery, and quantity, transforming a differentiated good (wheat from farm X) into a substitutable one (CBOT wheat contract).

- **Reducing  $\rho$  (bundling):** package deals that make goods more complementary, increasing  $T^*$  and the economy's robustness to information friction. Example: HMO health insurance bundles primary care, specialist care, and hospital services into a complementary package that is more robust to adverse selection than separate markets for each service.

The framework predicts which response is optimal: reduce  $T$  when information technology is available and the market has high intrinsic complementarity; increase  $\rho$  when the good is naturally commodity-like and information costs are irreducible; reduce  $\rho$  (bundle) when goods are naturally substitutable but the information environment is poor.

## 6.5 Dynamic Extension: Bistability

**Proposition 6.6** (Bistability of Market Information). *Let  $T(t)$  evolve as  $\dot{T} = \mu - \gamma \cdot \max(\Phi(T), 0)$ . Then  $T^*(\rho)$  is an unstable fixed point separating a functioning low- $T$  equilibrium (information production offsets depreciation) from a collapsed high- $T$  equilibrium. Low- $\rho$  markets have a larger functioning-equilibrium basin.*

*Proof.* The dynamics of information friction are governed by the balance between depreciation (at rate  $\mu > 0$ , reflecting the natural decay of market knowledge through agent turnover, product innovation, and regulatory change) and endogenous production (at rate  $\gamma \cdot \Phi(T)$ , where  $\Phi(T) = \max(\mathcal{F}(T), 0)$  is the positive part of the CES potential evaluated at the current participation rate).

*Step 1: Fixed points.* Setting  $\dot{T} = 0$  yields  $\mu = \gamma \cdot \Phi(T)$ . From Proposition 6.1,  $\Phi(T)$  is a hump-shaped function:  $\Phi(0) = 0$  (no surplus without participation),  $\Phi(T)$  rises to a maximum at some  $T_{\text{peak}}$ , and  $\Phi(T) = 0$  for  $T \geq T^*$ . The equation  $\mu/\gamma = \Phi(T)$  therefore has either zero or two solutions. When  $\mu/\gamma < \Phi(T_{\text{peak}})$ , there exist two fixed points:  $T_L < T_{\text{peak}} < T_H$ , plus the boundary fixed point  $T = T^*$  where  $\Phi = 0$ .

*Step 2: Stability.* Linearizing around each fixed point:  $d\dot{T}/dT = -\gamma\Phi'(T)$ . At  $T_L$ :  $\Phi'(T_L) > 0$  (ascending branch), so  $d\dot{T}/dT < 0$ —stable. At  $T_H$ :  $\Phi'(T_H) < 0$  (descending branch), so  $d\dot{T}/dT > 0$ —unstable. For  $T > T^*$ :  $\Phi(T) = 0$  so  $\dot{T} = \mu > 0$ , driving  $T \rightarrow \infty$  (total market collapse). The threshold  $T^*$  thus separates the basins of attraction of the functioning equilibrium  $T_L$  and the collapsed state.

*Step 3: Basin size depends on  $\rho$ .* The basin of attraction of  $T_L$  is the interval  $[0, T_H]$ . Since  $T^*$  is increasing in  $(1 - \rho)$  (Proposition 6.1), and  $T_H < T^*$ , the basin width  $T_H$  is also increasing in  $(1 - \rho)$ . Complementary markets have a wider functioning-equilibrium basin: a larger information friction shock is needed to tip them into collapse.

*Step 4: Hysteresis.* Once the market collapses ( $T > T^*$ ), restoring it requires reducing  $T$  below  $T_H$ , not merely below  $T^*$ . The hysteresis gap  $T^* - T_H > 0$  means that the

conditions for market revival are strictly more demanding than the conditions that would have prevented collapse. This matches the empirical observation that market institutions, once destroyed, are difficult to rebuild even when the original cause of failure is removed.

□

**Remark 6.7.** *The standard Akerlof result is the static limit:  $T$  is fixed at  $\infty$  (no information acquisition) or at 0 (perfect information). The dynamic extension shows that adverse selection is an absorbing state: once  $T$  crosses  $T^*$ , the market unravels and information production ceases, so  $T$  grows without bound. Markets for complements resist this because their basin of attraction is larger.*

## 7 Derivation II: Myerson's Optimal Mechanism

Myerson's virtual valuation ([Myerson, 1981](#)) is shown to be the CES potential gradient.

### 7.1 The Virtual Valuation as CES Potential Gradient

**Proposition 7.1** (Virtual Valuation Identification). *The virtual valuation  $\varphi(\theta) = \theta - (1 - F(\theta))/f(\theta)$  is the CES potential gradient: the marginal CES contribution  $\theta$  minus the marginal entropy cost  $(1 - F(\theta))/f(\theta)$  (the inverse Mills ratio, measuring residual uncertainty per type).*

*Proof.* The proof proceeds in four steps: formulating the mechanism design problem, deriving the information rent via the IC constraint, identifying the rent as an entropy gradient, and assembling the virtual valuation as the CES potential gradient.

*Step 1: The mechanism design problem.* By the revelation principle, restrict attention to direct mechanisms  $(x(\cdot), t(\cdot))$  where each agent reports a type and receives allocation  $x_j = x(\hat{\theta}_j)$  and transfer  $t_j = t(\hat{\theta}_j)$ . Agent  $j$  with true type  $\theta_j$  values receiving allocation  $x$  at  $\theta_j \cdot x$  (quasi-linear utility). Agent  $j$ 's payoff from reporting  $\hat{\theta}$  when the true type is  $\theta$  is:

$$U(\hat{\theta}, \theta) = \theta \cdot x(\hat{\theta}) - t(\hat{\theta}) \quad (41)$$

Define the equilibrium rent function  $U(\theta) \equiv U(\theta, \theta) = \theta \cdot x(\theta) - t(\theta)$ .

The principal maximizes expected CES welfare net of transfers:

$$\max_{x(\cdot), t(\cdot)} \mathbb{E}_{\theta} \left[ \Phi(x(\theta_1), \dots, x(\theta_J)) + \sum_{j=1}^J t(\theta_j) \right] \quad (42)$$

subject to incentive compatibility (IC:  $U(\theta) \geq U(\hat{\theta}, \theta)$  for all  $\theta, \hat{\theta}$ ) and individual rationality (IR:  $U(\theta) \geq 0$  for all  $\theta$ ).

*Step 2: IC implies the information rent is the integral of the allocation rule.* By standard arguments (Myerson, 1981), IC requires that  $x(\theta)$  is non-decreasing in  $\theta$  and the envelope condition holds. From (41), truth-telling requires  $\theta \in \arg \max_{\hat{\theta}} \theta \cdot x(\hat{\theta}) - t(\hat{\theta})$ . Taking the first-order condition and evaluating at  $\hat{\theta} = \theta$ :  $dU/d\theta = x(\theta)$ . Integrating from the lowest type  $\underline{\theta}$  and using IR binding at  $\underline{\theta}$ :

$$U(\theta) = \int_{\underline{\theta}}^{\theta} x(s) ds \quad (43)$$

Substituting  $t(\theta) = \theta \cdot x(\theta) - U(\theta)$  into (42) and applying Fubini's theorem to reverse the order of integration:

$$\mathbb{E}[t(\theta)] = \int_{\underline{\theta}}^{\bar{\theta}} \left[ \theta - \frac{1 - F(\theta)}{f(\theta)} \right] x(\theta) f(\theta) d\theta = \int_{\underline{\theta}}^{\bar{\theta}} \varphi(\theta) \cdot x(\theta) \cdot f(\theta) d\theta \quad (44)$$

*Step 3: The inverse Mills ratio as entropy gradient.* Define the cumulative residual entropy  $\mathcal{H}(\theta) \equiv - \int_{\underline{\theta}}^{\bar{\theta}} [1 - F(s)] \log[1 - F(s)] ds$ —the continuous analogue of Shannon entropy applied to the survival function  $\bar{F}(\theta) = 1 - F(\theta)$ . The information rent from (43) can be written as  $\mathbb{E}[U(\theta)] = \int x(\theta) [1 - F(\theta)] d\theta$ . The integrand  $[1 - F(\theta)]$  is the hazard weight: the mass of types above  $\theta$  who benefit from the information rent generated at  $\theta$ . Dividing by  $f(\theta)$  gives the per-unit-density rent  $(1 - F(\theta))/f(\theta)$ —the inverse of the hazard rate  $h(\theta) = f(\theta)/(1 - F(\theta))$ . This is the entropy contribution per agent at type  $\theta$ : low-hazard types contribute more residual uncertainty and therefore command higher information rents.

*Step 4: Assembly as the CES potential gradient.* From Step 2, the principal's pointwise objective is:

$$\mathcal{F}(\theta) = \underbrace{\theta \cdot x(\theta)}_{\text{CES marginal contribution } \partial \Phi / \partial \theta} - \underbrace{\frac{1 - F(\theta)}{f(\theta)} \cdot x(\theta)}_{\text{entropy cost } T \cdot \partial \mathcal{H} / \partial \theta}$$

The CES potential gradient with respect to  $\theta$  is  $\partial \mathcal{F} / \partial \theta = \theta - (1 - F(\theta))/f(\theta) = \varphi(\theta)$ , which is Myerson's virtual valuation.

*Explicit computation for the uniform distribution.* Let  $F(\theta) = (\theta - \underline{\theta})/(\bar{\theta} - \underline{\theta})$  on  $[\underline{\theta}, \bar{\theta}]$ . Then  $(1 - F(\theta))/f(\theta) = \bar{\theta} - \theta$ , giving  $\varphi(\theta) = 2\theta - \bar{\theta}$ . This is non-negative iff  $\theta \geq \bar{\theta}/2$ —so the optimal mechanism excludes the bottom half of the type distribution, which is Myerson's classical result.

*Non-uniform distributions.* The CES potential gradient interpretation extends naturally to non-uniform type distributions. Three illustrative cases:

- **Exponential types** ( $f(\theta) = \lambda e^{-\lambda\theta}$ ,  $\theta \geq 0$ ): The inverse Mills ratio is  $(1 - F)/f = 1/\lambda$  (constant), so  $\varphi(\theta) = \theta - 1/\lambda$ . The virtual valuation is a simple downward

shift by the mean type. The exclusion threshold  $\theta^* = 1/\lambda = \mathbb{E}[\theta]$ : the mechanism excludes below-average types. The PoIC depends on  $\rho$  through the CES welfare weight: for low  $\rho$ , excluding any type is costly (Leontief bottleneck), so the principal raises the transfer to include marginal types.

- **Power-law types** ( $f(\theta) = \alpha\theta^{-\alpha-1}$ ,  $\theta \geq 1$ ):  $(1 - F)/f = \theta/\alpha$ , giving  $\varphi(\theta) = \theta(1 - 1/\alpha)$ . For  $\alpha > 1$ : virtual valuation is positive everywhere and proportional to  $\theta$ —no exclusion is optimal. The lack of exclusion arises because the heavy tail means that high types are not uncommon enough to command large rents. This connects to the power-law firm size distributions predicted by the  $q$ -exponential equilibrium for  $q > 1$  (Remark 3.8).
- **Beta types** ( $f(\theta) \propto \theta^{a-1}(1-\theta)^{b-1}$  on  $[0, 1]$ ): The inverse Mills ratio varies smoothly, and the exclusion threshold depends on the shape parameters. For  $a < 1$  (many low types): high exclusion, high PoIC. For  $a > 1$  (few low types): low exclusion, low PoIC. The CES interaction: high  $K$  (low  $\rho$ ) makes exclusion more costly because each type contributes to the complementary bundle, shifting the optimal threshold downward.

□

## 7.2 The Optimal Mechanism Format

The CES curvature determines not only the magnitude of information rents but the optimal *format* of the mechanism.

**Proposition 7.2** (Mechanism Format Determination). *The optimal mechanism format transitions as  $\rho$  decreases:*

- (i)  $\rho$  near 1 (*perfect substitutes*): **Simple auction.** *The principal posts a price and agents self-select. The auction format is optimal because agents are interchangeable—the principal needs only to identify which agent has the lowest cost, not to learn detailed type information.*
- (ii)  $\rho$  moderate (*differentiated inputs*): **Scored evaluation.** *The principal evaluates multiple dimensions with explicit weights. The scoring reflects the CES production structure: each dimension receives weight proportional to its marginal contribution to the CES aggregate.*
- (iii)  $\rho$  near  $-\infty$  (*perfect complements*): **Sequential revelation with holdback.** *The principal reveals partial allocation information to each agent sequentially, exploiting the Leontief structure where each agent's optimal report depends on the reports*

of others. The mechanism must prevent each agent from free-riding on others' revelations.

The mechanism format prediction is testable in procurement data: defense contracts (low  $\rho$ ) should use scored evaluations and negotiated awards more frequently than commodity contracts (high  $\rho$ ), which should use sealed-bid auctions. The USAspending.gov data in Section 14.3 confirm this: 67% of LOW-substitutability contracts use negotiated awards vs. 23% of HIGH-substitutability contracts.

### 7.3 The Revelation Principle as Variational Equilibrium

**Proposition 7.3** (Optimality of Revelation). *A mechanism is incentive-compatible if and only if truth-telling minimizes each agent's CES potential:*

$$\mathcal{F}_j(\text{report } \theta') = -u_j(\theta, x(\theta')) + t(\theta') + T \cdot \text{KL}(\theta' \parallel \theta)$$

where  $\text{KL}(\theta' \parallel \theta)$  is the Kullback-Leibler divergence measuring the entropy cost of reporting  $\theta'$  when the true type is  $\theta$ . The VCG mechanism sets transfers  $t(\cdot)$  so that truth-telling ( $\theta' = \theta$ ,  $\text{KL} = 0$ ) is the global CES potential minimum.

The revelation principle—that any implementable outcome can be achieved through truthful direct revelation—is the economic analogue of the variational statement that equilibrium minimizes the CES potential at the point of maximum order (zero entropy of misrepresentation).

### 7.4 $\rho$ -Dependent Price of Incentive Compatibility

**Proposition 7.4** (Price of Incentive Compatibility). *The welfare ratio  $\text{PoIC}(\rho) = (W^{\text{FB}} - W^{\text{SB}})/W^{\text{FB}}$  is decreasing in  $\rho$ :*

- (a) As  $\rho \rightarrow 1$  (perfect substitutes):  $\text{PoIC} \rightarrow 0$ . Agents are replaceable; information rents vanish.
- (b) As  $\rho \rightarrow -\infty$  (perfect complements):  $\text{PoIC} \rightarrow 1$ . Each agent is essential; information rents consume the entire surplus.

*Proof.* The proof proceeds in four steps: solving the first-best, solving the second-best, computing the welfare ratio, and proving monotonicity with the two limits. For closed forms, work with the uniform distribution  $F(\theta) = (\theta - \underline{\theta})/(\bar{\theta} - \underline{\theta})$  on  $[\underline{\theta}, \bar{\theta}]$  with  $\underline{\theta} > 0$ .

*Step 1: First-best allocation ( $T = 0$ , no IC constraint).* With observable types, the principal solves  $\max_{x(\theta) \geq 0} \mathbb{E}_{\theta}[\theta \cdot x(\theta)^{\rho} - c \cdot x(\theta)]$ . Normalizing  $c = \rho$ , the first-order condition gives:

$$x^{\text{FB}}(\theta) = \theta^{1/(1-\rho)} = \theta^{\sigma} \tag{45}$$

where  $\sigma = 1/(1 - \rho)$ . The first-best per-type surplus is  $(1 - \rho)\theta^\sigma$ , giving:

$$W^{\text{FB}} = (1 - \rho) \cdot \mathbb{E}[\theta^\sigma] \quad (46)$$

*Step 2: Second-best allocation (with IC constraint).* From the virtual valuation identification, the principal maximizes virtual surplus. With the uniform distribution,  $\varphi(\theta) = 2\theta - \bar{\theta}$ . The second-best allocation for types with  $\varphi(\theta) > 0$  (i.e.,  $\theta > \bar{\theta}/2 \equiv \theta^*$ ) is:

$$x^{\text{SB}}(\theta) = \varphi(\theta)^\sigma = (2\theta - \bar{\theta})^\sigma, \quad \theta \geq \theta^* \quad (47)$$

and  $x^{\text{SB}}(\theta) = 0$  for  $\theta < \theta^*$  (exclusion of types with negative virtual valuation).

*Step 3: The welfare ratio.* The welfare loss decomposes into two sources: (i) *exclusion loss*—types  $\theta \in [\underline{\theta}, \theta^*]$  are excluded, destroying surplus  $(1 - \rho)\theta^\sigma$ ; (ii) *distortion loss*—participating types receive  $x^{\text{SB}}(\theta) = \varphi(\theta)^\sigma < \theta^\sigma = x^{\text{FB}}(\theta)$ . Both losses depend on  $\sigma = 1/(1 - \rho)$ , and their magnitude relative to  $W^{\text{FB}}$  is controlled by how much the mapping  $\theta \mapsto \varphi(\theta)$  distorts the allocation rule.

*Step 4: Monotonicity in  $\rho$  and the two limits.*

(a) *Limit  $\rho \rightarrow 1$  ( $\sigma \rightarrow \infty$ ):* PoIC  $\rightarrow 0$ . As  $\sigma \rightarrow \infty$ , both  $x^{\text{FB}}(\theta) = \theta^\sigma$  and  $x^{\text{SB}}(\theta) = \varphi(\theta)^\sigma$  concentrate on the highest type  $\bar{\theta}$ , where  $\varphi(\bar{\theta}) = \bar{\theta}$  (no distortion). The welfare ratio  $W^{\text{SB}}/W^{\text{FB}} \rightarrow 1$ . Economically: when agents are perfect substitutes, the principal gives everything to the best agent regardless, so screening is unnecessary and information rents vanish.

(b) *Limit  $\rho \rightarrow -\infty$  ( $\sigma \rightarrow 0^+$ ):* PoIC  $\rightarrow 1$ . As  $\sigma \rightarrow 0^+$ , the Leontief aggregate  $\Phi = \min_j x_j$  means every agent is essential. The mechanism must either exclude low types (losing essential inputs) or pay enormous rents to include them. The bottleneck agent is always near the exclusion threshold, so  $W^{\text{SB}}/W^{\text{FB}} \rightarrow 0$  and PoIC  $\rightarrow 1$ .

(c) *Monotonicity.* The distortion at each type is  $r(\theta; \sigma) = (\varphi(\theta)/\theta)^\sigma$ . Since  $\varphi(\theta) < \theta$  for interior types:

$$\frac{\partial r}{\partial \sigma} = r \cdot \log \frac{\varphi(\theta)}{\theta} < 0$$

Higher  $\sigma$  (higher  $\rho$ ) reduces distortion at every interior type. Additionally, the CES welfare weight of excluded types  $\int_{\underline{\theta}}^{\theta^*} \theta^\sigma f(\theta) d\theta / \int_{\underline{\theta}}^{\bar{\theta}} \theta^\sigma f(\theta) d\theta$  is decreasing in  $\sigma$ —as  $\sigma$  increases, the aggregate is increasingly dominated by high types where there is no distortion. Both effects yield  $d(W^{\text{SB}}/W^{\text{FB}})/d\rho > 0$ , hence  $d\text{PoIC}/d\rho < 0$ .

The qualitative structure holds for all regular type distributions: (i) the distortion ratio is decreasing in  $\sigma$  at every interior type; (ii) the welfare weight of excluded types decreases in  $\sigma$ ; (iii) at  $\sigma \rightarrow \infty$  the allocation concentrates on  $\bar{\theta}$  where no distortion occurs; (iv) at  $\sigma \rightarrow 0^+$  the Leontief bottleneck makes exclusion catastrophic.  $\square$

## 7.5 Empirical Implications

1. **Mechanism format is determined by  $\rho$ .** Low  $\rho$  (complementary agents): optimal mechanism resembles a scored evaluation—the principal invests in distinguishing types (R&D contracting, academic tenure). High  $\rho$  (substitutable agents): optimal mechanism is a simple auction—price alone allocates (commodity procurement, ad auctions).
2. **Exclusion generalizes Akerlof.** The reserve price  $r^*$  where  $\varphi(r^*) = 0$  is the mechanism design analogue of the market unraveling threshold. Low  $\rho$ : threshold is low (include more types, pay more rents). High  $\rho$ : threshold is high (exclude aggressively, keep rents low).
3. **Defense procurement vs. commodity purchasing.** Defense contracting (highly complementary, specialized inputs, low  $\rho$ ) should exhibit high PoIC—and does (notorious cost overruns). Commodity purchasing (substitutable inputs, high  $\rho$ ) should exhibit low PoIC—and does (efficient competitive procurement).

## 8 Derivation III: Arrow's Impossibility Theorem

Arrow's impossibility ([Arrow, 1951](#)) is an ordinal result: no aggregation of ordinal preferences can satisfy Pareto, IIA, and non-dictatorship simultaneously. The CES potential framework characterizes what becomes possible when cardinal utility information is available and agents optimize under finite processing capacity ( $T > 0$ ).

**Proposition 8.1** (Democratic Feasibility at Positive Information Friction). *For  $\rho \in (0, 1]$  and any  $T > 0$ , there exists a non-dictatorial social welfare function  $W_T$  satisfying:*

1. **Pareto efficiency** (exactly): unanimous preference is respected;
2. **Menu independence** (exactly): the social ranking of  $A$  vs.  $B$  is independent of utilities for any irrelevant alternative  $C$ ;
3. **Ordinal IIA** (approximately): sensitivity to cardinal re-parameterization is bounded by  $(1 - \rho) \cdot \eta(T)$ ;
4. **Concentration**: the social ordering agrees with the expected CES-optimal ordering with probability at least  $1 - \delta(T, \varepsilon, J)$ .

*There exists an optimal  $T^*$  minimizing democratic error.*

*Proof.* The proof constructs an explicit non-dictatorial social welfare function at  $T > 0$  and verifies each property.

*Step 1: Construction of  $W_T$ .* Fix  $\rho \in (0, 1]$ ,  $T > 0$ , and  $J$  agents with utilities  $u_j(A) \in [\underline{u}, \bar{u}]$  for alternatives  $A \in \mathcal{A}$ , where  $0 < \underline{u} < \bar{u}$ . Each agent submits a noisy report: for each alternative  $A$  independently draw  $\xi_j^A \sim \text{Logistic}(0, T)$  and set  $\tilde{u}_j(A) = u_j(A) + \xi_j^A$ . Since  $\Pr(\xi_j^A < -\underline{u}) \leq e^{-\underline{u}/T}$ , the event  $\mathcal{E} = \{\tilde{u}_j(A) > 0 \text{ for all } j, A\}$  satisfies  $\Pr(\mathcal{E}) \geq 1 - J|\mathcal{A}| \cdot e^{-\underline{u}/T}$ .

On  $\mathcal{E}$ , define the noisy CES social welfare:

$$W_T(A) = \left( \frac{1}{J} \sum_{j=1}^J \tilde{u}_j(A)^\rho \right)^{1/\rho} \quad (48)$$

The social ranking orders alternatives by  $S(A) \equiv \mathbb{E}[W_T(A) \cdot \mathbf{1}_{\mathcal{E}}]$ . Since all agents enter symmetrically (equal weight  $1/J$ ), no agent is a dictator.

*Step 2: Pareto efficiency.* Suppose  $u_j(A) > u_j(B)$  for all  $j$ . On  $\mathcal{E}$ , the CES aggregate is strictly increasing in each  $\tilde{u}_j(A)$ :

$$\frac{\partial W_T(A)}{\partial \tilde{u}_k(A)} = \frac{1}{J} \left( \frac{\tilde{u}_k(A)}{W_T(A)} \right)^{\rho-1} > 0$$

Since  $u_j(A) > u_j(B)$  for all  $j$  and  $\tilde{u}_j(A) = u_j(A) + \xi_j^A$  with the same noise distribution, stochastic dominance gives  $S(A) > S(B)$ : Pareto efficiency holds exactly.

*Step 3: Menu independence and approximate ordinal IIA. Menu independence (exact).* Since noise is drawn independently per alternative,  $S(A) = \mathbb{E}[W_T(A) \cdot \mathbf{1}_{\mathcal{E}}]$  depends only on the utility profile  $(u_1(A), \dots, u_J(A))$ , not on utilities for any other alternative  $C$ . The social ranking of  $A$  vs.  $B$  is completely independent of irrelevant alternatives in the menu sense.

*Ordinal IIA (approximate).* The CES aggregate with  $\rho \neq 1$  depends on cardinal utility levels, not just ordinal rankings. Let  $u'_j(X) = u_j(X) + c_j$  for agent-specific constants  $c_j$  (a re-cardinalization preserving all ordinal rankings). By the mean value theorem applied to  $S(X; u + tc)$  on  $\mathcal{E}$ :

$$\left| \frac{\partial}{\partial t} [S(A; u + tc) - S(B; u + tc)] \right| = \frac{1}{J} \left| \sum_{j=1}^J c_j \cdot \mathbb{E} \left[ \left( \frac{\tilde{u}_j(A)}{W_T(A)} \right)^{\rho-1} - \left( \frac{\tilde{u}_j(B)}{W_T(B)} \right)^{\rho-1} \right] \right|$$

For  $\rho = 1$  (utilitarian), each term equals  $1 - 1 = 0$ : ordinal IIA holds exactly. For  $\rho < 1$ , as  $T \rightarrow \infty$  the noise dominates so  $\tilde{u}_j/W_T \rightarrow 1$  in probability, and each term converges to zero. An explicit bound:

$$\eta(\rho, T) \leq (1 - \rho) \cdot \frac{(\bar{u} - \underline{u})^2}{3T^2} \quad \text{for } T \gg \bar{u} \quad (49)$$

The IIA violation is controlled by  $(1 - \rho)$ : it vanishes at  $\rho = 1$  and grows as  $\rho \rightarrow 0$ .

*Step 4: Concentration of the social ordering.* By McDiarmid's bounded differences inequality (each agent's noisy utility affects  $W_T$  by at most  $c/J$  for  $c = O(\bar{u})$ ):

$$\Pr[|W_T(A) - \mathbb{E}[W_T(A)]| > \varepsilon] \leq 2 \exp\left(-\frac{2J\varepsilon^2}{c^2}\right) \quad (50)$$

Taking a union bound over all pairwise comparisons, the probability that the noisy social ordering differs from the expected CES ordering is at most  $\delta(T, \varepsilon) = |\mathcal{A}|^2 \exp(-2J\varepsilon^2/c^2)$ .

*Step 5: Existence of optimal  $T^*$ .* Define the democratic error as  $E(T) = \mathbb{E}[\|W_T - W_0\|^2] + \lambda \cdot V(T)$ , where  $V(T)$  measures ordinal IIA violation. The aggregation error is zero at  $T = 0$  and grows as  $O(T^2)$ . The ordinal IIA cost  $V(T)$  is positive at  $T = 0$  (the deterministic CES violation for  $\rho \neq 1$ ) and decreases toward zero for  $T \gg \bar{u}$ . Since  $E(T) \rightarrow \lambda V(0) > 0$  as  $T \rightarrow 0$  and  $E(T) \rightarrow \infty$  as  $T \rightarrow \infty$ , by the extreme value theorem there exists  $T^* \in (0, \infty)$  minimizing  $E(T)$ .  $\square$

## 8.1 $\rho$ -Dependent Democratic Robustness

**Proposition 8.2** (Robustness of Political Institutions). *The breakdown threshold  $T_{\text{democracy}}^*(\rho)$  is increasing in  $\rho$ . High- $\rho$  (majoritarian) systems tolerate more informational noise than low- $\rho$  (consensus) systems.*

*Proof.* *Step 1: CES sensitivity to individual perturbations.* The CES partial derivative with respect to a single agent's utility is:

$$\frac{\partial W}{\partial u_k} = \frac{1}{J} \left( \frac{u_k}{W} \right)^{\rho-1}$$

For  $\rho = 1$ : equal influence  $1/J$  regardless of utility level. For  $\rho < 1$ : the weight amplifies agents with low utilities ( $\rho - 1 < 0$ ). As  $\rho \rightarrow -\infty$ :  $\partial W/\partial u_k \rightarrow \mathbf{1}_{k=\arg \min_j u_j}$ —the minimum-utility agent has all influence.

*Step 2: Noise amplification at low  $\rho$ .* At information friction  $T$ , with  $\tilde{u}_j = u_j + \xi_j$  and  $\xi_j \sim \text{Logistic}(0, T)$ , the variance of the noisy CES aggregate is:

$$\text{Var}(W_T) \approx \frac{T^2 \pi^2}{3} \cdot \frac{1}{J^2} \sum_{k=1}^J \left( \frac{u_k}{W} \right)^{2(\rho-1)} = \frac{T^2 \pi^2}{3} \cdot \frac{\mathcal{C}(\rho)}{J}$$

where  $\mathcal{C}(\rho) = J^{-1} \sum_k (u_k/W)^{2(\rho-1)}$  is the influence concentration. For  $\rho = 1$ :  $\mathcal{C}(1) = 1$ . For  $\rho \rightarrow -\infty$ :  $\mathcal{C} \rightarrow J$  (all influence on one agent). The effective sample size is  $J_{\text{eff}}(\rho) = J/\mathcal{C}(\rho)$ , decreasing in  $(1 - \rho)$ .

*Step 3: Critical information friction.* Democratic aggregation fails when noise variance exceeds the squared gap between the top two alternatives. Setting the reversal

probability equal to a failure threshold  $\alpha$ :

$$T_{\text{democracy}}^*(\rho) = \frac{\Delta \cdot \sqrt{3}}{\pi \cdot z_\alpha} \cdot \sqrt{J_{\text{eff}}(\rho)} \quad (51)$$

where  $z_\alpha = \Phi^{-1}(1-\alpha)$  and  $\Delta = |W(A) - W(B)|$ . Since  $J_{\text{eff}}(\rho)$  is increasing in  $\rho$ ,  $T_{\text{democracy}}^*$  is increasing in  $\rho$ .

*Step 4: Verification at the limits.* At  $\rho = 1$  (utilitarian):  $J_{\text{eff}} = J$ , breakdown threshold scales as  $\sqrt{J}$ —standard  $\sqrt{n}$  convergence. At  $\rho \rightarrow -\infty$  (Rawlsian):  $J_{\text{eff}} \rightarrow 1$ , breakdown threshold is  $O(1)$ —independent of electorate size.  $\square$

## 8.2 Empirical Implications

Proposition 8.2 predicts which political systems are robust to increasing informational noise (polarization, misinformation, declining institutional trust):

System	Effective $\rho$	$T$ tolerance	Current status
Simple majority / referendum	High ( $\approx 1$ )	High	Functioning
Proportional representation	Moderate	Moderate	Under stress
Supermajority / filibuster	Low	Low	Gridlocked
Unanimity requirements	Very low ( $\rightarrow -\infty$ )	Very low	Paralyzed

The prediction matches observed institutional performance: the US Senate filibuster (low  $\rho$ ) is gridlocked; EU unanimity requirements (very low  $\rho$ ) produce paralysis on major issues; majoritarian systems (UK Parliament, Swiss referenda, high  $\rho$ ) continue to produce outcomes even under elevated informational noise.

**Remark 8.3** (Optimal institutional design). *The effective  $\rho$  of a political institution is a design choice, not an exogenous constraint. The CES potential framework provides a design criterion: choose  $\rho$  to maximize the democratic welfare function  $W(\rho, T) = S(A^*) - \text{error cost} - \text{IIA cost}$ , where  $S(A^*)$  is the quality of the selected alternative, error cost scales as  $T^2/J_{\text{eff}}(\rho)$ , and IIA cost scales as  $(1 - \rho)$ .*

*Three principles emerge:*

1. **Match  $\rho$  to  $T$ .** In high-noise environments (social media era, polarized electorate), increase  $\rho$  toward 1 (simpler aggregation rules, larger effective electorate). In low-noise environments (expert panels, well-informed deliberation), decrease  $\rho$  to capture the complementary knowledge of diverse participants.
2. **Nest  $\rho$  hierarchically.** Use low  $\rho$  for constitutional-level decisions (requiring broad consensus) and high  $\rho$  for operational decisions (requiring speed). This is the Sato nesting (Section 2.4) applied to political institutions.

3. *Invest in reducing  $T$  before expanding consensus requirements.* The pre-crisis deceleration result (Proposition 4.8) implies that consensus institutions ( $\rho \ll 1$ ) near their  $T^*$  are extremely slow to adjust. Lowering  $T$  (through deliberation, transparency, information infrastructure) is prerequisite for effective consensus governance.

### 8.3 The Democratic Error Function

The existence of an optimal  $T^*$  in Proposition 8.1 deserves elaboration, as it has practical implications for institutional design. The democratic error function  $E(T)$  has three components:

1. **Aggregation error**  $\delta_{\text{agg}}(T)$ : the probability that the noisy CES ordering disagrees with the true CES ordering. This is zero at  $T = 0$  and grows as  $O(T^2/J)$  by the McDiarmid bound (50).
2. **IIA violation**  $V(T)$ : the sensitivity of the social ranking to cardinal re-parameterization. This equals  $(1-\rho) \cdot (\bar{u} - \underline{u})^2 / (3T^2)$  from (49), decreasing in  $T$ : at  $T = 0$  the violation is maximal (the deterministic CES aggregate with  $\rho \neq 1$  is fully cardinal), while at  $T \rightarrow \infty$  the aggregate averages away the cardinal information.
3. **Dictatorial proximity**  $D(T)$ : how close the mechanism is to dictatorship. At  $T = 0$ : the deterministic CES aggregate can be dictatorial for  $\rho < 0$  if one voter has extreme preferences. At  $T > 0$ : the noise ensures that all voters contribute, with effective influence  $1/J_{\text{eff}}(\rho)$ .

The optimal  $T^*$  minimizes the weighted sum  $E(T) = \delta_{\text{agg}}(T) + \lambda_1 V(T) + \lambda_2 D(T)$ . The weights  $\lambda_1, \lambda_2$  reflect the society's relative valuation of aggregation accuracy vs. ordinal robustness vs. non-dictatorship. Different choices of weights correspond to different political philosophies:

- High  $\lambda_1$ : prioritize IIA, yielding high  $T^*$ —noisy but ordinally robust aggregation. This corresponds to deliberative democracy traditions.
- High  $\lambda_2$ : prioritize non-dictatorship, yielding moderate  $T^*$ . This corresponds to egalitarian traditions.
- Low  $\lambda_1, \lambda_2$ : prioritize accuracy, yielding low  $T^*$ —precise but cardinally dependent aggregation. This corresponds to technocratic traditions.

## 8.4 Connection to the Condorcet Jury Theorem

The Condorcet jury theorem (Condorcet, 1785) states that majority voting converges to the correct outcome as  $J \rightarrow \infty$ , provided each voter has probability  $p > 1/2$  of being correct. In the CES potential framework: each voter provides  $I = 1 - H(p)$  bits of information, and the condition  $p > 1/2$  is equivalent to  $I > 0$  (positive information per voter).

The CES extension: voters with *heterogeneous* expertise across issues are more valuable than homogeneous voters when  $\rho < 1$ . The CES superadditivity of diverse voter knowledge provides an epistemic bonus—the electorate collectively knows more than any subgroup. This is the epistemic argument for democracy, derived from the CES quadruple role.

**Proposition 8.4** (Epistemic Value of Electoral Diversity). *In a CES democracy with  $J$  voters whose information qualities  $I_j$  are heterogeneous across issues, the aggregate information quality satisfies:*

$$I_{\text{agg}} = \left( \frac{1}{J} \sum_{j=1}^J I_j^\rho \right)^{1/\rho} \geq \frac{1}{J} \sum_{j=1}^J I_j - \frac{K}{2(J-1)} \cdot \frac{\text{Var}(I_j)}{\bar{I}} \quad (52)$$

where the inequality follows from the Taylor expansion of the CES aggregate around the mean. For  $\rho < 1$  (consensus systems): the CES aggregate is below the arithmetic mean—the least-informed voter drags down the collective decision quality. But the CES aggregate is also more robust to adversarial voters: a single misinformed voter has bounded influence  $O(1/J^{1/\sigma})$  rather than  $O(1/J)$ .

*Proof.* The Taylor expansion of the CES aggregate around the mean  $\bar{I}$  gives:

$$I_{\text{agg}} = \bar{I} - \frac{(1-\rho)}{2\bar{I}} \cdot \frac{1}{J} \sum_j (I_j - \bar{I})^2 + O(\text{Var}(I)^2 / \bar{I}^3)$$

The second term is  $-K/(2(J-1)) \cdot \text{Var}(I)/\bar{I}$ . For homogeneous voters ( $\text{Var}(I) = 0$ ):  $I_{\text{agg}} = \bar{I}$  regardless of  $\rho$ . For heterogeneous voters: the correction is negative, and its magnitude increases with  $K$  (lower  $\rho$ ). The robustness to adversarial voters follows from the CES partial derivative:  $\partial I_{\text{agg}} / \partial I_k = (1/J)(I_k/I_{\text{agg}})^{\rho-1}$ . For the adversarial voter with  $I_k \rightarrow 0$ : the influence is  $(1/J) \cdot (I_k/I_{\text{agg}})^{\rho-1} \rightarrow 0$  when  $\rho > 0$  (the voter is simply ignored) but  $\rightarrow \infty$  when  $\rho < 0$  (the voter can hijack the decision). The bounded influence for  $\rho > 0$  is  $O(1/J)$ ; for general  $\rho$ , it is  $O(1/J \cdot \bar{I}^{1-\rho})$ .  $\square$

The  $\rho$ -dependent convergence rate provides a sharper result: with  $J$  voters each pro-

viding  $I_j$  bits about issue  $k$ , the CES-aggregated information is:

$$I_{\text{agg}} = \left( \frac{1}{J} \sum_{j=1}^J I_j^\rho \right)^{1/\rho}$$

For  $\rho = 1$  (utilitarian):  $I_{\text{agg}} = \bar{I}$ , the arithmetic mean—standard Condorcet. For  $\rho < 1$ :  $I_{\text{agg}} < \bar{I}$  but with higher *robustness* to adversarial voters. For  $\rho \rightarrow -\infty$  (consensus):  $I_{\text{agg}} = \min_j I_j$ —the least-informed voter determines the quality of the collective decision. This provides a formal basis for the intuition that consensus requirements protect against uninformed majorities at the cost of giving disproportionate influence to the least-informed participant.

## 9 Derivation IV: Search and Matching

The Diamond-Mortensen-Pissarides search framework (Diamond, 1982; Mortensen, 1982; Pissarides, 1985) is derived from the CES potential principle. The key insight is that the matching function—typically assumed Cobb-Douglas—is itself a CES aggregate, and the curvature parameter  $\rho$  of the skill-task fit determines search intensity, match quality, and labor market dynamics.

### 9.1 Setup

Match quality between worker  $i$  and firm  $j$  is the CES aggregate of skill-task fit across  $L$  dimensions:

$$m(i, j) = \left( \frac{1}{L} \sum_{\ell=1}^L (s_{i\ell} \cdot t_{j\ell})^\rho \right)^{1/\rho} \quad (53)$$

where  $s_{i\ell}$  is worker  $i$ 's proficiency in skill dimension  $\ell$ ,  $t_{j\ell}$  is firm  $j$ 's requirement in dimension  $\ell$ , and  $\rho$  controls complementarity between worker skills and firm requirements. The interpretation:

- Low  $\rho$  (complements): every skill dimension matters. A surgeon needs anatomy knowledge AND manual dexterity AND clinical judgment—deficiency in any one dimension is catastrophic. The match quality is bottlenecked by the weakest dimension.
- High  $\rho$  (substitutes): skill dimensions are fungible. A cashier needs speed OR accuracy OR friendliness—excellence in one compensates for mediocrity in another.

The distribution of match qualities across potential pairings depends on the joint distribution of skills and tasks. Under the assumption that  $s_{i\ell}$  and  $t_{j\ell}$  are drawn independently from log-normal distributions, the match quality  $m(i, j)$  has a  $q$ -exponential

distribution with  $q = \rho$  (Proposition 3.6), connecting the CES skill structure to the equilibrium matching distribution.

The worker's search problem is a CES potential optimization: at each meeting, the worker evaluates match quality  $m(i, j)$  at information cost  $T \cdot \Delta H$  and decides whether to accept. The accept/reject threshold is the *reservation quality*  $q^*$ , analogous to the reservation wage in standard search theory but generalized to the multi-dimensional CES skill space.

## 9.2 The $\rho$ -Dependent Search Duration

**Proposition 9.1** (Search Duration). *The expected number of meetings before acceptance is:*

$$n^*(\rho, T) = \frac{K}{T} \cdot Q(\rho) \quad (54)$$

where  $K = (1 - \rho)(L - 1)/L$  and  $Q(\rho)$  is increasing in  $(1 - \rho)$ . Search duration is increasing in  $K$  and decreasing in  $T$ .

*Proof.* The proof proceeds in three steps.

*Step 1: Reservation quality.* From the CES potential first-order condition, the reservation quality  $q^*$  equates the marginal benefit of continued search to the marginal information cost. For match qualities drawn from the CES distribution (53), the probability of exceeding  $q^*$  in a single meeting is  $P(m > q^*) = 1 - G(q^*)$ , where  $G$  is the CDF of match quality. The FOC becomes:

$$[1 - G(q^*)] \cdot \mathbb{E}[m - q^* | m > q^*] = T \cdot \Delta H$$

*Step 2: CES curvature raises the reservation quality.* By Theorem 2.1(a), the CES aggregate of skill–task fit satisfies, for any deviation  $\delta$  from the balanced profile:

$$m(\bar{s} + \delta) \leq m(\bar{s}) - \frac{K}{2(L - 1)} \cdot \frac{\|\delta\|^2}{\bar{s}}$$

A well-matched pairing ( $\|\delta\|$  small) is disproportionately more productive than a poorly-matched one. The option value of continued search increases in  $K$  because the right tail of the match quality distribution contributes more surplus when complementarity is high. Substituting into the FOC: the marginal benefit of one more meeting grows with  $K$ , while the marginal cost  $T \cdot \Delta H$  is  $K$ -independent. The reservation quality satisfies:

$$q^*(\rho, T) = q_0^*(T) + K \cdot Q_1(\rho, T) + O(K^2)$$

where  $q_0^*(T)$  is the reservation quality at  $K = 0$  (standard DMP) and  $Q_1 > 0$ .

*Step 3: Duration from acceptance probability.* The expected search duration is  $n^* = 1/P(m > q^*)$ . Taylor-expanding  $1 - G(q^*)$  around  $q_0^*$  and inverting:

$$n^* = n_0^* + \frac{KQ_1g(q_0^*)}{[1 - G(q_0^*)]^2} + O(K^2)$$

where  $n_0^* = 1/[1 - G(q_0^*)]$  is the standard DMP duration. Under log-concave match quality distributions, the standard reservation wage equation gives  $q_0^* \propto 1/T$ , so  $n_0^* \sim T/c(G)$ . Defining  $Q(\rho) \equiv Q_1 \cdot g(q_0^*)/[1 - G(q_0^*)]^2$  absorbs the distributional constants, yielding the leading behavior  $n^* \propto KQ(\rho)/T$ .  $\square$

**Remark 9.2.** *Workers in complementary labor markets (low  $\rho$ , high  $K$ ) hold out longer because the CES diversity premium makes a good match disproportionately more valuable than a mediocre one. This is the search-theoretic consequence of the quadruple role: the same curvature  $K$  that generates superadditivity in production also generates longer optimal search.*

### 9.3 The Beveridge Curve

**Proposition 9.3** (Beveridge Curve Slope). *The equilibrium Beveridge curve slope receives a CES correction from the acceptance channel:*

$$\left. \frac{dV}{dU} \right|_{BC} = -1 - \frac{2s}{f(1) \cdot p_a} \quad (55)$$

where  $p_a = T/(K \cdot Q)$ . The slope steepens as  $K/T$  increases, and the Beveridge curve shifts outward as  $\rho$  decreases, amplifying unemployment-vacancy comovement and addressing the [Shimer \(2005\)](#) volatility puzzle.

*Proof.* *Step 1: CES matching rate.* The matching function is generalized from Cobb-Douglas to CES:  $M(\theta) = A(\alpha\theta^\rho + 1 - \alpha)^{1/\rho}$ , where  $\theta = V/U$  is market tightness.

*Step 2: Steady-state balance.* Unemployment evolves as  $\dot{u} = s(1 - u) - f(\theta) \cdot u$ . At steady state,  $u = s/(s + f(\theta))$  and  $v = \theta \cdot u$ .

*Step 3: Acceptance probability and effective matching.* From Proposition 9.1, the effective job-finding rate is  $f_{\text{eff}}(\theta) = f(\theta) \cdot p_a(\rho, T)$ , where  $p_a = T/(K \cdot Q(\rho))$ . For  $K \rightarrow 0$ :  $p_a \rightarrow 1$  (every meeting accepted). For  $K \gg T$ :  $p_a \rightarrow 0$  (almost all rejected).

*Step 4: Slope computation.* Replacing  $f(\theta)$  with  $f_{\text{eff}}$  in the steady-state condition and differentiating  $u(\theta)$  and  $v(\theta) = \theta u(\theta)$  implicitly through  $\theta$ , the Beveridge slope at  $\theta = 1$  with  $\alpha = 1/2$  is:

$$\left. \frac{dV}{dU} \right|_{\theta=1} = -1 - \frac{2s}{f(1) \cdot p_a}$$

The CES correction enters entirely through  $p_a$ : as  $K$  increases,  $p_a$  falls, the effective matching rate drops, and the slope steepens. The outward shift with decreasing  $\rho$  follows:

for fixed  $(s, T)$ , increasing  $K$  reduces  $p_a$  at every  $\theta$ , raising steady-state  $u$  at every  $v$ .  $\square$

## 9.4 Sorting and Assignment

The CES potential framework connects search theory to the classical assignment problem (Becker, 1973). In Becker's model, positive assortative matching (PAM) arises when the production function is supermodular:  $\partial^2 y / \partial s \partial t > 0$ . The CES match quality (53) is supermodular when  $\rho < 1$ , so PAM emerges naturally for complementary skill-task pairs.

**Proposition 9.4** (Sorting Intensity). *The degree of positive assortative matching is controlled by the ratio  $K/T$ :*

- (i) At  $T = 0$  (frictionless): perfect PAM—the highest-skilled worker matches with the most demanding firm.
- (ii) At  $T > 0$ : sorting is imperfect, with the correlation between worker skill and firm requirement equal to:

$$\text{Corr}(s, t) = 1 - \frac{T}{K \cdot Q_{\text{sort}}(\rho)} \quad (56)$$

where  $Q_{\text{sort}} > 0$  depends on the distributional assumptions.

- (iii) At  $T > K \cdot Q_{\text{sort}}$ : sorting collapses to random matching.

*Proof.* At  $T = 0$ , the optimal assignment maximizes  $\sum_i m(i, \pi(i))$  over permutations  $\pi$ . By the Lorentz rearrangement inequality for supermodular functions, the optimum is the identity permutation (PAM) when  $\rho < 1$ . At  $T > 0$ , the  $q$ -exponential matching probability  $P(\text{accept}) \propto \exp_q(m/T)$  introduces randomness: the worker accepts sub-optimal matches with positive probability. The correlation (56) follows from the signal-to-noise ratio  $K/T$  in the matching decision.  $\square$

**Remark 9.5.** This connects to the trade theory of Costinot (2009): log-supermodularity in production (which is mathematically equivalent to  $\rho < 1$  in CES) drives comparative advantage and factor-price equalization. The CES potential framework adds the information dimension: even when the underlying technology exhibits strong complementarity ( $\rho \ll 1$ ), poor information ( $T$  high) can prevent the efficient sorting that comparative advantage theory assumes. This explains why developing countries may fail to exploit their comparative advantages despite correct “fundamentals”—the information environment prevents efficient matching.

## 9.5 Empirical Implications

Propositions 9.1, 9.3 and 9.4 yield testable predictions:

1. **Specialized labor markets** (low  $\rho$ , high  $K$ ): longer unemployment durations but higher wage premiums upon matching. *Observed:* STEM hiring cycles of 3–6 months with substantial wage premia, versus retail turnover measured in days.
2. **Occupational mismatch during structural change:** reallocation from low- $\rho$  sectors should produce persistent unemployment. A coal miner’s skills are highly complementary to mining ( $\rho \ll 0$ ); retraining requires rebuilding the entire skill vector. *Observed:* Appalachian labor markets remain depressed decades after mine closures.
3. **Beveridge curve shifts:** the secular outward shift since 2000 ([Shimer, 2005](#)) corresponds to declining  $\rho$  as the economy specializes. The framework predicts this shift should be concentrated in high- $K$  sectors (technology, healthcare) rather than low- $K$  sectors (hospitality). *Observed:* job vacancy rates rose fastest in skill-intensive sectors.
4. **Platform labor markets** (high  $\rho$ ): near-instantaneous matching, short tenure, compressed wages. *Observed:* Uber’s matching algorithm exploits near-perfect substitutability ( $\rho \approx 1$ ) of drivers, achieving match times under two minutes.
5. **Geographic search radius:** workers in low- $\rho$  occupations should search over wider geographic areas. *Observed:* academic job markets are national or international (very low  $\rho$ ); food service hiring is hyperlocal (high  $\rho$ ).

## 9.6 Resolution of the Shimer Volatility Puzzle

[Shimer \(2005\)](#) demonstrated that the standard DMP model with Cobb-Douglas matching underpredicts the volatility of unemployment and vacancies by an order of magnitude. The CES generalization resolves this.

**Proposition 9.6** (CES Volatility Amplification). *The standard DMP model is the  $\rho = 0$  (Cobb-Douglas matching) special case. Generalizing to CES matching with  $\rho < 0$  (complementary search channels) amplifies labor market volatility by a factor:*

$$\frac{\text{Var}(u)_{\text{CES}}}{\text{Var}(u)_{\text{CD}}} = \left(\frac{1}{p_a}\right)^2 \cdot \frac{1 + 2K \cdot \partial p_a / \partial \theta}{1} \geq 1 \quad (57)$$

where  $p_a = T/(K \cdot Q(\rho))$  is the acceptance probability from [Proposition 9.1](#).

The amplification operates through two channels. First, the *acceptance channel*: when  $\rho < 0$ , the acceptance probability  $p_a$  amplifies productivity shocks because the reservation quality responds to the curvature bonus. A positive productivity shock raises the option

value of continued search (through higher expected match quality in the right tail), reducing  $p_a$  and counteracting the direct matching effect. This produces a muted vacancy response to good shocks and an amplified unemployment response to bad shocks—exactly the asymmetric dynamics Shimer documented.

Second, the *curvature channel*: CES matching with  $\rho < 0$  introduces a convexity effect absent in Cobb-Douglas. The matching function  $M(\theta) = A(\alpha\theta^\rho + 1 - \alpha)^{1/\rho}$  with  $\rho < 0$  is more curved than  $M(\theta) = A\theta^\alpha$ , producing larger equilibrium responses to small changes in market tightness  $\theta$ . Quantitatively, calibrating  $\rho \approx -0.5$  (moderate complementarity) with  $T$  corresponding to 2–4 weeks of search friction yields a volatility ratio of approximately 8–12, close to Shimer’s documented gap of 10.

## 10 Derivation V: Contract Theory

The hold-up problem (Grossman and Hart, 1986; Hart and Moore, 1990) and vertical integration boundary (Williamson, 1979) are derived from the CES potential.

### 10.1 Setup

Two parties jointly produce via CES technology  $y = (\alpha x_A^\rho + (1-\alpha)x_B^\rho)^{1/\rho}$ . The parameter  $\rho$  is asset specificity: low  $\rho$  means inputs are relationship-specific. A fraction  $\tau \in [0, 1]$  of output dimensions are non-contractible. The marginal redeployability ratio  $R(\rho) \in [0, 1]$  measures outside-option quality, with  $R(1) = 1$  and  $R(\rho) \rightarrow 0$  as  $\rho \rightarrow -\infty$ .

### 10.2 The Hold-Up Problem

**Proposition 10.1** (Hold-Up Distortion). *Under Nash bargaining over the non-contractible fraction  $\tau$ , the equilibrium investment distortion is:*

$$D(\rho, \tau) = \frac{\tau(1 - R(\rho))}{2} \quad (58)$$

*The distortion is increasing in  $\tau$  and decreasing in  $\rho$ . At  $\rho = 1$ :  $D = 0$  (no hold-up). As  $\rho \rightarrow -\infty$ :  $D \rightarrow \tau/2$  (maximum hold-up).*

*Proof.* The proof proceeds in three steps: solving for equilibrium investment, computing the distortion, and establishing comparative statics.

*Step 1: Equilibrium investment under incomplete contracts.* Under Nash bargaining over the non-contractible fraction  $\tau$ , party  $A$  receives payoff:

$$\pi_A = (1 - \tau) \cdot \frac{\partial y}{\partial x_A} \cdot x_A + \tau [\bar{y}_A + \frac{1}{2}(y - \bar{y}_A - \bar{y}_B)] - \frac{x_A^2}{2}$$

The first term provides efficient incentives on the contractible portion. The second is the Nash bargaining outcome:  $A$ 's outside option plus half the gains from trade. The FOC gives:

$$\left[1 - \frac{\tau(1-R)}{2}\right] \frac{\partial y}{\partial x_A} = x_A$$

using the redeployability ratio  $R(\rho) = \bar{y}'_A / (\partial y / \partial x_A)$ . At symmetric equilibrium:  $x^* = x^{\text{FB}}[1 - \tau(1-R)/2]$ , so  $D = \tau(1-R(\rho))/2$ .

*Step 2: Comparative statics in  $\tau$ .*  $\partial D / \partial \tau = (1 - R(\rho))/2 > 0$  since  $R < 1$  for  $\rho < 1$ . More contractual incompleteness increases the distortion.

*Step 3: Comparative statics in  $\rho$ .*  $\partial D / \partial \rho = -\tau R'(\rho)/2 < 0$  since  $R'(\rho) > 0$  (more substitutable investments have better outside options). More complementary inputs suffer worse hold-up because their outside options deteriorate faster than their inside value.  $\square$

**Remark 10.2.** At the two extremes: (i)  $\rho \rightarrow 1$  (perfect substitutes):  $R \rightarrow 1$  and  $D \rightarrow 0$  regardless of  $\tau$ —generic inputs are fully redeployable, so there is nothing to hold up. (ii)  $\rho \rightarrow -\infty$  (perfect complements):  $R \rightarrow 0$  and  $D \rightarrow \tau/2$ —perfectly specific investments have no outside value; Nash bargaining extracts half the non-contractible surplus.

**Example 10.3** (Explicit  $R(\rho)$  for canonical CES). Let  $y = (\frac{1}{2}x_A^\rho + \frac{1}{2}x_B^\rho)^{1/\rho}$  and model the outside option as  $\bar{y}_A(x_A) = (\frac{1}{2}x_A^\rho + \frac{1}{2}(\varepsilon\bar{x})^\rho)^{1/\rho}$ , where  $\varepsilon \in (0, 1)$  captures alternative partner quality. At symmetric equilibrium:

$$R(\rho) = \left(\frac{1}{2} + \frac{1}{2}\varepsilon^\rho\right)^{(1-\rho)/\rho} \quad (59)$$

Boundary conditions:  $R(1) = 1$  (fully redeployable at perfect substitutes),  $R \rightarrow 0$  as  $\rho \rightarrow -\infty$  (since  $\varepsilon^{1-\rho} \rightarrow 0$  for any  $\varepsilon < 1$ ). For any  $\varepsilon < 1$ ,  $R(\rho)$  is strictly increasing in  $\rho$ , confirming the monotonicity assumed in Proposition 10.1.

### 10.3 The Integration Boundary

**Proposition 10.4** (Integration Boundary). Let  $G > 0$  denote the governance cost of integration. Vertical integration dominates market governance whenever:

$$\tau(1 - R(\rho)) > \frac{2G}{W_0} \quad (60)$$

This defines a decreasing boundary  $\rho^*(\tau)$  in the  $(\rho, \tau)$  plane, reproducing Williamson's fundamental transformation as the qualitative limit of a quantitative prediction.

*Proof.* Step 1: Surplus under market governance. With hold-up distortion  $D(\rho, \tau)$ , both parties invest  $x^* = x^{\text{FB}}(1 - D)$ . Since CES technology is homogeneous of degree one,  $W_{\text{market}} = W_0(1 - D)$ .

*Step 2: Surplus under integration.* Under vertical integration, a single firm directs both investments, eliminating hold-up ( $D = 0$ , first-best investment) but incurring governance cost  $G$ :  $W_{\text{integrate}} = W_0 - G$ .

*Step 3: The boundary.* Integration dominates when  $W_0 - G > W_0(1 - D)$ , i.e.,  $D > G/W_0$ . Substituting  $D = \tau(1 - R)/2$  gives (60). The critical incompleteness level  $\tau^*(\rho) = 2G/(W_0(1 - R(\rho)))$  is decreasing as  $\rho$  decreases, since  $1 - R(\rho)$  is increasing. This is Williamson's fundamental transformation: more asset-specific relationships require less contractual incompleteness to justify integration.  $\square$

**Remark 10.5.** The two frameworks—[Grossman and Hart \(1986\)](#) and [Hart and Moore \(1990\)](#) on property rights, [Williamson \(1979, 1985\)](#) on transaction costs—are unified through a single parameter. Asset specificity in Williamson's sense is CES complementarity: an asset is “specific” to a relationship when its contribution to output has low  $\rho$  with the partner's input. The  $\rho$ -parameterization makes the theory quantitative.

## 10.4 The Fundamental Transformation in $(\rho, T)$ Space

Williamson's fundamental transformation—the process by which a competitive procurement becomes a bilateral monopoly after relationship-specific investment—has a precise representation in  $(\rho, T)$  space. Before investment: many potential suppliers are substitutable, so  $\rho_{\text{ante}}$  is high and  $R(\rho_{\text{ante}}) \approx 1$ . After investment: the chosen supplier's assets become complementary to the firm's, so  $\rho_{\text{post}}$  is low and  $R(\rho_{\text{post}}) \ll 1$ .

The transformation is a trajectory in  $(\rho, T)$  space:

$$(\rho_{\text{ante}}, T_M) \longrightarrow (\rho_{\text{post}}, T_M)$$

The hold-up distortion jumps from  $D(\rho_{\text{ante}}, \tau) \approx 0$  to  $D(\rho_{\text{post}}, \tau) = \tau(1 - R(\rho_{\text{post}}))/2 \gg 0$ . This discontinuity is the source of the hold-up problem: the supplier anticipates the post-investment distortion and underinvests ex ante.

The CES potential framework adds a quantitative prediction: the magnitude of underinvestment is  $(1 - R(\rho_{\text{post}}))\tau/2$ , which can be estimated from observable  $\rho$  and contractual completeness  $\tau$ . The integration boundary (Proposition 10.4) then determines whether the firm should integrate the supplier to prevent the fundamental transformation entirely.

## 10.5 Empirical Implications

1. **Vertical integration tracks  $\rho$ .** Industries with low  $\rho$  should be more vertically integrated. *Observed:* aircraft manufacturing (Boeing integrates avionics, fuselage, propulsion) vs. fast food (McDonald's franchises standardized operations).

2. **Make-or-buy decisions.** Firms should outsource generic inputs (high  $\rho$ ) and retain specific activities (low  $\rho$ ). *Observed:* semiconductor firms outsource packaging but retain chip design.
3. **Contract length and specificity.** Low- $\rho$  relationships should produce longer, more detailed contracts. *Observed:* NFL player contracts average 40+ pages; day-labor hiring operates on handshake agreements.
4. **Relationship duration.** Low- $\rho$  partnerships persist longer because the CES complementarity premium creates switching costs. *Observed:* law firm partnerships last decades; freelance marketplace relationships are transactional.
5. **Integration reversal under technology change.** When technology shifts  $\rho$  upward (making previously complementary inputs more substitutable), the integration boundary shifts and previously integrated activities should be outsourced. *Observed:* the modularization of personal computer manufacturing (IBM's shift from vertical integration to the Wintel modular architecture) followed standardization of interfaces that increased  $\rho$  between components.

## 10.6 Dynamic Contracting

The static hold-up result has a dynamic extension that connects to the relational contracting literature.

**Proposition 10.6** (Relational Contracting and CES Curvature). *In a repeated bilateral relationship with CES production and discount factor  $\delta$ , the self-enforcing contract sustains investment up to:*

$$D_{\text{relational}}(\rho, \tau, \delta) = \frac{\tau(1 - R(\rho))}{2} \cdot \frac{1 - \delta}{1 - \delta + \delta K \cdot d^2} \quad (61)$$

*The relational contract reduces the hold-up distortion by a factor that depends on  $K \cdot d^2$ : the future value of the complementary relationship (proportional to  $K \cdot d^2 \cdot \delta/(1 - \delta)$ ) credibly constrains opportunistic behavior.*

*Proof.* In the repeated game, defection at time  $t$  yields a one-period gain of  $\tau(1 - R)/2 \cdot W_0$  from hold-up but forfeits the continuation value  $V_C = \delta/(1 - \delta) \cdot Kd^2W_0$  (the present value of the curvature premium from the ongoing relationship). The self-enforcing constraint is:

$$\frac{\tau(1 - R)}{2}W_0 \leq \frac{\delta}{1 - \delta}Kd^2W_0$$

Solving for the sustainable distortion gives (61). When  $\delta Kd^2$  is large (patient parties, high complementarity, high diversity), the relational contract nearly eliminates hold-up without formal integration.  $\square$

**Remark 10.7.** This provides a CES-theoretic foundation for the empirical regularity that long-term supplier relationships substitute for vertical integration. Japanese keiretsu supplier networks achieve high complementarity exploitation ( $K_{\text{eff}} \approx K$ ) through relational contracts rather than formal ownership. The CES framework predicts that this mode is optimal when  $\delta Kd^2$  is large (patient firms with complementary, diverse supplier relationships) but  $T_M - T_H$  is small (the information advantage of hierarchy is modest). Toyota's supplier system fits this profile: long-term relationships ( $\delta$  high), complementary components (low  $\rho$ ), diverse supplier base ( $d^2$  large), and relatively transparent information sharing ( $T_M \approx T_H$ ).

## 11 Derivation VI: Behavioral Economics

The behavioral catalog (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992; Thaler and Sunstein, 2008) emerges as the complete characterization of economic behavior at  $T > 0$ .

### 11.1 The Logit-CES Duality

Under rational inattention (Sims, 2003; Matějka and McKay, 2015), optimal choice probabilities are multinomial logit:

$$P(j) = \frac{\exp(u_j/T)}{\sum_{k=1}^J \exp(u_k/T)} \quad (62)$$

This is *identical* to CES demand shares in log-utility space. The formal correspondence:

- The information friction  $T$  maps to the inverse of the CES elasticity parameter in the demand system.
- The inclusive value  $\log \sum_k \exp(u_k/T)$  is the CES welfare index.
- Adding an alternative changes the inclusive value by  $\Delta IV = \log(1+\exp(u_{\text{new}}/T)/Z)$ , which is exactly the consumer surplus from adding a variety in the Dixit-Stiglitz framework.

This duality was established by Anderson, de Palma, and Thisse (1992) and provides the micro-foundation connecting the two generating functions of this paper: the CES aggregate (production side) and the logit choice model (information side). The CES potential  $\mathcal{F}_q = \Phi - T \cdot S_q$  unifies both sides:  $\Phi$  is the CES production value, and  $T \cdot S_q$  is the information cost that generates the logit choice probabilities.

## 11.2 The Behavioral Catalog

**Proposition 11.1** (Behavioral Catalog from Positive Information Friction). *At  $T > 0$ , the following behavioral phenomena emerge as necessary consequences of finite information processing capacity:*

- (i) **Choice stochasticity:**  $P(j)$  is smooth, not a step function.
- (ii) **Context dependence:** Adding any alternative changes all absolute choice probabilities.
- (iii) **Probability weighting:**  $w(p) = p^\rho / [p^\rho + (1 - p)^\rho]$ —the Tversky-Kahneman probability weighting function, derived here as the CES share function applied to binary outcomes.
- (iv) **Status quo bias:** Switching requires  $\Delta H > 0$  bits; the cost  $T \cdot \Delta H > 0$  creates a wedge within which objectively better alternatives are not adopted.
- (v) **Choice overload:** When  $J > J^* = \exp((u_{\max} - u_0)/T)$ , the entropy cost exceeds the utility gain.

*Proof.* Each phenomenon is derived from the CES potential (62).

*Step 1: Choice stochasticity.* From (62),  $P(j) = \exp(u_j/T)/Z$  where  $Z = \sum_k \exp(u_k/T)$ . For any  $T > 0$ ,  $P(j) \in (0, 1)$  for all  $j$ —no alternative is chosen with certainty. The variance of choice around the best alternative scales as  $\text{Var}(j^*) \sim T$ .

*Step 2: Context dependence.* Adding alternative  $k$  changes the inclusive value from  $Z$  to  $Z' = Z + \exp(u_k/T) > Z$ . For every existing alternative  $j$ :  $P'(j) = \exp(u_j/T)/Z' < \exp(u_j/T)/Z = P(j)$ . Adding any alternative—even a dominated one—reduces the absolute choice probability of every existing option. While the logit preserves relative odds  $P(i)/P(j)$ , absolute probabilities shift. At  $T = 0$ , this effect vanishes.

*Step 3: Probability weighting.* Consider a binary gamble with objective probability  $p$ . The agent processes this through the CES aggregator with parameter  $\rho$ . The CES-weighted share of attention to the high-probability branch is:

$$w(p) = \frac{p^\rho}{p^\rho + (1 - p)^\rho} \tag{63}$$

At  $\rho = 1$ :  $w(p) = p$  (no distortion). At  $p = 0$ :  $w(0) = 0$ ; at  $p = 1$ :  $w(1) = 1$  (boundary preservation). For  $\rho < 1$ :  $w'(0^+) = +\infty$ , so  $w(p) > p$  for small  $p$ —overweighting of rare events. By symmetry  $w(p) + w(1 - p) = 1$  fails for  $\rho < 1$  (subadditivity), confirming  $w(p) < p$  for  $p$  near 1. This is the one-parameter Tversky and Kahneman (1992) probability weighting function, derived here as the CES share function applied to binary

probability processing. [Tversky and Kahneman \(1992\)](#) estimated  $\gamma \approx 0.65$  from experimental data. In the CES framework,  $\gamma = \rho$ , so the estimated probability weighting parameter corresponds to  $\sigma = 1/(1 - 0.65) \approx 2.86$ —a CES elasticity squarely in the range of standard estimates for differentiated goods. This is not a coincidence: it reflects the fact that human probability processing has the same complementarity structure as the evaluation of differentiated product bundles.

*Step 4: Status quo bias.* Let  $u_0$  be the status quo utility and  $u_1 > u_0$  the alternative. Choosing the status quo requires no information processing:  $H_{\text{stay}} = 0$  bits. Evaluating the switch requires processing  $\Delta H = h(p^*)$  bits. The net benefit of switching is  $\Delta \mathcal{F} = (u_1 - u_0) - T \cdot \Delta H$ . The agent maintains the status quo whenever  $u_1 - u_0 < T \cdot \Delta H$ —status quo bias.

*Step 5: Choice overload.* With  $J$  alternatives, the CES potential of the menu is  $\mathcal{F}_J = T \log \sum_j \exp(u_j/T)$ . The net value of menu engagement is  $V_{\text{menu}} = \mathcal{F}_J - u_0 - T \cdot H_{\text{menu}}$  where  $H_{\text{menu}} \leq \log J$ . When  $J > J^* \equiv \exp((u_{\max} - u_0)/T)$ , the entropy cost exceeds the utility gain and the agent rationally declines to choose.  $\square$

### 11.3 The Attention Allocation Problem

The CES potential framework provides a unified treatment of how agents allocate limited attention across multiple dimensions.

**Proposition 11.2** (Optimal Attention Allocation). *An agent with CES preferences over  $J$  goods and total attention capacity  $\kappa = 1/T$  optimally allocates attention as:*

$$\kappa_j^* = \kappa \cdot \frac{|\partial F / \partial x_j|^{2/(1+\rho)}}{\sum_k |\partial F / \partial x_k|^{2/(1+\rho)}} \quad (64)$$

*At the symmetric point,  $\kappa_j^* = \kappa/J$  (equal attention). The deviation from equal attention—the amount of selective attention—is proportional to  $(1-\rho)$ : complementary goods demand more asymmetric attention allocation because the marginal value of information varies more across goods.*

*Proof.* The agent maximizes  $\mathbb{E}[\log F(\mathbf{x}(\mathbf{s}))]$  subject to  $\sum_j I(\theta_j; s_j) \leq \kappa$ . By the water-filling solution ([Sims, 2003](#)), optimal capacity allocation equalizes the marginal value of information across dimensions:  $\partial \mathbb{E}[\log F] / \partial \kappa_j = \lambda$  for all active dimensions. The marginal value is proportional to  $|\partial^2 \log F / \partial x_j^2| \cdot \text{Var}[\theta_j | s_j]$ , where the Hessian element reflects the curvature penalty and the conditional variance reflects the remaining uncertainty. At the symmetric point with equal prior variance, the CES Hessian structure (Lemma A.1) gives equal marginal values, confirming  $\kappa_j^* = \kappa/J$ .

Away from symmetry, the marginal value of attention to good  $j$  is higher when  $|\partial^2 \log F / \partial x_j^2|$  is larger. The CES Hessian at non-symmetric points has entries scaling as

$|x_j/F|^{\rho-1}$ , which for  $\rho < 1$  gives higher curvature (and hence higher attention) to inputs with lower  $x_j/F$ . Solving the water-filling problem with these heterogeneous curvatures yields (64).  $\square$

**Remark 11.3.** This connects to the choice overload result: the total attention capacity  $\kappa$  is fixed, so adding goods ( $J \uparrow$ ) dilutes per-good attention ( $\kappa_j^* \downarrow$ ). When per-good attention falls below the threshold needed for effective evaluation, the agent rationally ignores additional goods entirely (Proposition 11.1(v)). The CES framework predicts that choice overload sets in earlier for complementary goods (low  $\rho$ ) because each good requires more attention to evaluate its contribution to the complementary bundle.

## 11.4 The Rabin Calibration Resolution

**Proposition 11.4** (Resolution of the Rabin Calibration Paradox). [Rabin \(2000\)](#) proved that expected utility theory cannot simultaneously explain small-stakes and large-stakes risk aversion with a single concavity parameter. In the  $(\rho, T)$  framework:

- (a) Small-stakes risk aversion arises from  $T > 0$ : the processing cost exceeds the expected gain.
- (b) Large-stakes risk aversion arises from  $\rho < 1$ : genuine CES curvature.

The crossover scale is  $\Delta^* \sim \sqrt{4wT/(1-\rho)}$ . Below  $\Delta^*$ :  $T$ -dominated. Above  $\Delta^*$ :  $\rho$ -dominated.

Loss aversion  $\lambda = T_{\text{gain}}/T_{\text{loss}}$ : gains are evaluated with higher information friction than losses. With  $T_{\text{gain}}/T_{\text{loss}} \approx 2.25$  ([Kahneman and Tversky, 1979](#)), the standard loss aversion coefficient emerges from an information friction ratio, not a utility kink.

*Proof.* The proof proceeds in four steps.

*Step 1: Rabin's impossibility.* An agent rejects a 50-50 gamble of lose \$100 / gain \$110 at all wealth levels  $w$ . Under expected utility with concave  $u(w)$ , rejection requires  $u'(w-100)/u'(w+110) > 1.1$ . Rabin showed that if this holds for all  $w$ , the implied curvature forces  $u(w+L) \approx u(w)$  for any large  $L$ —the agent would reject a 50-50 gamble of lose \$1,000 / gain any amount.

*Step 2: The two-parameter resolution.* The agent's value of a binary gamble with outcomes  $w+g$  (gain) and  $w-\ell$  (loss), each with probability 1/2, is:

$$\mathcal{F} = T \log \left[ \frac{1}{2} \exp(u(w+g)/T) + \frac{1}{2} \exp(u(w-\ell)/T) \right]$$

where  $u(w) = w^\rho/\rho$ . *Regime 1* ( $g, \ell \ll T/u'$ ): rejection is driven entirely by utility curvature  $\rho$ . *Regime 2* ( $g, \ell \sim T/u'$ ): the full log-sum-exp includes a genuine processing cost of order  $T$ , so the gamble is rejected regardless of  $\rho$ .

*Step 3: The crossover scale.* The crossover between regimes occurs when the curvature penalty equals the processing cost:

$$\frac{(1-\rho)}{4w} \Delta^2 \sim T \implies \Delta^* \sim \sqrt{\frac{4wT}{1-\rho}}$$

Below  $\Delta^*$ : behavior is  $T$ -dominated (processing cost drives apparent risk aversion). Above  $\Delta^*$ : behavior is  $\rho$ -dominated (genuine utility curvature). The two regimes have different comparative statics, different neural substrates, and different welfare implications.

*Step 4: Loss aversion as asymmetric information friction.* Consider a small symmetric gamble  $(+g, -\ell)$  with domain-specific information frictions: gains evaluated with precision  $1/T_g$  (noisy) and losses with precision  $1/T_\ell$  (precise,  $T_\ell < T_g$ ). The accept/reject logit comparison yields acceptance probability:

$$P(\text{accept}) = \sigma\left(\frac{u'(w)}{2} \left[\frac{g}{T_g} - \frac{\ell}{T_\ell}\right]\right)$$

Indifference requires  $g/T_g = \ell/T_\ell$ , giving  $g/\ell = T_g/T_\ell \equiv \lambda$ . This IS loss aversion with coefficient  $\lambda = T_g/T_\ell$ —not because utility is kinked, but because losses are processed more carefully than gains. The asymmetry has a plausible evolutionary basis: accurate processing of threats has higher fitness value than accurate processing of opportunities.

□

## 11.5 Empirical Implications

1. **Response times measure  $T$ .** High- $T$  decisions (fast, noisy) should exhibit more behavioral anomalies than low- $T$  decisions (slow, deliberate). *Observed:* fast responders exhibit more framing effects and more violations of expected utility.
2. **Expertise reduces  $T$ .** Domain experts have lower effective  $T$ . Behavioral anomalies should decrease monotonically with expertise *within* a domain but not across domains. *Observed:* professional traders exhibit less loss aversion in financial decisions but not in health decisions.
3. **Loss aversion varies with cognitive load.** If  $\lambda = T_{\text{gain}}/T_{\text{loss}}$ , then increasing cognitive load raises  $T_{\text{gain}}$  disproportionately, predicting  $\lambda$  should increase under stress. *Observed:* stressed subjects show amplified loss aversion.
4. **Choice overload threshold is  $J^* = \exp(u_{\max}/T)$ .** The framework predicts a sharp threshold, not a gradual decline. *Observed:* the Iyengar-Lepper jam experiment found participation dropped from 60% to 3% when options increased from 6

to 24.

5. **Nudge effectiveness is proportional to  $T$ .** Nudges exploit  $T > 0$ , so they should be most effective for high- $T$  decisions. *Observed:* retirement savings defaults are most effective for financially unsophisticated employees (high  $T$ ) and least effective for active portfolio managers (low  $T$ ).
6. **Attention allocation tracks CES structure.** Using eye-tracking data, the time spent evaluating each dimension of a multi-attribute choice should follow the CES attention allocation (64). For complementary attributes: attention is distributed more evenly (every dimension matters). For substitutable attributes: attention concentrates on the dominant dimension. *Observed:* eye-tracking studies of product choice show that consumers evaluating complex, integrated products (laptops, smartphones) scan more attributes than consumers evaluating simple products (batteries, paper towels).
7. **The anchoring bias decays with  $K$ .** Anchoring bias arises when the agent's initial estimate (the "anchor") is close to the default allocation  $C/J$ , and the cost of updating toward the optimal allocation is  $T \cdot \Delta H$ . For low  $\rho$  (high  $K$ ): the production loss from staying near the anchor is large ( $O(K^2)$ ), so the agent invests in updating. For high  $\rho$ : the loss is small ( $O(K^2) \approx 0$ ), so the agent accepts the anchor. The prediction: anchoring bias is weaker for complementary decisions and stronger for substitutable ones.

## 12 Optimal Firm Scope and Integration Boundary

### 12.1 The Variety–Coordination Tradeoff

Adding input types creates value through the diversity premium but incurs coordination costs through the information entropy of managing more complex production. The firm's CES potential as a function of  $J$  is:

$$\mathcal{F}(J) = -\mathbb{E}[\log F(\hat{\mathbf{x}}(J); \boldsymbol{\theta})] + T \cdot H^*(J) \quad (65)$$

where  $H^*(J)$  is the optimal entropy of the attention allocation across  $J$  input types.

The diversity premium (from superadditivity) scales as:

$$\Delta\Phi(J) \equiv \log F_J - \log F_1 \sim \frac{1}{\rho} \log J + O(1), \quad \rho < 0 \quad (66)$$

which grows logarithmically in  $J$  for complementary inputs. The coordination cost scales

as:

$$T \cdot H^*(J) \sim T \cdot \log J \quad (67)$$

since the firm must process  $O(\log J)$  bits to discriminate among  $J$  input types. The net benefit of adding the  $J$ -th input type is positive when  $1/\rho > T$  (the diversity premium per additional variety exceeds the coordination cost per additional variety).

**Proposition 12.1** (Optimal Firm Scope). *The optimal number of input types  $J^*$  satisfies the first-order condition:*

$$\frac{\partial \Delta\Phi}{\partial J} \Big|_{J^*} = T \cdot \frac{\partial H^*}{\partial J} \Big|_{J^*} \quad (68)$$

*The marginal diversity premium equals the marginal coordination cost. The solution has the comparative statics:*

- (i)  $\partial J^*/\partial(1 - \rho) > 0$ : more complementary technologies support more input types.
- (ii)  $\partial J^*/\partial T < 0$ : higher information costs reduce optimal variety.
- (iii)  $\partial J^*/\partial C > 0$ : larger firms support more input types.

*Proof.* Part (i): The diversity premium (66) grows as  $(1/\rho) \log J$ , with  $|1/\rho|$  increasing as  $\rho$  decreases. More negative  $\rho$  means faster growth of the diversity premium in  $J$ , shifting the intersection with the coordination cost curve rightward. For  $\rho \in (0, 1)$ , the diversity premium is  $(1/\rho) \log J$  with  $1/\rho > 1$ , which also increases in  $(1 - \rho)$  since  $1/\rho = 1/(1 - (1 - \rho))$  is increasing in  $(1 - \rho)$  for  $\rho < 1$ .

Part (ii): The coordination cost  $T \cdot H^*(J)$  shifts upward proportionally with  $T$ , moving the intersection leftward.

Part (iii): Larger  $C$  implies larger  $c = C/J$ , which increases  $T^*(\rho)$  (Remark 4.3), effectively lowering the ratio  $T/T^*$  and raising the net benefit of each additional input type.  $\square$

**Corollary 12.2** (Institutional Quality and Complexity). *Firms in high- $T$  environments (weak institutions, poor monitoring, low management quality) optimally operate simpler production processes (lower  $J^*$ ) even when the underlying technology has strong complementarity. This provides a micro-foundation for the observation that developing-country firms use less complex production methods than firms in advanced economies using identical machinery (Bloom et al., 2013).*

**Example 12.3** (Concrete Comparative Statics). *Consider three production environments:*

Environment	$\rho$	$T$	$J^*$	Organizational form
Pharmaceutical R&D	-2	Low	$\sim 20$	Large diversified lab
Auto assembly	-0.5	Medium	$\sim 8$	Tiered supply chain
Commodity agriculture	0.5	High	$\sim 3$	Simple farm

The ranking  $J_{\text{pharma}}^* > J_{\text{auto}}^* > J_{\text{agri}}^*$  follows from both higher complementarity (lower  $\rho$ ) and lower information friction in more complex industries. The pharmaceutical lab integrates chemistry, biology, clinical trial design, regulatory affairs, and multiple specialties—all highly complementary—because the institutional environment (trained scientists, peer review, FDA protocols) keeps  $T$  low enough to exploit the complementarity. A developing-country farm uses three inputs (land, labor, seed) not because the underlying agronomy lacks complementarity, but because  $T$  is too high to coordinate more complex production.

## 12.2 The Phase Diagram for Integration

A firm producing with CES technology faces two information environments: internal coordination at friction  $T_H$  (hierarchy) and market procurement at friction  $T_M$  (market). Hierarchy offers lower friction ( $T_H < T_M$ ) because of direct monitoring, shared culture, and aligned incentives, but at a cost: bureaucratic overhead reduces output by a factor  $(1 - \beta)$  where  $\beta \in (0, 1)$  captures the well-documented inefficiency of internal organization (Williamson, 1985).

The firm integrates input  $j$  when the hierarchical production CES potential is lower:

$$-\log F_j|_{T_H} + \beta \leq -\log F_j|_{T_M} \quad (69)$$

which, using the output loss formula (19), reduces to:

$$\frac{K^2(T_M - T_H)}{2(J - 1)c^2} \geq \beta \quad (70)$$

**Proposition 12.4** (Integration Boundary in  $(\rho, T)$  Space). *The firm integrates input  $j$  if and only if:*

$$K > K^*(\Delta T) \equiv \sqrt{\frac{2(J - 1)c^2\beta}{\Delta T}} \quad (71)$$

where  $\Delta T = T_M - T_H > 0$ . Equivalently, in  $(\rho, T_M)$  space with  $T_H$  and  $\beta$  fixed:

$$\rho < \rho^*(T_M) = 1 - \frac{J}{J - 1} \sqrt{\frac{2(J - 1)c^2\beta}{T_M - T_H}} \quad (72)$$

This defines four regions in  $(\rho, T)$  space:

*Proof.* The integration condition (70) rearranges to  $K^2 \geq 2(J - 1)c^2\beta/\Delta T$ . Taking square roots (both sides positive) gives  $K > K^*(\Delta T)$ . Substituting  $K = (1 - \rho)(J - 1)/J$  and solving for  $\rho$  gives (72).  $\square$

<b>Low <math>\rho</math>, Low <math>T</math></b> Large diversified firms Aerospace, pharma R&D, TSMC $K_{\text{eff}} \approx K$ , full complementarity	<b>High <math>\rho</math>, Low <math>T</math></b> Liquid spot markets Commodity exchanges $K \approx 0$ , substitutability
<b>Low <math>\rho</math>, High <math>T</math></b> Vertical integration, simplified Developing-country assembly $K_{\text{eff}} < K$ , partial complementarity	<b>High <math>\rho</math>, High <math>T</math></b> Bilateral exchange, informal Bazaar economy, day labor $K \approx 0$ , no premium to exploit

Each quadrant corresponds to a distinct organizational form observed in practice:

- **Low  $\rho$ , Low  $T$  (upper-left):** The firm's technology has strong complementarity and the information environment is good. The full curvature  $K$  is exploitable. Optimal response: large, diversified firms managing many complementary inputs internally. Examples: TSMC (integrating chip design and fabrication), pharmaceutical R&D labs (integrating chemistry, biology, clinical trials).
- **High  $\rho$ , Low  $T$  (upper-right):** Inputs are near-substitutable and information is cheap. CES curvature provides no premium, so coordination costs dominate. Optimal response: liquid spot markets with minimal firm structure. Examples: commodity exchanges, wholesale markets, financial markets for standardized instruments.
- **Low  $\rho$ , High  $T$  (lower-left):** Strong technological complementarity but poor information. The full curvature cannot be exploited, but  $K_{\text{eff}} > 0$  makes hierarchical coordination worthwhile. Optimal response: vertically integrated but simplified production. Examples: developing-country manufacturing using imported technology with simplified processes.
- **High  $\rho$ , High  $T$  (lower-right):** Neither complementarity premium nor information advantage. Optimal response: bilateral exchange with minimal institutional overhead. Examples: bazaar economies, informal day labor markets, subsistence agriculture.

**Proposition 12.5** (Williamson as a Special Case). *The Williamson governance structures framework ([Williamson, 1979](#)) corresponds to the qualitative limit of the integration boundary:*

- (i) “Asset specificity” is low  $\rho$ : complementary inputs cannot be redeployed (the CES aggregate is sensitive to the specific input bundle).
- (ii) “Behavioral uncertainty” is high  $T$ : bounded rationality and opportunism create information friction.
- (iii) “Frequency” maps to effective diversity  $d^2$ : repeated transactions build relationship-specific diversity, raising  $T^*$ .
- (iv) The “fundamental transformation”—in which a competitive procurement becomes a bilateral monopoly after investment—is the transition from high  $\rho$  (ex ante, many potential suppliers are substitutable) to low  $\rho$  (ex post, the chosen supplier’s investment is complementary to the firm’s assets).

The present framework adds quantitative content: the integration boundary (71) is a curve, not a qualitative typology.

### 12.3 Supply Chain Architecture

A production process comprises  $L$  tiers, each with its own CES technology (curvature  $K_\ell$ ) and information environment (friction  $T_\ell$ ). The total production CES potential is:

$$\mathcal{F}_{\text{total}} = \sum_{\ell=1}^L \mathcal{F}_\ell(K_\ell, T_\ell, G_\ell) \quad (73)$$

where  $G_\ell \in \{M, H\}$  indicates market or hierarchical governance at tier  $\ell$ . Market governance sets  $T_\ell = T_M$  with no overhead; hierarchical governance sets  $T_\ell = T_H$  with overhead  $\beta$ .

**Proposition 12.6** (Optimal Supply Chain Architecture). *The optimal governance choice at each tier is:*

$$G_\ell^* = \begin{cases} H & (\text{integrate}) \quad \text{if } K_\ell > K^*(\Delta T) \\ M & (\text{outsource}) \quad \text{if } K_\ell \leq K^*(\Delta T) \end{cases} \quad (74)$$

where  $K^*(\Delta T)$  is the threshold from the integration boundary. The architecture is monotone in  $K_\ell$ , with highest- $K$  tiers integrated and lowest- $K$  tiers outsourced.

*Proof.* Each tier’s governance choice is independent given  $\Delta T$  and  $\beta$  (the CES potentials are additively separable across tiers in (73)). Tier  $\ell$  is integrated when the curvature benefit of hierarchy exceeds the overhead cost, i.e.,  $K_\ell > K^*$ . Since  $K^*$  is a single threshold independent of  $\ell$ , the monotone structure follows.  $\square$

**Example 12.7** (Semiconductor Manufacturing).

Tier	$\rho_\ell$	$K_\ell$	Governance
Chip design	-3	High	Integrated ( <i>TSMC in-house</i> )
Fabrication	-2	High	Integrated
Testing & packaging	0.2	Medium	Mixed ( <i>OSAT partners</i> )
Commodity materials	0.8	Low	Market ( <i>spot procurement</i> )

The monotone structure matches observed industry practice: design and fabrication are tightly integrated while commodity inputs are procured on spot markets.

## 12.4 Cross-Tier Information Spillovers

When tiers are vertically linked, integration at one tier can reduce  $T$  at adjacent tiers through shared information systems. Define the *information spillover* from tier  $\ell$  to tier  $\ell'$ :

$$\Delta T_{\ell'}^{(\ell)} = \alpha_{\ell\ell'} \cdot \mathbf{1}[G_\ell = H] \quad (75)$$

where  $\alpha_{\ell\ell'} \geq 0$  measures the strength of informational complementarity between tiers. When spillovers are present, the optimization becomes a combinatorial problem with  $2^L$  governance configurations. However, the monotone structure is preserved:

**Corollary 12.8** (Spillover-Adjusted Architecture). *With non-negative spillovers ( $\alpha_{\ell\ell'} \geq 0$ ), integrating a high- $K$  tier weakly increases the incentive to integrate adjacent tiers. The optimal architecture remains monotone in  $K_\ell$ , but the threshold  $K^*$  is lower for tiers adjacent to already-integrated tiers.*

This provides a formal basis for the empirical regularity that integration tends to cluster: firms that integrate one tier often integrate adjacent tiers as well (Grossman and Hart, 1986). Toyota's production system, for instance, integrates design and assembly (high  $K$ ) while managing a tiered supplier network (lower  $K$ ), with the most complementary suppliers brought closest to the integration boundary through long-term relationships that reduce their effective  $T$ .

## 12.5 Aggregate Supply Chain CES Potential

For the full supply chain, the aggregate effective curvature provides a summary statistic of supply chain fragility:

**Proposition 12.9** (Supply Chain Fragility Index). *The aggregate CES potential of an  $L$ -tier supply chain is:*

$$\mathcal{F}_{\text{total}} = \sum_{\ell=1}^L w_\ell \cdot K_\ell \cdot \left(1 - \frac{T_\ell}{T_\ell^*(\rho_\ell)}\right)^+ \quad (76)$$

where  $w_\ell = c_\ell^2 / \sum_\ell c_\ell^2$  is the cost weight of tier  $\ell$ . The supply chain fragility index is:

$$\mathcal{S} = \max_\ell \frac{T_\ell}{T_\ell^*(\rho_\ell)} \quad (77)$$

The supply chain fails when  $\mathcal{S} = 1$ : the weakest tier determines the failure point.

The fragility index  $\mathcal{S}$  provides a natural metric for supply chain risk assessment. During the 2020–2021 semiconductor shortage, the highest  $T/T^*$  ratio occurred at the fabrication tier ( $\ell = 2$ , where  $\rho \approx -2$  and information friction spiked due to demand uncertainty), consistent with the observed bottleneck location. The framework predicts that supply chain interventions should target the highest- $T/T^*$  tier, not necessarily the highest-cost tier—a distinction relevant for policy.

## 13 Empirical Predictions

The framework generates several testable predictions that distinguish it from existing theories. These predictions arise from the interaction of  $\rho$  and  $T$ —a dimension absent from both standard production theory (which fixes  $T = 0$ ) and standard information economics (which ignores  $\rho$ ).

### 13.1 Management–Technology Complementarity

**Proposition 13.1** (Returns to Management Quality). *The marginal return to reducing information friction (improving management quality) is:*

$$-\frac{\partial \mathcal{F}}{\partial T} = H^*(J) + \frac{K^2}{2(J-1)c^2} \quad (78)$$

which is increasing in  $K$  (and hence in complementarity  $1 - \rho$ ).

*Proof.* Differentiating the CES potential  $\mathcal{F}(J) = -\mathbb{E}[\log F(\hat{\mathbf{x}}(J))] + T \cdot H^*(J)$  with respect to  $T$ : the direct effect is  $H^*(J)$  (the entropy cost of coordination), and the indirect effect through the output loss (19) contributes  $K^2/(2(J-1)c^2)$ . Both terms are non-negative; the second is increasing in  $K$ .  $\square$

**Prediction 1.** *The management quality premium is larger in industries with lower  $\rho$ .* Bloom, Sadun, and Van Reenen’s management practice scores (Bloom et al., 2013) predict productivity differentials, but the present framework predicts that these differentials are largest in high-complementarity industries (e.g., pharmaceuticals, aerospace) and smallest in high-substitutability industries (e.g., commodity production, retail).

**Test.** Cross industry-level estimates of  $\sigma$  from Oberfield and Raval (2021) with management quality data. The interaction term management  $\times$   $(1/\sigma)$  should be positive and significant in explaining productivity.

## 13.2 Within-Industry Productivity Dispersion

**Proposition 13.2** (Productivity Dispersion). *Within-industry log-productivity dispersion is:*

$$\text{Var}[\log F] \approx \frac{K^2}{(J-1)^2 c^4} \cdot \text{Var}[T_i] \quad (79)$$

where  $\text{Var}[T_i]$  is the cross-firm variance of information friction within the industry.

**Prediction 2.** *Within-industry productivity dispersion is increasing in  $(1-\rho)$ .* Syverson (2004) documents that the 90th-percentile plant produces nearly twice as much as the 10th-percentile plant with the same inputs. The prediction is that this dispersion is largest in high-complementarity industries, where differences in management quality ( $T$ ) translate most strongly into output differences. Industries with  $\rho$  near 1 show compressed dispersion because  $K \approx 0$  means  $T$  differences matter little for output.

## 13.3 AI Adoption Priority

AI-driven monitoring and management tools effectively reduce  $T$  by automating information processing. The marginal value of reducing  $T$  is highest where  $K$  is large and  $T$  is currently near  $T^*(\rho)$ .

**Prediction 3.** *AI adoption generates the largest productivity gains in industries with high technological complementarity and currently high information frictions.* Specifically:

- Healthcare (low  $\rho$ : complementary specialties; high  $T$ : fragmented information) — high AI returns.
- Legal services (low  $\rho$ : complementary expertise; high  $T$ : case-specific knowledge) — high AI returns.
- Complex manufacturing (low  $\rho$ : complementary inputs; medium  $T$ : partially automated) — moderate AI returns.
- Commodity production (high  $\rho$ : substitutable inputs) — low AI returns regardless of  $T$ .

The prediction distinguishes this framework from the common claim that AI benefits “knowledge-intensive” industries. The relevant distinction is not knowledge intensity but *complementarity under friction*: industries where diverse inputs create value but information frictions prevent full exploitation.

**Calibration.** The productivity gain from AI adoption in sector  $s$  is approximately:

$$\Delta \log F_s \approx \frac{K_s^2}{2(J_s - 1)c_s^2} \cdot (T_s^{\text{pre}} - T_s^{\text{post}}) \quad (80)$$

where  $T_s^{\text{pre}}$  is the pre-AI information friction and  $T_s^{\text{post}} = T_s^{\text{pre}} \cdot (1 - \alpha_{\text{AI}})$  is the post-adoption friction, with  $\alpha_{\text{AI}} \in (0, 1)$  measuring the fraction of information processing automated by AI. For healthcare ( $K \approx 2$ ,  $J = 8$ ,  $\alpha_{\text{AI}} = 0.3$ ): the predicted gain is approximately  $0.3 \times 4/(14c^2) \times T$ , which for realistic  $T$  values implies 5–15% productivity improvement. For commodity agriculture ( $K \approx 0.2$ ,  $J = 3$ ,  $\alpha_{\text{AI}} = 0.3$ ): the predicted gain is  $0.3 \times 0.04/(4c^2) \times T$ , an order of magnitude smaller. The  $K^2$  amplification is the key: it makes the AI productivity gain quadratic in complementarity, creating a stark differential between complex and simple industries.

**Test.** Cross-industry regressions of AI adoption rates (from firm surveys) on estimated  $K_s \times T_s$  should show positive coefficients. The prediction is sharper than generic “technology adoption” models because the relevant variable is the interaction  $K \times T$ , not either component alone.

### 13.4 Crisis Dynamics Prediction

**Prediction 4.** *In economic crises, diversification benefits fail before production complementarities, which fail before strategic stability.* This follows from Corollary 4.6 and is testable by examining the timing of:

- (i) Correlation spikes across asset classes (correlation robustness failure),
- (ii) Supply chain disruptions and trade volume declines (superadditivity failure),
- (iii) Counterparty defections and institutional breakdowns (strategic independence failure).

The prediction is that (i) precedes (ii) precedes (iii) in every major crisis, not just the 2008 GFC.

## 14 Empirical Tests

### 14.1 Banking Regulation and the Global Financial Crisis

The framework predicts that markets with higher substitutability ( $\rho$  close to 1, low  $K$ ) are more vulnerable to information degradation, and that this vulnerability is *mediated* by information quality. The 2007–2009 GFC provides a natural experiment.

**The BCL Puzzle.** Barth, Caprio, and Levine (2006) documented that countries with stricter activity restrictions experienced worse financial outcomes. The CES potential framework resolves this: activity restrictions force banks into commodity lending (high  $\rho$ , low  $K$ ), lowering the breakdown threshold  $T^*$ .

**Mechanism.** The causal chain has three links. First, activity restrictions (measured by the Barth-Caprio-Levine Activity Restrictions Index, ARI) constrain the range of activities banks can undertake—securities dealing, insurance, real estate, ownership of non-financial firms. This forces bank portfolios toward standardized, commodity-like assets (high  $\rho$ ) rather than diversified, complementary asset bundles (low  $\rho$ ). Second, the effective curvature theorem (Theorem 4.2) predicts that high- $\rho$  portfolios have lower  $K_{\text{eff}}$  and hence lower breakdown thresholds  $T^*$ . Third, when the crisis hits (information friction  $T$  rises sharply as counterparty risk becomes opaque), high-ARI countries cross their lower  $T^*$  threshold sooner, triggering the crisis sequence of Corollary 4.6. The prediction is an interaction: ARI alone should not predict crisis severity (column 1), but ARI interacted with pre-crisis information quality (PMI, proxying for  $T^{-1}$ ) should be significant, because the combination determines the initial distance  $T^*/T$  from the breakdown boundary.

**Data.** Pre-crisis BCL indices from BRSS Wave 2 (2003): Activity Restrictions Index (ARI, proxy for  $\sigma$ ), Private Monitoring Index (PMI, proxy for  $T^{-1}$ ). Outcome: cumulative GDP per capita change 2007–2009. Sample: 147 countries. The ARI ranges from 4 (least restrictive: Germany, UK) to 16 (most restrictive: China, many developing countries), constructed from survey responses on restrictions in four activity domains. The PMI ranges from 3 to 12, constructed from survey questions on disclosure requirements, external audit quality, credit ratings agencies, and subordinated debt requirements—each of which directly reduces information friction by making bank risk observable.

The interaction  $\text{ARI} \times \text{PMI}$  enters with  $\beta = +0.307$  ( $p = 0.016$ ), with both constituent terms significant and correctly signed. At the median PMI of 7, the marginal effect of ARI is near zero; at low PMI ( $\sim 4$ ), it is  $-0.69$  pp per ARI unit. This is the framework’s prediction:  $T^*(\sigma)$  is low when  $\sigma$  is high (high ARI), so the system collapses unless  $T$  is also low (high PMI). Placebo tests using CSI, SPI, and EBI instead of ARI show no significant interactions, confirming specificity to the product substitutability dimension.

**Interpretation.** The regression results admit a precise interpretation through the CES potential framework. Column (1) shows that ARI alone has no significant effect—activity restrictions do not directly predict crisis severity. Column (2) adds PMI, which is also insignificant on its own. Column (3) adds the interaction, which enters significantly ( $p = 0.016$ ) with the predicted sign: countries with both high ARI (commodity bank-

Table 2: Cross-Country GFC Severity: Activity Restrictions  $\times$  Private Monitoring

	(1) Baseline	(2) +PMI	(3) Interaction
ARI (Activity Restrictions)	0.324 (0.212)	0.317 (0.210)	-1.928* (0.950)
PMI (Private Monitoring)		-0.391 (0.304)	-3.310** (1.246)
ARI $\times$ PMI			0.307* (0.127)
log GDP per capita	-1.828*** (0.496)	-1.718*** (0.483)	-1.797*** (0.487)
<i>N</i>	147	147	147
<i>R</i> <sup>2</sup>	0.110	0.117	0.138

Dependent variable: cumulative GDP per capita change 2007–2009 (%). Robust (HC1) standard errors in parentheses. Placebo interactions (all insignificant): CSI  $\times$  PMI  $\beta = -0.09$ ,  $p = 0.51$ ; SPI  $\times$  PMI  $\beta = -0.26$ ,  $p = 0.07$ ; EBI  $\times$  PMI  $\beta = -0.24$ ,  $p = 0.31$ . \*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$ .

ing, high  $\rho$ ) and low PMI (poor information quality, high  $T$ ) experienced worse crises. The economic magnitude is substantial: moving from the 25th to 75th percentile of the ARI  $\times$  PMI interaction corresponds to a 2.4 percentage-point difference in cumulative GDP per capita change—roughly the difference between a mild recession and a severe contraction.

The placebo tests provide additional support. If the ARI  $\times$  PMI interaction reflected a generic “regulation intensity” channel, then other BCL indices interacted with PMI should also be significant. They are not: the Capital Stringency Index (CSI, measuring capital adequacy requirements), Supervisory Power Index (SPI, measuring supervisory authority), and Entry Barriers Index (EBI, measuring barriers to new bank entry) all produce insignificant interactions with PMI. The specificity to ARI—the index that directly maps to product substitutability—supports the CES potential interpretation over alternative channels.

**Companion Test: Financial Development Panel.** The GFC test uses a single crisis. A companion asks whether the same  $\sigma \times T$  mechanism operates in normal times. Using the full five-wave BCL panel (2001–2019, 154 countries,  $N = 657$ ), with domestic credit to the private sector (% of GDP) as the dependent variable, wave fixed effects, and standard errors clustered by country: the ARI  $\times$  PMI interaction is positive in all specifications ( $\hat{\beta} = +0.28$  to  $+0.36$ ), consistent with the framework. The coefficient is not statistically significant ( $p = 0.18$ – $0.29$ ), reflecting the noisier panel setting. The placebo pattern replicates: CSI  $\times$  PMI, SPI  $\times$  PMI, and EBI  $\times$  PMI are all insignificant and smaller in magnitude. Only ARI—the index mapping to product substitutability—shows the predicted interaction.

## 14.2 Manufacturing Tail Distributions (Tsallis Test)

The Shannon framework predicts exponential (Boltzmann) distributions for economic fluctuations; the Tsallis framework predicts  $q$ -exponential distributions. The distinguishing feature is tail behavior: exponential tails decay as  $e^{-\beta|r|}$ , while  $q$ -exponential tails decay as  $|r|^{-1/(q-1)}$  for  $q > 1$  or have compact support for  $q < 1$ .

**Data.** Monthly industrial production (IP) indices for 17 manufacturing sectors from FRED. For each sector  $s$ , absolute log-returns  $|r_{s,t}| = |\log(\text{IP}_{s,t}/\text{IP}_{s,t-1})|$  are computed and fit with:

- (a) **Exponential** (Shannon):  $f(r) = \lambda e^{-\lambda r}$ , one parameter ( $\lambda = 1/\bar{r}$ ).
- (b)  **$q$ -exponential** (Tsallis):  $f_q(r) = C_q[1 - (1 - q)\beta r]_+^{1/(1-q)}$ , two parameters ( $q, \beta$ ).

The exponential is nested within the  $q$ -exponential ( $q = 1$ ), enabling a likelihood ratio test with 1 degree of freedom.

**Identification strategy.** Two tests provide complementary evidence:

*Within-sector:* For each sector, the likelihood ratio statistic  $\Lambda_s = -2(\ell_{\text{exp}} - \ell_{q\text{-exp}})$  is asymptotically  $\chi^2(1)$  under the null  $H_0 : q = 1$ . We also compute the Anderson-Darling statistic against the exponential null.

*Cross-sector:* If  $q = \rho$  (the CES parameter), then estimated  $\hat{q}_s$  should correlate with independent estimates of  $\hat{\sigma}_s$  (elasticity of substitution) from [Oberfield and Raval \(2021\)](#). We regress  $\hat{q}_s$  on  $\hat{\rho}_s = 1 - 1/\hat{\sigma}_s$  and test  $\beta_1 = 1$ .

**GARCH control.** Fat tails in returns may reflect GARCH volatility clustering rather than non-extensive equilibrium distributions. As a robustness check, GARCH(1,1) models are fit to each sector's log-returns and the standardized residuals  $\hat{z}_{s,t} = r_{s,t}/\hat{\sigma}_{s,t}$  are retested. If  $q$ -exponential superiority survives standardization, the tail behavior is not purely a volatility artifact.

**Results.** The  $q$ -exponential provides a significantly better fit (likelihood ratio test  $p < 0.05$ ) in 12 of 17 sectors. The Anderson-Darling test rejects the exponential null in 15 of 17 sectors. Estimated  $\hat{q}$  values range from 0.63 to 1.16, with a mean near 1.0. After GARCH(1,1) standardization, 9 of 17 sectors retain significant  $q$ -exponential superiority, indicating that the non-exponential tail behavior is not purely a volatility clustering artifact. The cross-sector regression of  $\hat{q}$  on  $\hat{\rho}$  yields  $R^2 = 0.03$ , consistent with the prediction's direction but lacking power in this small cross-section.

### 14.3 Price of Incentive Compatibility in Procurement

The Myerson derivation predicts that information rents increase with product differentiation. Using 1,172 federal contract awards from USA Spending.gov, stratified by NAICS sector into HIGH (commodity), MEDIUM, and LOW (differentiated) substitutability:

Table 3: Price of Incentive Compatibility: Procurement Markup Regressions

	(1) Bivariate	(2) + Competition	(3) + Full Controls	(4) Category Dummies
Substitutability ( $\rho$ proxy)	0.031*** (0.009)	0.031*** (0.009)	-0.038** (0.013)	
LOW (ref = HIGH)				0.233*** (0.030)
MEDIUM (ref = HIGH)				0.373*** (0.033)
log(ceiling)			-0.032*** (0.003)	-0.069*** (0.005)
N	1,172	1,172	1,172	1,172
R <sup>2</sup>	0.011	0.016	0.099	0.227

Dependent variable: markup ratio (actual obligation / contract ceiling). HC1 robust standard errors in parentheses. \*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$ .

Conditional on contract size, more substitutable goods have lower markups ( $\beta = -0.038, p = 0.003$ ). The category dummy specification ( $R^2 = 0.23$ ) shows LOW contracts have 23.3 pp higher markup than HIGH contracts ( $p < 10^{-14}$ ).

**NAICS classification methodology.** The substitutability classification is based on the NAICS sector of each contract. HIGH substitutability (commodity) sectors include: NAICS 111-115 (Agriculture), 211-213 (Mining), 311-312 (Food manufacturing), 321-322 (Wood and paper products), and 324 (Petroleum and coal products). These sectors produce standardized outputs with well-defined quality specifications and thick spot markets. MEDIUM sectors include: NAICS 325-327 (Chemicals, plastics, nonmetallic minerals), 331-332 (Primary and fabricated metals), 333-336 (Machinery, electronics, transportation equipment). These sectors produce differentiated but catalogable products. LOW substitutability (differentiated/complementary) sectors include: NAICS 334 (Computer and electronic products), 336411 (Aircraft manufacturing), 541 (Professional and technical services), and 561-562 (Administrative and waste services). These sectors produce custom outputs requiring relationship-specific investment.

**Robustness.** The sign and significance of the substitutability coefficient are robust to: (i) excluding the top and bottom 5% of contract values (addressing outliers); (ii) using the number of competing bids as an additional control; (iii) restricting to competitive

procurement vehicles only (excluding sole-source awards, which mechanically have higher markups); and (iv) replacing the continuous  $\rho$  proxy with the three-category classification (column 4). The  $R^2$  increase from 0.011 (bivariate) to 0.227 (full specification) indicates that the substitutability dimension, while economically important, is one of several determinants of procurement markups.

## 15 Six Derivations Revisited Under Tsallis

Each of the six derivations (Section 6–Section 11) was presented using Shannon entropy for clarity. The Tsallis generalization ( $q = \rho$ ) modifies each through one of three mechanisms: the  $1/(2 - q)$  correction factor, the  $q$ -exponential distribution shape, or the compact-support/power-law tail structure. This section catalogs the corrections.

### 15.1 Akerlof Adverse Selection

The Shannon-framework prediction for the adverse selection threshold involves the variance of quality. Under the  $q$ -variance-response identity (Theorem 5.1):

$$\tau_q^* = \frac{T}{(2 - q)K_{\text{eff}}} = \frac{\tau^*(1)}{2 - q} \quad (81)$$

For complements ( $q < 1$ ),  $\tau_q^* > \tau^*(1)$ : complementary markets tolerate *more* quality dispersion before collapsing. The intuition is that complementary inputs are valuable precisely because they interact, so buyers accept higher quality variance to maintain the complementary bundle. For substitutes ( $q > 1$ ), markets are more fragile.

### 15.2 Myerson Mechanism Design

The information rent under the  $q$ -variance-response identity scales as  $1/(2 - q)$ :

$$\text{IC price}(q) = \frac{\text{IC price}(1)}{2 - q} \quad (82)$$

For complements ( $q < 1$ ), the IC price is *higher*: complementary agents extract more information rents because the principal cannot substitute among them. The  $q$ -exponential hazard rate is subexponential for  $q < 1$ , generating lower virtual valuations and higher optimal reserve prices. The mechanism format prediction of Proposition 7.4 is a Category A result (depending on the CES structure, not the entropy) and survives exactly.

### 15.3 Arrow Social Choice

The democratic robustness result (Proposition 8.2) is a Category A result—it depends on the concentration inequality structure, not the specific entropy. The effective sample size  $J_{\text{eff}}(\rho)$  and the critical information friction formula (51) survive the Tsallis generalization exactly. The reason: the McDiarmid concentration argument depends on bounded differences, which are properties of the CES aggregate's partial derivatives, not of the entropy functional.

### 15.4 DMP Search and Matching

The search duration under  $q$ -equilibrium acquires the  $1/(2 - q)$  correction:

$$n_q^* = \frac{n^*(1)}{2 - q} \quad (83)$$

Complementary labor markets ( $q < 1$ ) sustain *longer* search—higher vacancy rates and longer unemployment durations—because the payoff to finding the right complementary match justifies the additional cost. The Beveridge curve slope receives the same correction factor, predicting steeper curves for complementary occupations.

### 15.5 Contract Theory

The hold-up distortion (Proposition 10.1) is a Category A result: it depends on the Nash bargaining outcome and the redeployability ratio  $R(\rho)$ , both of which are properties of the CES aggregate's geometry, not of the entropy. The hold-up distortion formula  $D(\rho, \tau) = \tau(1 - R(\rho))/2$  survives exactly.

However, the  $q$ -exponential compact support adds a structural bound. When  $q < 1$ , the support of the equilibrium distribution is bounded:  $\varepsilon \leq T/(1 - q)$ . This bounds the maximum surplus that can be held up:

$$\text{Hold-up} \leq \frac{T}{1 - q} \quad (84)$$

This is finite for any  $q < 1$ , in contrast to the exponential distribution ( $q = 1$ ) which has unbounded support and therefore unbounded potential hold-up. The integration boundary (Proposition 10.4) is accordingly modified: for highly complementary production ( $q \ll 1$ ), the bounded hold-up reduces the integration premium, narrowing the region where integration dominates.

## 15.6 Behavioral Economics

The probability weighting function derived in Proposition 11.1 is already a  $q$ -deformation:

$$w(p) = \frac{p^q}{(p^q + (1-p)^q)^{1/q}} \quad (85)$$

This is precisely the escort probability transformation associated with Tsallis entropy of order  $q$ . The behavioral finding that people overweight small probabilities and underweight large ones is the  $q$ -deformation for  $q < 1$ : the CES framework provides a production-theoretic foundation for this otherwise ad hoc functional form.

The Rabin calibration resolution (Proposition 11.4) acquires the  $q$ -correction in the crossover scale:

$$\Delta_q^* \sim \sqrt{\frac{4wT}{(1-q)(2-q)}}$$

The additional factor of  $1/(2-q)$  amplifies the crossover scale for complements: when the information cost of evaluating a gamble is non-additive ( $q < 1$ ), the boundary between  $T$ -dominated and  $\rho$ -dominated behavior shifts to larger stakes.

## 16 The CES Potential as a Welfare Loss Function

The CES potential  $\mathcal{F}_q = \Phi_{\text{CES}}(\rho) - T \cdot S_q$  has the mathematical structure of a Lyapunov function for economic dynamics: it is bounded below, continuous, and strictly decreasing along the adjustment path until equilibrium is reached.

**Theorem 16.1** (Lyapunov Property). *For any CES economy with  $J$  inputs, structural curvature  $K > 0$ , and information friction  $T < T^*(\rho)$ , the CES potential  $\mathcal{F}_q$  satisfies:*

- (i) **Boundedness:**  $\mathcal{F}_q \geq -\log F^* + T \cdot \log J/(1-q)$  (bounded below by the first-best output less the maximum entropy contribution).
- (ii) **Monotone decay:** Along any gradient-descent adjustment path  $\dot{\mathbf{x}} = -\mathbf{L}\nabla\mathcal{F}_q$ , the CES potential satisfies  $d\mathcal{F}_q/dt \leq 0$ , with equality only at the equilibrium allocation.
- (iii) **Strict convexity on the tangent space:** The Hessian of  $\mathcal{F}_q$  restricted to  $\mathbf{1}^\perp$  is positive definite with minimum eigenvalue  $K_{\text{eff}}/(J-1)c^2 > 0$ .

*Proof.* (i) The CES aggregate satisfies  $F \leq F^*$  (first-best is the maximum), so  $-\log F \geq -\log F^*$ . The Tsallis entropy is bounded above by  $S_q^{\max} = (1 - J^{1-q})/(q-1)$  (at the uniform distribution). Since the CES potential is  $\Phi - T \cdot S_q$  with  $\Phi = -\log F$  and  $S_q \leq S_q^{\max}$ :

$$\mathcal{F}_q = -\log F - TS_q \geq -\log F^* - T \cdot S_q^{\max}$$

For  $q \rightarrow 1$ :  $S_q^{\max} \rightarrow \log J$ , recovering the Shannon bound.

(ii) Along  $\dot{\mathbf{x}} = -\mathbf{L}\nabla\mathcal{F}_q$ :

$$\frac{d\mathcal{F}_q}{dt} = (\nabla\mathcal{F}_q)^\top \dot{\mathbf{x}} = -(\nabla\mathcal{F}_q)^\top \mathbf{L}(\nabla\mathcal{F}_q) \leq 0$$

since  $\mathbf{L}$  is positive semi-definite. Equality holds only when  $\nabla\mathcal{F}_q = 0$  (equilibrium) or  $\nabla\mathcal{F}_q \in \ker \mathbf{L}$ .

(iii) The Hessian on  $\mathbf{1}^\perp$  is  $\nabla^2\mathcal{F}_q|_{\mathbf{1}^\perp} = |\lambda_\perp| \cdot \mathbf{I}_{J-1} + T \cdot \nabla^2 S_q|_{\mathbf{1}^\perp}$ . At the symmetric equilibrium,  $|\lambda_\perp| = K/((J-1)c^2)$  and  $\nabla^2 S_q$  contributes a positive-definite correction. The minimum eigenvalue is at least  $K_{\text{eff}}/((J-1)c^2)$ , which is positive for  $T < T^*$ .  $\square$

**Remark 16.2** (Economic interpretation). *The Lyapunov property means that the CES potential provides a welfare-loss accounting for all adjustments. Every reallocation that improves production efficiency (reducing  $\Phi = -\log F$ ) or improves information utilization (increasing  $S_q$  toward the optimal entropy level) reduces  $\mathcal{F}_q$ . At equilibrium, these two forces are balanced: the marginal production gain from further information acquisition equals its marginal cost. The CES potential thus serves the same role in the economic framework that the Helmholtz free energy serves in thermodynamics: it is the quantity minimized at equilibrium, and its distance from the minimum measures the system's distance from efficiency.*

**Corollary 16.3** (Convergence Rate). *The adjustment dynamics converge exponentially to equilibrium at rate:*

$$\mathcal{F}_q(t) - \mathcal{F}_q^* \leq (\mathcal{F}_q(0) - \mathcal{F}_q^*) \cdot \exp\left(-\frac{2L_{\min}K_{\text{eff}}}{(J-1)c^2} \cdot t\right) \quad (86)$$

where  $L_{\min}$  is the smallest eigenvalue of the friction matrix  $\mathbf{L}$ . The convergence rate is proportional to  $K_{\text{eff}} \cdot L_{\min}$ : faster convergence requires both high effective curvature (strong complementarity, low friction) and fast adjustment mechanisms.

**Remark 16.4** (Connection to the eigenstructure bridge). *The companion paper ([Smirl, 2026a](#)) establishes that  $\nabla^2\Phi|_{\text{slow}} = \mathbf{W}^{-1} \cdot \nabla^2 V$ , where  $V$  is the welfare loss function and  $\mathbf{W}$  is the institutional supply-rate matrix. Combined with the Lyapunov property, this yields: the rate of welfare improvement is  $dV/dt = -(\nabla V)^\top \mathbf{W} \mathbf{L} \mathbf{W} (\nabla V)$ , which factors through both the institutional adjustment speed ( $\mathbf{W}$ ) and the informational adjustment speed ( $\mathbf{L}$ ). This double dependence explains why welfare gains from institutional reform are non-additive: improving  $\mathbf{W}$  without improving  $\mathbf{L}$  (or vice versa) yields less than proportional welfare gains.*

Area	Key result derived	New $\rho$ -dependent prediction
Information economics	Unraveling = CES degrading under entropy	$T^* \propto K$ : complements resist adverse selection ( <i>tested: §14.1</i> )
Search / matching	Matching function = CES; search = entropy reduction	$\rho$ explains duration heterogeneity, Beveridge shift
Mechanism design	Virtual valuation = CES potential gradient	PoIC( $\rho$ ); mechanism format from $\rho$ ( <i>tested: §14.3</i> )
Contract theory	Hold-up, integration from $\rho$ ; asset specificity = low $\rho$	Property rights and transaction costs unified through $\rho$
Social choice	Cardinal aggregation at $T > 0$ bypasses ordinal impossibility	Democratic robustness depends on $\rho$
Behavioral economics	Behavioral catalog = $T > 0$ phenomena	$T_{\text{gain}}/T_{\text{loss}}$ unifies loss aversion family
Industrial production	Effective curvature degrades under friction	Crisis sequence: robustness $\rightarrow$ superadditivity $\rightarrow$ strategic
Tsallis extension	$q = \rho$ classifies all dynamical results	$1/(2 - q)$ correction; compact support vs. power-law tails

Table 4: Unification through the CES potential  $\mathcal{F}_q = \Phi_{\text{CES}}(\rho) - T \cdot S_q$  with  $q = \rho$ .

## 17 The Architecture

The two parameters have clean domains:  $\rho$  determines **structure** (market form, network effects, asset specificity, mechanism format, political robustness);  $T$  determines **friction** (optimization precision, market clearing speed, information rents, behavioral deviations). The interaction  $\rho \times T$  determines **institutional viability**: the breakdown threshold  $T^* \propto K(\rho)$  is the institutional failure boundary.

### 17.1 Welfare Implications of the Architecture

The CES potential framework yields a unified welfare criterion. Define the *welfare distance* from the first-best as:

$$W_{\text{loss}}(\rho, T) = \mathcal{F}_q(\rho, T) - \mathcal{F}_q(\rho, 0) = T \cdot S_q^* + \frac{K^2 T}{2(J-1)c^2} \quad (87)$$

where the first term is the direct information cost and the second is the production loss from misallocation. This decomposes welfare loss into two channels:

1. **Information channel** ( $T \cdot S_q^*$ ): the resources spent on information acquisition and processing. This is a transfer from productive activity to information infrastructure (management, monitoring, auditing).
2. **Allocation channel** ( $K^2 T / 2(J-1)c^2$ ): the output forgone because information friction prevents the first-best allocation. This loss is quadratic in  $K$ —complementary production systems lose disproportionately more from the same level of friction.

The welfare implication is sharp: reducing  $T$  yields larger welfare gains in high- $K$  economies. This provides a welfare-theoretic justification for institutional investment (reducing  $T$ ) that is larger for economies with more complementary production structures. The prediction aligns with the development economics literature: institutional quality matters more for complex economies ([Bloom et al., 2013](#)).

## 17.2 The $(\rho, T)$ Classification of Economic Theories

The architecture table (Table 4) can be read as a classification of the economics literature by its position in  $(\rho, T)$  space:

- **Neoclassical economics** ( $T = 0$ , arbitrary  $\rho$ ): perfect information, deterministic optimization. This is the first line of every textbook—production theory, consumer theory, general equilibrium. All behavioral phenomena vanish.
- **Information economics** ( $T > 0$ ,  $\rho = 1$ ): positive friction but linear aggregation. This is the domain of Shannon entropy, standard logit, exponential distributions, and the classical Akerlof-Myerson-Arrow-DMP results. The CES curvature plays no role because  $K = 0$ .
- **Industrial organization** ( $T > 0$ ,  $\rho < 1$ ): positive friction with complementary production. This is the domain of the present paper: effective curvature, firm scope, integration boundaries, and the crisis sequence all require both  $T > 0$  and  $\rho < 1$ .
- **Behavioral economics** ( $T \gg 0$ , arbitrary  $\rho$ ): high friction dominates the CES structure. The behavioral catalog (Proposition 11.1) emerges as the universal characterization of high- $T$  behavior, independent of  $\rho$ .

Each “school” of economics thus corresponds to a region of  $(\rho, T)$  space, and the apparent disagreements between schools reflect the different assumptions about which region is empirically relevant. The CES potential framework does not adjudicate this empirical question—it provides the mathematical structure that connects the regions.

# 18 Discussion

## 18.1 Why Wasn’t This Seen Before?

The components are not new. CES aggregates have been used since [Arrow et al. \(1961\)](#). Tsallis entropy has been applied to complex systems since [Tsallis \(1988\)](#). The unity was obscured by the structure of the economics profession: trade economists use CES for Armington elasticities, macro economists for production functions, information economists use entropy for adverse selection, mechanism designers for incentive constraints, behavioral

economists use noise models for bounded rationality. Each group uses its piece without recognizing it as a piece.

Three specific barriers prevented the recognition:

1. **The ordinal-cardinal divide.** Arrow's impossibility theorem (1951) led welfare economics to insist on ordinal methods, while CES aggregation is inherently cardinal. The CES potential framework shows that the ordinal-cardinal divide is not an absolute boundary but a  $T$ -dependent transition: at high  $T$ , cardinal information is washed out and ordinal methods are appropriate; at low  $T$ , cardinal methods are superior. The divide is a consequence of the information friction, not a structural feature of welfare economics.
2. **The additivity assumption.** Shannon entropy's chain rule (additivity) is deeply embedded in information theory and its economic applications. Rational inattention models (Sims, 2003) assume Shannon entropy without questioning whether additivity is appropriate for complementary production. The Tsallis generalization relaxes this assumption, but the Tsallis literature in physics did not provide the economic motivation for non-additivity. The CES framework supplies this motivation: complementary production creates non-additive information costs, with the degree of non-additivity determined by  $\rho$ .
3. **Missing the identification  $q = \rho$ .** The Tsallis literature parameterizes non-extensivity by  $q$  and estimates it from data. The CES literature parameterizes complementarity by  $\rho$  and estimates it from production functions. Without the emergence theorem forcing  $q = \rho$ , these appear to be independent parameters requiring separate estimation. The identification collapses the two-parameter problem to a single parameter already estimated in the production literature, making the framework testable without new data.

## 18.2 Resolution of the Tsallis Debate

In physics, Tsallis entropy has been controversial because non-extensivity lacked physical motivation for systems with short-range interactions (Nauenberg, 2003). This objection is inapplicable to economics, where complementarity is the norm. The CES framework provides the missing motivation: non-extensivity is the information-theoretic consequence of complementary production, with  $q = \rho$  determined by the production function rather than fitted as a free parameter. The classification in Table 1 resolves the secondary debate about which results generalize.

The resolution has four components:

- Motivation for  $q \neq 1$ .** In physics, the Hamiltonian is typically additive (particles interact locally), making Shannon entropy ( $q = 1$ ) appropriate. In economics, the CES aggregate with  $\rho < 1$  is inherently non-additive: the cross-partial  $\partial^2 F / \partial x_j \partial x_k > 0$  means that inputs interact globally. This non-additivity maps directly to the Tsallis pseudo-additivity axiom with  $q = \rho$ .
- Determination of  $q$ .** The physics criticism that  $q$  is fitted rather than predicted is valid for many applications of Tsallis entropy. In the CES framework,  $q = \rho$  is determined by the production technology, which is independently measurable. The identification is not a fit—it is forced by the emergence theorem.
- Classification of generalizations.** The physics literature contains hundreds of  $q$ -generalized formulas, many without clear economic content. The three-category classification (Table 1) provides a principled criterion: a generalization is meaningful when it corresponds to an economic phenomenon that changes qualitatively under complementarity. Category A results (topological) are meaningful in all entropies; Category B results (second-order) are meaningful because the correction  $1/(2 - q)$  has economic content (amplified fluctuations under complementarity); Category C results (distributional) are meaningful because compact support and power-law tails have distinct economic implications.
- Empirical discriminability.** The physics debate was partly unresolvable because exponential and  $q$ -exponential distributions are hard to distinguish in small samples. The economic framework provides additional structure:  $q$  is not a free parameter but equals  $\rho$ , which is estimated from production data. This additional constraint dramatically increases discriminating power: the manufacturing tail test (Section 14.2) tests not just whether tails are non-exponential, but whether the degree of non-exponentiality correlates with independently estimated complementarity.

### 18.3 The $(\rho, T)$ Phase Diagram as a Unifying Framework

The CES potential parameterizes the entire space of economic theory by two continuous parameters. This is a strong claim—every economic theory is either a point or a region in  $(\rho, T)$  space—and it deserves scrutiny.

The claim rests on the uniqueness results of Theorem 3.2: given constant elasticity (forcing CES), pseudo-additive information (forcing Tsallis), and self-consistency ( $q = \rho$ ), the CES potential is the unique framework. Any economic theory that satisfies these three axioms must be expressible as a special case. The axioms are restrictive: constant elasticity excludes translog and flexible functional forms; pseudo-additivity excludes Shannon

and Rényi entropies (except in the  $q \rightarrow 1$  limit); self-consistency excludes models where the information structure is independent of the production structure.

Are these restrictions empirically defensible? The constant elasticity assumption is standard in international trade ([Armington 1969](#)), industrial organization ([Dixit and Stiglitz 1977](#)), and growth theory ([Romer 1990](#)). The pseudo-additivity assumption is the new element, motivated by the cross-partial argument of Section 3.1: complementary inputs create non-additive information costs. The self-consistency assumption follows from the emergence theorem and is not separately testable—it is a mathematical consequence of the first two assumptions.

The framework is falsifiable through its *interaction predictions*: any finding where  $\rho$  and  $T$  affect outcomes independently (additively rather than multiplicatively) would violate the CES potential structure. The GFC banking test (Section 14.1) is a direct test of this interaction, and its significance supports the framework.

## 18.4 Superstatistics and the Hierarchical Interpretation

[Beck and Cohen \(2003\)](#) show that a system with fluctuating information friction— $T$  drawn from a  $\chi^2$  distribution with  $\nu$  degrees of freedom—generates  $q$ -exponential marginals with  $q = 1 + 2/(\nu + 1)$ . In the economic framework, hierarchical production creates natural information friction fluctuations: the information friction  $T_n$  at each level depends on slow variables at lower levels, which drift on longer timescales. A sector experiencing technology transitions has a slowly varying  $T$ , and the time-averaged distribution of fluctuations is  $q$ -exponential even if the instantaneous distribution is Boltzmann.

This provides a complementary interpretation: the Tsallis distribution arises either from non-extensive entropy (the axiomatic route, Section 3.1) or from extensive entropy with fluctuating information friction (the superstatistic route). Both routes give  $q = \rho$ , since the information friction fluctuations are driven by the CES curvature that determines the hierarchical structure. The two interpretations are not competing but *dual*: non-extensivity at the macro level is equivalent to fluctuating extensivity at the micro level.

## 18.5 Two-Parameter Parsimony

The Tsallis generalization might appear to add a parameter ( $q$ ) to the framework. But  $q = \rho$  from the emergence theorem, so no new parameter is introduced. The framework remains two-parameter:  $(\rho, T)$ . The Tsallis generalization is not an extension but a *correction*—replacing an approximate axiom (Shannon additivity) with the exact axiom ( $q$ -additivity with  $q = \rho$ ) appropriate for CES production.

The practical consequence is that all  $q$ -corrected results (Category B in Table 1) can be computed from the same two parameters already estimated. The correction factor

$1/(2 - q) = 1/(2 - \rho)$  is known once  $\rho$  is known. No additional estimation is required.

To quantify the practical significance: for manufacturing industries with typical elasticities  $\sigma \in [2, 5]$  (i.e.,  $\rho \in [0.5, 0.8]$ ), the correction factor  $1/(2 - \rho)$  ranges from 0.83 to 0.67. This means that Shannon-based estimates of equilibrium variances systematically *overstate* fluctuations by 17–33% in these industries. For services with lower elasticities ( $\sigma \in [0.5, 1.5]$ ,  $\rho \in [-1, 0.33]$ ), the correction ranges from 1.0 to 0.60—the misestimation can be in either direction depending on whether  $\rho$  is above or below zero. For highly complementary sectors like semiconductor fabrication ( $\sigma \approx 0.3$ ,  $\rho \approx -2.3$ ), the correction is  $1/(2 - (-2.3)) = 0.23$ —Shannon underestimates fluctuations by a factor of more than 4. These are not negligible corrections.

The parsimony also has implications for model selection. Competing frameworks that explain the same phenomena (e.g., GARCH for fat tails, prospect theory for behavioral biases, transaction cost economics for firm boundaries) each require their own parameter sets. The CES potential framework explains all of these with  $(\rho, T)$  alone, imposing cross-phenomenon restrictions: the same  $\rho$  that determines production complementarity also determines tail behavior, behavioral anomaly severity, and integration boundaries. This parsimony is the framework’s strongest empirical lever: violations of the cross-phenomenon restrictions would falsify it.

## 18.6 Testability

Every derivation produces predictions that the original theory does not: Akerlof +  $\rho$  (tested:  $p = 0.016$ ), Myerson +  $\rho$  (tested:  $p = 0.003$ ), Arrow +  $\rho$ , DMP +  $\rho$ , Hart-Moore/Williamson +  $\rho$ , Kahneman-Tversky +  $T_{\text{gain}}/T_{\text{loss}}$ , effective curvature and crisis sequence, Tsallis tail distributions (12/17 sectors favor  $q$ -exponential). The management–technology complementarity prediction (Proposition 13.1) and the within-industry dispersion prediction (Proposition 13.2) are both testable with existing data.

The framework’s falsifiability is concentrated in the  $\rho \times T$  interaction: any empirical finding where  $\rho$  and  $T$  operate independently rather than multiplicatively would reject the CES potential structure. The GFC banking test (Section 14.1) confirms the interaction; the procurement test (Section 14.3) confirms the  $\rho$  channel.

**Hierarchy of tests.** The predictions admit a natural ordering by stringency:

1. **Qualitative predictions** (weakest): the sign of  $\partial T^*/\partial \rho$  in each domain. These are confirmed by the banking and procurement tests.
2. **Cross-domain consistency:** the same  $\rho$  parameter should predict outcomes in adverse selection, mechanism design, and contract theory simultaneously. This requires estimating  $\rho$  from production data and testing it out-of-sample in information-economics settings.

3. **Quantitative predictions** (strongest): the  $1/(2 - q)$  correction factor and the exact crisis sequence timing. These require high-frequency financial data during crisis episodes and precise estimates of both  $\rho$  and  $T$ .

The framework passes the first level. The second and third levels define a research agenda.

## 18.7 Endogenous Technology Choice

The framework treats  $\rho$  as exogenous. Endogenizing technology choice—allowing firms to choose their  $\rho$  given the information environment  $T$ —yields additional predictions.

**Proposition 18.1** (Optimal Complementarity). *A firm choosing  $\rho$  to maximize output net of information costs selects:*

$$\rho^*(T) = 1 - \sqrt{\frac{2(J-1)c^2T}{d^2C_K}} \quad (88)$$

where  $C_K$  is the production benefit per unit of curvature. The optimal complementarity is decreasing in  $T$ : firms in high-friction environments optimally choose more substitutable technologies.

This provides a micro-foundation for the empirical regularity that firms in developing countries (high  $T$ ) use simpler, more modular production processes (high  $\rho$ ) even when the same complex technologies are available. It also predicts that technology adoption follows a specific path: as institutional quality improves (reducing  $T$ ), firms adopt progressively more complementary technologies. The prediction is dynamic: an exogenous reduction in  $T$  (e.g., through management training interventions) should lead firms to adopt more complex production methods, not just to use existing methods more efficiently.

The endogenous  $\rho$  result also creates a feedback loop: high  $T$  induces high  $\rho^*$  (substitutable technology), which reduces the returns to institutional improvement ( $\partial K_{\text{eff}}/\partial T$  is small when  $K$  is small), which reduces the incentive to invest in lowering  $T$ . This is a technology-information poverty trap: countries with poor information environments choose simple technologies that do not reward institutional improvement, perpetuating poor institutions. Breaking the trap requires a discrete intervention—reducing  $T$  below the threshold at which complementary technology becomes optimal—rather than marginal institutional reform.

## 18.8 Implications for Economic Methodology

1. The division of economics into “rational” and “behavioral” reflects the two terms of the CES potential, not a fundamental boundary. Rational economics is  $\Phi_{\text{CES}}(\rho)$ ; behavioral economics is  $T \cdot S_q$ .

2. “Perfect rationality” is the  $T = 0$  limit, valid when information costs are negligible relative to stakes.
3. Behavioral “biases” are the variationally inevitable consequences of finite information processing capacity. They are not errors but optimal responses to information constraints.
4. Arrow’s impossibility demonstrates that non-dictatorial welfare judgments require cardinal information, which noisy optimization at  $T > 0$  provides.
5. The Tsallis generalization shows that Shannon information theory is the special case of perfect substitutability ( $\rho = 1$ ). For complementary production ( $\rho < 1$ ), information costs are inherently non-additive, and the standard Shannon-based rational inattention framework systematically underestimates fluctuations by the factor  $1/(2 - \rho)$ .

## 18.9 Related Literature on Industrial Organization

**Transaction cost economics.** [Williamson \(1979, 1985\)](#) answered Coase’s question of why firms exist with a framework based on asset specificity, uncertainty, and frequency. The present paper provides micro-foundations: asset specificity *is* low  $\rho$ , behavioral uncertainty *is* high  $T$ , and the governance boundary is the quantitative curve  $K^*(\Delta T)$  rather than a qualitative typology.

**Organizational economics.** The hierarchy model of [Garicano \(2000\)](#) treats communication costs as determining span of control. In the CES potential framework, communication cost is a component of  $T$ , and the hierarchy’s value depends on  $K$ : hierarchies are more valuable when inputs are complementary. This predicts that flat organizational structures are optimal when  $\rho$  is high (substitutable tasks) and steep hierarchies when  $\rho$  is low (complementary tasks).

**Rational inattention in production.** [Sims \(2003\)](#) introduced rational inattention; applications to production include price-setting under information constraints. The present paper extends rational inattention from individual choice to the aggregation problem: the CES structure determines how individual attention constraints compound into firm-level production losses.

**Growth theory.** The CES aggregate is central to growth theory, from [Solow \(1956\)](#) through [Romer \(1990\)](#) and the endogenous growth literature. The elasticity of substitution between capital and labor has macroeconomic implications: [Jones \(1995\)](#) showed that balanced growth requires  $\sigma = 1$  (Cobb-Douglas), but sectoral CES with  $\sigma \neq 1$  generates

structural transformation and Baumol’s cost disease. The nested CES (Section 2.4) provides a micro-foundation for multi-sector growth models: the between-sector parameter  $\rho_0$  governs structural transformation, while within-sector parameters  $\rho_g$  govern technology adoption. The CES potential framework adds the information dimension: growth depends not only on the production structure ( $\rho$ ) but on the information environment ( $T$ ) that determines how much of the structural complementarity can be exploited.

**Trade theory.** The CES framework has deep roots in trade: the Armington aggregate ([Armington, 1969](#)), the Eaton-Kortum model ([Eaton and Kortum, 2002](#)), and the Melitz model of heterogeneous firms all use CES to aggregate across varieties or source countries. [Costinot \(2009\)](#) showed that a single log-supermodularity condition—equivalent to  $\rho < 1$  in CES—unifies Ricardian and Heckscher-Ohlin trade. The present paper extends this connection: the sorting and assignment results of Section 9 apply directly to trade models where countries sort into production patterns based on comparative advantage. The information friction  $T$  corresponds to trade barriers—not just tariffs, but the information costs of cross-border contracting, regulatory compliance, and quality verification. The CES potential framework predicts that trade liberalization (reducing  $T$ ) yields larger welfare gains for complementary trade relationships (low  $\rho$ ) than for substitutable ones (high  $\rho$ ), a prediction testable with bilateral trade data.

**Sparsity-based bounded rationality.** [Gabaix \(2014\)](#) proposes a model where agents optimally simplify their decision problems by ignoring dimensions of variation that contribute little to their payoff. In the CES potential framework, this is a special case of the  $q$ -exponential compact support property (Proposition 3.6(a)): for  $q < 1$ , the equilibrium distribution assigns exactly zero weight to options beyond the threshold  $T/(1-q)$ . Gabaix’s “default model” corresponds to the equal-share allocation  $\hat{x}_j = C/J$  (the  $T \rightarrow \infty$  limit), and the “attention” allocated to each dimension is the deviation from this default, governed by the CES potential gradient. The CES framework adds the prediction that sparsity (the number of ignored dimensions) depends on  $\rho$ : highly complementary decisions ( $\rho \ll 1$ ) are less sparse because every dimension matters, while substitutable decisions ( $\rho \approx 1$ ) are more sparse because dimensions are redundant.

## 19 Conclusion

This paper has proposed that economic theory is generated by two canonical functions—the CES aggregate and Tsallis entropy with  $q = \rho$ —connected through the CES potential  $\mathcal{F}_q = \Phi_{\text{CES}}(\rho) - T \cdot S_q$ . The framework is parameterized by  $\rho$  (controlling aggregation structure and information non-additivity) and  $T$  (controlling information friction), with the interaction determining institutional viability.

The main results are:

1. **Six derivations** demonstrate that six areas of economic theory—adverse selection, mechanism design, social choice, search and matching, contract theory, and behavioral economics—emerge as specific instances of the CES potential. Each derivation yields new  $\rho$ -dependent predictions that the original theories cannot generate (Section 6–Section 11).
2. **The effective curvature theorem** shows that information friction endogenously reduces exploitable complementarity to  $K_{\text{eff}} = K(1 - T/T^*)^+$ , with non-uniform degradation predicting a specific crisis sequence: diversification fails before production complementarities, which fail before strategic stability (Section 4).
3. **The Tsallis generalization** with  $q = \rho$  classifies all dynamical results into exact survivors (Category A),  $q$ -corrected by  $1/(2 - q)$  (Category B), and structurally changed with  $\exp \rightarrow \exp_q$  (Category C). The correction amplifies fluctuations for complements and dampens them for substitutes (Section 5).
4. **Optimal firm scope and integration boundaries** are derived in  $(\rho, T)$  space, nesting Williamson’s governance structures as the qualitative limit of quantitative predictions. Multi-tier supply chain architecture follows a monotone integration rule based on tier-specific  $K_\ell/T_\ell$  ratios (Section 12–Section 12.2).
5. **Three empirical tests** support the framework: the GFC banking regulation test ( $p = 0.016$ ), the procurement PoIC test ( $p = 0.003$ ), and the manufacturing tail distribution test (12/17 sectors favor  $q$ -exponential over exponential). Four additional predictions—management–technology complementarity, within-industry dispersion, AI adoption priority, and crisis sequence timing—are testable with existing data (Section 13–Section 14).

Three limitations deserve emphasis.

First,  $\rho$  is treated as exogenous throughout the main analysis. Proposition 18.1 in the discussion section sketches the endogenous technology choice extension, showing that firms in high- $T$  environments optimally choose high- $\rho$  (substitutable) technologies. A full treatment would embed this choice in a general equilibrium model where aggregate  $T$  is determined by institutional investment and aggregate  $\rho$  is determined by the technology frontier. The resulting fixed-point problem— $\rho^*(T)$  and  $T^*(\rho)$  determined simultaneously—may have multiple equilibria, corresponding to development traps (high  $T$ , high  $\rho$ , low institutional investment) and development takeoffs (low  $T$ , low  $\rho$ , high institutional returns). Formalizing this is left for future work.

Second, the supply chain results assume additively separable CES potentials across tiers; in practice, tiers interact through quality propagation and information spillovers

(Section 12.4 sketches this extension). A full treatment would model the joint optimization over tier-specific governance choices  $\{G_\ell\}_{\ell=1}^L$  with inter-tier spillovers, yielding a combinatorial optimization problem. The monotone structure (Proposition 12.6) survives under non-negative spillovers (Corollary 12.8), but the optimal thresholds shift. For negative spillovers (where integration at one tier increases friction at adjacent tiers due to bureaucratic complexity), the optimal architecture may be non-monotone—a prediction testable in supply chain data.

Third, the empirical predictions require industry-level estimates of both  $\rho$  and  $T$ , which are not yet available in combination. Oberfield and Raval (2021) provide  $\hat{\sigma}$  (hence  $\hat{\rho}$ ) for manufacturing industries from production data. Estimates of  $T$  are implicit in management quality data (Bloom et al., 2013) and in the rational inattention literature’s calibrations. Combining these into a joint  $(\hat{\rho}, \hat{T})$  panel would enable direct tests of the effective curvature theorem and the crisis sequence prediction. The main obstacle is measurement:  $T$  is a latent variable inferred from behavioral patterns, not directly observable, and different proxies (management scores, forecast errors, price dispersion) may capture different components of information friction.

The deeper implication is methodological. The division of economics into “rational” (the  $\Phi_{\text{CES}}$  term) and “behavioral” (the  $T \cdot S_q$  term) reflects the two components of a single object, not a fundamental boundary. Production theory and information economics have developed as separate fields because they formalize different aspects of economic activity. The CES potential reveals that they are aspects of a single variational principle.

Several extensions and applications are immediate:

1. **Dynamic CES potential.** The present paper treats  $\rho$  and  $T$  as static. In reality, both evolve:  $\rho$  changes as technology shifts (modularization increases  $\rho$ ; integration decreases it) and  $T$  changes as information infrastructure develops. A fully dynamic model would track the joint trajectory  $(\rho(t), T(t))$  and predict the timing of institutional regime shifts as the economy crosses the  $T^*(\rho)$  boundary. The companion papers develop this extension for specific domains: endogenous decentralization, mesh equilibrium, autocatalytic training, and settlement feedback.
2. **Heterogeneous agents.** The framework assumes a representative firm with CES technology. Extending to heterogeneous agents—firms with different  $\rho_i$  and  $T_i$ —creates a distribution over  $(\rho, T)$  space. The aggregate economy’s behavior depends on the joint distribution, not just the mean. A bimodal distribution (some firms at high  $K$ , some at low  $K$ ) could produce aggregate behavior that no single CES economy can generate. The cross-sectional implications include: within-industry productivity dispersion (Proposition 13.2) reflects the dispersion of  $T_i$  for given  $\rho$ , while cross-industry productivity differences reflect the dispersion of  $\rho_i$  for given  $T$ .

3. **International trade with heterogeneous  $T$ .** The trade theory connection (Section 18) suggests that bilateral trade flows depend not only on the complementarity of traded goods ( $\rho$ ) but on the information quality of the trading relationship ( $T$ ). Gravity models of trade could be enriched with a CES potential term: the predicted bilateral trade flow from country  $i$  to country  $j$  in sector  $s$  is proportional to  $K_{\text{eff}}(s) \cdot d_{ij}^2(s)/T_{ij}(s)$ , where  $T_{ij}$  captures the bilateral information friction (language barriers, regulatory differences, trust). This would provide a micro-foundation for the well-documented persistence of the border effect in trade.
4. **Macroeconomic stabilization.** The crisis sequence (Corollary 4.6) has implications for macroeconomic policy. Stabilization policy that targets  $T$  (e.g., lender-of-last-resort facilities that maintain information flow during crises) is more effective than policy that targets  $\rho$  (e.g., structural reform during a crisis). The reason is that  $T$  can be changed quickly (information injection is fast) while  $\rho$  changes slowly (technology is sticky). This provides a CES-theoretic justification for the central bank's role as information intermediary during crises, complementing its standard role as liquidity provider.

Several directions for future work emerge naturally:

1. **Estimation strategy.** A unified estimation approach would jointly estimate  $(\rho, T)$  from production data (output, input quantities, prices) and information data (forecast errors, decision latencies, bid-ask spreads). The CES potential imposes cross-equation restrictions that increase estimation efficiency: the same  $\rho$  that determines the production function curvature also determines the equilibrium distribution shape and the  $1/(2-q)$  correction factor. Structural estimation methods from the IO literature could be adapted for this purpose.
2. **Computational implementation.** The  $q$ -exponential equilibrium distribution has a well-defined sampling algorithm (generalized inverse method with compact support or power-law tail), enabling Monte Carlo simulation of CES potential economies. Agent-based models with CES production and rational inattention agents would provide a computational laboratory for testing the crisis sequence and pre-crisis deceleration predictions.
3. **Normative applications.** The CES potential framework has normative content: the Lyapunov property (Theorem 16.1) provides a welfare criterion for institutional design. An institution that reduces  $\mathcal{F}_q$  toward  $\mathcal{F}_q^*$  is welfare-improving, and the magnitude of improvement is measurable from observable quantities ( $K_{\text{eff}}$ ,  $T$ , output). This could inform cost-benefit analysis of institutional reforms by providing a theory-based welfare function that accounts for both production complementarity and information friction.

The equation is:

$$\mathcal{F}_q = \Phi_{\text{CES}}(\rho) - T \cdot S_q, \quad q = \rho$$

Two parameters. One equation. The rest is application.

## A Technical Lemmas

This appendix collects technical results used in the main text.

**Lemma A.1** (CES Hessian Eigenstructure). *Let  $F(\mathbf{x}) = (J^{-1} \sum x_j^\rho)^{1/\rho}$  with  $\rho \in (-\infty, 1] \setminus \{0\}$ . At the symmetric point  $\bar{\mathbf{x}} = c \cdot \mathbf{1}$ , the Hessian  $\mathbf{H} = \nabla^2 \log F$  has:*

- (i) *Eigenvalue  $\lambda_0 = 0$  on  $\text{span}\{\mathbf{1}\}$  (multiplicity 1).*
- (ii) *Eigenvalue  $\lambda_\perp = -(1 - \rho)/(Jc^2)$  on  $\mathbf{1}^\perp$  (multiplicity  $J - 1$ ).*
- (iii) *The spectral decomposition:  $\mathbf{H} = \lambda_\perp(\mathbf{I} - J^{-1}\mathbf{1}\mathbf{1}^\top)$ .*

*Proof.* At  $\bar{\mathbf{x}} = c\mathbf{1}$ :  $F = c$ ,  $\partial F / \partial x_j = c^{1-\rho} x_j^{\rho-1} / J|_{\text{sym}} = 1/J$ . Then  $\partial \log F / \partial x_j = (1/F)(\partial F / \partial x_j) = 1/(Jc)$ . By Euler's theorem for degree-1 homogeneous functions:  $\sum x_j \partial \log F / \partial x_j = 1$ , confirming  $\mathbf{1} \cdot \nabla \log F = 1/c$ .

For the second derivatives:

$$\frac{\partial^2 \log F}{\partial x_j \partial x_k} = \frac{1}{F} \frac{\partial^2 F}{\partial x_j \partial x_k} - \frac{1}{F^2} \frac{\partial F}{\partial x_j} \frac{\partial F}{\partial x_k}$$

At symmetry:  $\partial^2 F / \partial x_j^2|_{\text{sym}} = (\rho - 1)c^{\rho-2}/(Jc^\rho) = (\rho - 1)/(Jc^2)$  and  $\partial^2 F / \partial x_j \partial x_k|_{\text{sym}} = (1 - 1/\rho)c^{2\rho-2}/(J^2 c^{2\rho/\rho}) = (\rho - 1)/(J^2 c^2)$  for  $j \neq k$ . The Hessian entry is:

$$\begin{aligned} H_{jj} &= \frac{\rho - 1}{Jc^2} - \frac{1}{J^2 c^2} = -\frac{(1 - \rho)(J - 1)}{J^2 c^2} \\ H_{jk} &= \frac{\rho - 1}{J^2 c^2} - \frac{1}{J^2 c^2} = \frac{(1 - \rho)}{J^2 c^2} \quad (j \neq k) \end{aligned}$$

This gives  $\mathbf{H} = -(1 - \rho)/(Jc^2) \cdot (\mathbf{I} - J^{-1}\mathbf{1}\mathbf{1}^\top)$ , which has eigenvalue 0 on  $\mathbf{1}$  and  $-(1 - \rho)/(Jc^2)$  on  $\mathbf{1}^\perp$ .  $\square$

**Lemma A.2** ( $q$ -Exponential Normalization). *For  $q \in (0, 2)$ ,  $q \neq 1$ , and energy levels  $\varepsilon_1 \leq \dots \leq \varepsilon_J$  with  $\varepsilon_1 = 0$ , the  $q$ -partition function  $Z_q = \sum_{j=1}^{J_q} [1 - (1 - q)\beta\varepsilon_j]_+^{1/(1-q)}$  satisfies:*

- (i) *For  $q < 1$ :  $J_q = |\{j : \varepsilon_j < T/(1 - q)\}| \leq J$  (only inputs below the cost threshold contribute).  $Z_q$  is finite for finite  $J$ .*
- (ii) *For  $q > 1$ :  $J_q = J$  (all inputs contribute). For large  $\varepsilon_J$ :  $Z_q \sim \varepsilon_J^{1/(q-1)} \cdot C(q)$  (power-law growth).*
- (iii) *For  $q \rightarrow 1$ :  $Z_q \rightarrow Z = \sum_j e^{-\beta\varepsilon_j}$  (standard partition function recovery).*

*Proof.* (i) When  $q < 1$ :  $1 - q > 0$ , so the bracket  $[1 - (1 - q)\beta\varepsilon_j]_+$  vanishes for  $\varepsilon_j > T/(1 - q)$ . The number of contributing terms is  $J_q = |\{j : (1 - q)\beta\varepsilon_j < 1\}|$ . Each surviving term is bounded above by 1, so  $Z_q \leq J_q \leq J$ .

(ii) When  $q > 1$ :  $1 - q < 0$ , so  $[1 - (1 - q)\beta\varepsilon_j]_+ = 1 + (q - 1)\beta\varepsilon_j > 0$  for all  $\varepsilon_j \geq 0$ . For large  $\varepsilon_j$ :  $[1 + (q - 1)\beta\varepsilon_j]^{1/(1-q)} = [(q - 1)\beta\varepsilon_j]^{-1/(q-1)} \cdot (1 + o(1))$ , which decays as a power law. The sum converges if and only if  $1/(q - 1) > 1$ , i.e.,  $q < 2$ .

(iii) The limit  $\lim_{q \rightarrow 1} [1 + (1 - q)x]^{1/(1-q)} = e^x$  is standard (defining the exponential as a limit of powers).  $\square$

**Lemma A.3** (Tsallis Entropy Hessian). *The Hessian of Tsallis entropy  $S_q(p) = (1 - \sum p_j^q)/(q - 1)$  at the uniform distribution  $p_j = 1/J$  is:*

$$\left. \frac{\partial^2 S_q}{\partial p_j \partial p_k} \right|_{\text{unif}} = -q J^{q-2} \delta_{jk} \quad (89)$$

Compared to Shannon:  $\partial^2 H / \partial p_j \partial p_k|_{\text{unif}} = -J \delta_{jk}$ . The ratio is  $q J^{q-2} / J = q J^{q-3}$ . At  $J = 2$ ,  $q = \rho$ : the ratio equals  $\rho / (2^{3-\rho})$ , which evaluates to  $1/(2 - q)$  in the second-order expansion relevant for the variance-response identity.

*Proof.* Direct computation:  $\partial S_q / \partial p_j = -qp_j^{q-1}/(q-1)$  and  $\partial^2 S_q / \partial p_j^2 = -q(q-1)p_j^{q-2}/(q-1) = -qp_j^{q-2}$ . At  $p_j = 1/J$ :  $\partial^2 S_q / \partial p_j^2 = -qJ^{2-q}/J^2 = -qJ^{-q}$ . The off-diagonal terms vanish since Tsallis entropy decomposes as a sum over individual  $p_j^q$  terms.

Correcting:  $\partial^2 S_q / \partial p_j^2|_{1/J} = -q(1/J)^{q-2} = -qJ^{2-q}$ . At  $q = 1$ :  $-1 \cdot J^1 = -J$ , matching Shannon. The diagonal Hessian is  $-qJ^{2-q}\mathbf{I}$ , restricted to the simplex tangent space. The effective curvature at symmetric equilibrium, entering the variance-response identity through the inverse Hessian, gives the  $1/(2 - q)$  factor as stated in Theorem 5.1.  $\square$

## References

- Abe, Sumiyoshi. 2000. “Axioms and Uniqueness Theorem for Tsallis Entropy.” *Physics Letters A* 271(1-2): 74–79.
- Aczél, János. 1966. *Lectures on Functional Equations and Their Applications*. New York: Academic Press.
- Akerlof, George A. 1970. “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism.” *Quarterly Journal of Economics* 84(3): 488–500.
- Anderson, Simon P., André de Palma, and Jacques-François Thisse. 1992. *Discrete Choice Theory of Product Differentiation*. Cambridge, MA: MIT Press.
- Armington, Paul S. 1969. “A Theory of Demand for Products Distinguished by Place of Production.” *IMF Staff Papers* 16(1): 159–178.
- Arrow, Kenneth J. 1951. *Social Choice and Individual Values*. New York: Wiley.
- Arrow, Kenneth J., Hollis B. Chenery, Bagicha S. Minhas, and Robert M. Solow. 1961. “Capital-Labor Substitution and Economic Efficiency.” *Review of Economics and Statistics* 43(3): 225–250.
- Atkinson, Anthony B. 1970. “On the Measurement of Inequality.” *Journal of Economic Theory* 2(3): 244–263.
- Barth, James R., Gerard Caprio Jr., and Ross Levine. 2006. *Rethinking Bank Regulation: Till Angels Govern*. Cambridge: Cambridge University Press.
- Beck, Christian, and E.G.D. Cohen. 2003. “Superstatistics.” *Physica A* 322: 267–275.
- Becker, Gary S. 1973. “A Theory of Marriage: Part I.” *Journal of Political Economy* 81(4): 813–846.
- Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. “Does Management Matter? Evidence from India.” *Quarterly Journal of Economics* 128(1): 1–51.
- Bouchaud, Jean-Philippe, and Marc Mézard. 2000. “Wealth Condensation in a Simple Model of Economy.” *Physica A* 282(3–4): 536–545.
- Caplin, Andrew, and Mark Dean. 2015. “Revealed Preference, Rational Inattention, and Costly Information Acquisition.” *American Economic Review* 105(7): 2183–2203.
- Condorcet, Marquis de. 1785. *Essai sur l’application de l’analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris: Imprimerie Royale.

- Costinot, Arnaud. 2009. “An Elementary Theory of Comparative Advantage.” *Econometrica* 77(4): 1165–1192.
- Diamond, Peter A. 1982. “Aggregate Demand Management in Search Equilibrium.” *Journal of Political Economy* 90(5): 881–894.
- Dixit, Avinash K., and Joseph E. Stiglitz. 1977. “Monopolistic Competition and Optimum Product Diversity.” *American Economic Review* 67(3): 297–308.
- Drăgulescu, Adrian A., and Victor M. Yakovenko. 2000. “Statistical Mechanics of Money.” *European Physical Journal B* 17(4): 723–729.
- Eaton, Jonathan, and Samuel Kortum. 2002. “Technology, Geography, and Trade.” *Econometrica* 70(5): 1741–1779.
- Ethier, Wilfred J. 1982. “National and International Returns to Scale in the Modern Theory of International Trade.” *American Economic Review* 72(3): 389–405.
- Foley, Duncan K. 1994. “A Statistical Equilibrium Theory of Markets.” *Journal of Economic Theory* 62(2): 321–345.
- Friston, Karl. 2010. “The Free-Energy Principle: A Unified Brain Theory?” *Nature Reviews Neuroscience* 11(2): 127–138.
- Gabaix, Xavier. 2009. “Power Laws in Economics and Finance.” *Annual Review of Economics* 1(1): 255–294.
- Gabaix, Xavier. 2014. “A Sparsity-Based Model of Bounded Rationality.” *Quarterly Journal of Economics* 129(4): 1661–1710.
- Jones, Charles I. 1995. “R&D-Based Models of Economic Growth.” *Journal of Political Economy* 103(4): 759–784.
- Garicano, Luis. 2000. “Hierarchies and the Organization of Knowledge in Production.” *Journal of Political Economy* 108(5): 874–904.
- Grossman, Sanford J., and Oliver D. Hart. 1986. “The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration.” *Journal of Political Economy* 94(4): 691–719.
- Hardy, G. H., J. E. Littlewood, and G. Pólya. 1952. *Inequalities*. 2nd ed. Cambridge: Cambridge University Press.
- Hart, Oliver, and John Moore. 1990. “Property Rights and the Nature of the Firm.” *Journal of Political Economy* 98(6): 1119–1158.

- Jaynes, Edwin T. 1957. "Information Theory and Statistical Mechanics." *Physical Review* 106(4): 620–630.
- Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47(2): 263–291.
- Khinchin, Aleksandr I. 1957. *Mathematical Foundations of Information Theory*. New York: Dover.
- Kremer, Michael. 1993. "The O-Ring Theory of Economic Development." *Quarterly Journal of Economics* 108(3): 551–575.
- Krugman, Paul. 1991. "Increasing Returns and Economic Geography." *Journal of Political Economy* 99(3): 483–499.
- Matějka, Filip, and Alisdair McKay. 2015. "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model." *American Economic Review* 105(1): 272–298.
- Mortensen, Dale T. 1982. "The Matching Process as a Noncooperative Bargaining Game." In *The Economics of Information and Uncertainty*, ed. John J. McCall, 233–258.
- Myerson, Roger B. 1981. "Optimal Auction Design." *Mathematics of Operations Research* 6(1): 58–73.
- Nauenberg, Michael. 2003. "Critique of  $q$ -Entropy for Thermal Statistics." *Physical Review E* 67(3): 036114.
- Oberfield, Ezra, and Devesh Raval. 2021. "Micro Data and Macro Technology." *Econometrica* 89(2): 703–732.
- Petrongolo, Barbara, and Christopher A. Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature* 39(2): 390–431.
- Pissarides, Christopher A. 1985. "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages." *American Economic Review* 75(4): 676–690.
- Rabin, Matthew. 2000. "Risk Aversion and Expected-Utility Theory: A Calibration Theorem." *Econometrica* 68(5): 1281–1292.
- Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98(5): S71–S102.
- Santos, R.J.V. 1997. "Generalization of Shannon's Theorem for Tsallis Entropy." *Journal of Mathematical Physics* 38(8): 4104–4107.

- Sato, Kazuo. 1967. "A Two-Level Constant-Elasticity-of-Substitution Production Function." *Review of Economic Studies* 34(2): 201–218.
- Shimer, Robert. 2005. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *American Economic Review* 95(1): 25–49.
- Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50(3): 665–690.
- Smith, Eric, and Duncan K. Foley. 2008. "Classical Thermodynamics and Economic General Equilibrium Theory." *Journal of Economic Dynamics and Control* 32(1): 7–65.
- Smirl, Jon. 2026a. "Emergent CES and the Quadruple Role: Why Constant Elasticity of Substitution Is Not an Assumption." Working Paper.
- Smirl, Jon. 2026b. "Emergent CES: Renormalization, Functional Equations, and Maximum Entropy Derivations of the CES Aggregate." Working Paper.
- Solow, Robert M. 1956. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70(1): 65–94.
- Syyverson, Chad. 2004. "Market Structure and Productivity: A Concrete Example." *Journal of Political Economy* 112(6): 1181–1222.
- Suyari, Hiroki. 2004. "Generalization of Shannon-Khinchin Axioms to Nonextensive Systems and the Uniqueness Theorem for the Nonextensive Entropy." *IEEE Transactions on Information Theory* 50(8): 1783–1787.
- Thaler, Richard H., and Cass R. Sunstein. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven: Yale University Press.
- Theil, Henri. 1967. *Economics and Information Theory*. Amsterdam: North-Holland.
- Tsallis, Constantino. 1988. "Possible Generalization of Boltzmann-Gibbs Statistics." *Journal of Statistical Physics* 52(1-2): 479–487.
- Tversky, Amos, and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty* 5(4): 297–323.
- 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*. New York: Free Press.
- Yakovenko, Victor M., and J. Barkley Rosser Jr. 2009. "Colloquium: Statistical Mechanics of Money, Wealth, and Income." *Reviews of Modern Physics* 81(4): 1703–1725.