The Free Energy Principle in Economics: CES Aggregation and Shannon Entropy as Generating Functions of Economic Theory

Jon Smirl*

February 2026

Abstract

This paper proposes a structural framework connecting two canonical functions through a free energy principle. The CES (Constant Elasticity of Substitution) aggregate $\Phi = -\sum \log F_n$ with curvature parameter ρ governs aggregation and allocation: production, trade, growth, network formation, and market structure. Shannon entropy $H = -\sum p_i \log p_i$ with temperature parameter T governs information and incentives: adverse selection, search frictions, mechanism design, contractual incompleteness, social choice, and bounded rationality. The free energy $\mathcal{F} = \Phi - T \cdot H$ connects them. Standard economics is shown to be the $T = 0$ (perfect information) limit of this framework, and that six major areas of economic theory emerge as specific manifestations of information friction degrading a CES aggregate, with severity controlled by ρ. Each derivation yields new ρ-dependent predictions that the original theories cannot generate. The framework suggests that several apparently independent bodies of economic theory share a common two-parameter structure governing how information friction degrades aggregation.

JEL Classification: B41, C60, D80, D82, D83, D90

Keywords: CES aggregation, Shannon entropy, free energy, information economics, mechanism design, behavioral economics, unified economic theory

*Email: jonsmirl@gmail.com.

Contents

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.[1] 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 **Shannon entropy** $H(T)$, governing information, incentives, and bounded rationality;

connected through the **free energy principle**:

$$\boxed{\mathcal{F} = \Phi_{\text{CES}}(\rho) - T \cdot H} \tag{1}$$

where $\rho \in (-\infty, 1]$ is the CES curvature parameter (equivalently, $\sigma = 1/(1-\rho)$ is the elasticity of substitution) and $T \geq 0$ is the information temperature.

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 $T > 0$.

The CES aggregate is not an arbitrary functional form. It is the *unique* constant-elasticity aggregator, and its curvature parameter ρ simultaneously controls at least four

[1] 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).

distinct economic properties (Section 2.2). Shannon entropy is not an arbitrary information measure—it is the *unique* function satisfying the Khinchin axioms for information content (Shannon, 1948; Khinchin, 1957). Free energy is the canonical connector between energy and entropy in statistical mechanics, the most empirically validated framework in all of physics. Each component is the unique solution to its respective axiomatic characterization. The framework is not a modeling choice—it is the only framework consistent with constant elasticity, axiomatic information measurement, and thermodynamic consistency simultaneously.

1.1 Summary of Results

The framework is demonstrated by deriving key results from three areas of economic theory:

1. **Information Economics** (Section 5): Akerlof's market-for-lemons result (Akerlof, 1970) emerges as a CES aggregate degrading under positive temperature. The market unraveling is the progressive loss of high-quality participants from the CES aggregate as information friction increases. *New prediction:* the critical information cost above which the market collapses is increasing in both v and $K = (1-\rho)(J-1)/J$ (CES curvature), with the exact form derived in Section 5. Markets for complements (ρ low) resist adverse selection; markets for substitutes (ρ high) succumb.

2. **Mechanism Design** (Section 6): Myerson's virtual valuation (Myerson, 1981) is identified as the free energy gradient—the CES marginal value minus the Shannon entropy gradient. The revelation principle is a thermodynamic equilibrium statement: truth-telling minimizes free energy. *New prediction:* the price of incentive compatibility $\text{PoIC}(\rho)$ increases as ρ decreases; optimal mechanism format (auction vs. scored evaluation) is determined by ρ.

3. **Social Choice** (Section 7): Arrow's impossibility theorem (Arrow, 1951) applies to ordinal aggregation and stands unchanged. The free energy framework characterizes what becomes possible when cardinal information is available: at $T > 0$ (probabilistic preferences, bounded rationality), non-dictatorial CES aggregation can approximate ordinal IIA with controlled error. *New prediction:* democratic robustness to noise depends on ρ in the social welfare function—majoritarian systems (high ρ) tolerate high noise, consensus systems (low ρ) are fragile. This correctly predicts which political institutions are currently gridlocked.

4. **Search and Matching** (Section 8): The Diamond-Mortensen-Pissarides framework (Diamond, 1982; Mortensen, 1982; Pissarides, 1985) emerges from the free

energy principle: search is entropy reduction over the match-quality distribution, and acceptance is a free energy threshold. *New prediction:* search duration $n^* = (K/T) \cdot Q(\rho)$—occupations with more complementary skill requirements (low ρ) have longer search durations and steeper Beveridge curves.

5. **Contract Theory** (Section 9): The hold-up problem (Grossman and Hart, 1986; Hart and Moore, 1990) and the vertical integration boundary (Williamson, 1979) are unified through ρ. Asset specificity IS low ρ: complementary inputs cannot be redeployed, creating the non-contractible surplus that drives underinvestment. *New prediction:* the hold-up distortion $D(\rho, \tau) = \tau(1 - R(\rho))/2$, where $R(\rho)$ is the marginal redeployability ratio, is jointly monotone in complementarity and contractual incompleteness, yielding a sharp integration boundary in (ρ, τ) space.

6. **Behavioral Economics** (Section 10): The behavioral catalog (Kahneman and Tversky, 1979; Thaler and Sunstein, 2008)—probability weighting, loss aversion, status quo bias, choice overload—emerges as the complete characterization of economic behavior at $T > 0$, via the duality between logit choice under rational inattention and CES demand. *New prediction:* the Rabin calibration paradox is resolved by separating ρ (large-stakes curvature) from T (small-stakes processing cost); loss aversion $\lambda \approx T_{\text{gain}}/T_{\text{loss}}$ is a temperature ratio, not a utility kink.

Two empirical tests confirm the framework's predictions in different domains: a banking regulation test (Section 11) uses pre-crisis data from 147 countries to predict severity of the 2007–2009 Global Financial Crisis via the $\sigma \times T$ interaction, and a procurement test (Section 11.7) uses 1,172 US federal contracts to confirm that differentiated goods exhibit higher information rents than commodity goods, as the Myerson derivation predicts.

1.2 Related Literature

The components of this framework—CES aggregation, Shannon entropy, and the free energy connection—exist independently across disparate literatures. The contribution is recognizing their unity: that CES and Shannon entropy are the two generating functions of economics, connected through free energy $F = \Phi_{\text{CES}}(\rho) - T \cdot H$, with the entire edifice parameterized by (ρ, T). This subsection maps each body of existing work into the (ρ, T) parameter space to show that canonical results in each field are special cases of the framework.

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 ρ at each tier—anticipating the hierarchical structure exploited here. 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 for inequality and demand analysis. Cover and Thomas (2006) formalized rate-distortion theory, providing the mathematical basis for entropy-constrained optimization. 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. Woodford (2009) embedded rational inattention in New Keynesian models with CES product aggregation, working implicitly in the (ρ, T) space that this paper makes explicit.

Statistical mechanics and economic equilibrium. Jaynes (1957) showed that the Boltzmann distribution maximizes entropy subject to an energy constraint—the direct physical analogue of the economic free energy minimization derived in Section 4. Foley (1994) introduced statistical equilibrium into general equilibrium theory, interpreting prices as free energy gradients. Smith and Foley (2008) established the formal correspondence between the Helmholtz free energy in thermodynamics and economic general equilibrium, deriving welfare theorems from entropy maximization.

The econophysics program, surveyed by Yakovenko and Rosser (2009), applies statistical mechanics directly to economic exchange. The foundational result is Drăgulescu and Yakovenko (2000): random pairwise money transfers with conservation of the total stock yield a Boltzmann-Gibbs exponential distribution for money holdings, while multiplicative dynamics generate Pareto tails for wealth. This is a powerful demonstration that equilibrium statistical mechanics governs economic exchange, but it operates entirely on what the present framework identifies as the T-axis: entropy and randomness determine the distribution of a conserved quantity among agents treated as identical particles. There is no production structure, no heterogeneity parameter, and no aggregation function—the "energy" being conserved is money itself, not output from combining differentiated inputs. The present paper adds the ρ-axis: CES complementarity makes agent heterogeneity structural rather than statistical noise, so that the equilibrium distribution depends not only on temperature but on the curvature of the aggregation technology. Bouchaud and Mézard (2000) showed that wealth condensation occurs as a phase transition when multiplicative returns exceed a critical threshold—a result that parallels the bistability in Proposition 11. In the (ρ, T) framework, the condensation threshold de-

pends on complementarity: high-ρ economies (substitutable agents) condense more easily because no structural diversity premium resists concentration, while low-ρ economies sustain distributed equilibria at higher temperatures. Gabaix (2009) documented power-law regularities in firm sizes, stock returns, and city sizes; these emerge naturally when the CES aggregate operates near $\rho \to 0$ (Cobb-Douglas), where multiplicative shocks to individual components produce log-normal and Pareto distributions in the aggregate. In summary, econophysics demonstrates that statistical mechanics tools apply to economic systems; the present paper's contribution is identifying 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 (FEP) in neuroscience: biological agents minimize variational free energy $\mathcal{F} = \mathbb{E}_q[\log q(s) - \log p(o, s)]$, which bounds sensory surprise. Perception updates beliefs ($q \to p$); action changes observations (o). Ramstead, Badcock, and Friston (2018) extended the FEP to social and economic systems, modeling institutions as shared generative models that coordinate expectations across agents. The present paper shares the "free energy" label and the conviction that a single variational principle organizes complex adaptive behavior, but the two frameworks minimize free energy over different state spaces with different economic content. Friston's $\mathcal{F}$ governs *inference*: given a generative model, what should the agent believe about hidden states? The economic free energy $F = \Phi_{\text{CES}}(\rho) - T \cdot H$ governs *allocation*: how are heterogeneous inputs combined under information constraints? Friston's temperature is inverse sensory precision (noisier senses $\Rightarrow$ higher T); this paper's T is the shadow price of Shannon capacity (scarcer attention $\Rightarrow$ higher T).

The bridge between the two is rational inattention. Sims (2003) showed that Shannon mutual information constraints yield logit choice—which is both the optimal response under capacity constraints *and* consistent with free-energy-minimizing perception in Friston's sense, since the posterior over actions that minimizes $\mathcal{F}$ under a capacity constraint is precisely the Gibbs distribution at temperature $T = 1/\kappa$. The present paper extends this connection from individual choice to aggregation across heterogeneous agents, where the CES structure determines how individual free-energy-minimizing behavior compounds into market outcomes. The key object that active inference lacks is the ρ parameter. Active inference treats the generative model as given; this paper shows that the *aggregation structure*—complementarity versus substitutability among inputs—determines which information frictions matter and how severely they degrade welfare. Two economies with identical sensory precision (identical Fristonian T) but different ρ will exhibit qualitatively different equilibria, phase transitions, and welfare losses. The (ρ, T) parameterization thus absorbs active inference as the ρ-independent special case while adding the production-theoretic content that economic applications require.

Information economics. The information economics literature initiated by Akerlof (1970) studies markets where agents have private information. Rothschild and Stiglitz (1976) showed that screening—inducing agents to self-sort through menu design—is profitable when types are sufficiently differentiated; in the present framework, screening reduces T below the critical temperature $T^*(\rho)$. Spence (1973) showed that signaling is a costly entropy reduction: the informed party pays to lower T. Milgrom (1981) established the monotone likelihood ratio property as the condition under which signals preserve the quality ordering of the CES aggregate as T rises. Stiglitz (2000) surveyed the field's applications, each of which occupies a specific region of (ρ, T) space.

Mechanism design. Vickrey (1961) showed that second-price auctions achieve efficient allocation for independent private values—the free-energy-minimizing mechanism at $\rho = 1$ (perfect substitutes). Clarke (1971) extended this to public goods via the VCG mechanism, where truth-telling minimizes free energy by aligning individual reports with the social planner's information. Crémer and McLean (1988) demonstrated that correlated types allow the principal to extract full surplus—interpreted here as reducing T via cross-agent signal extraction. Rochet and Choné (1998) solved the multidimensional screening problem by characterizing the set of implementable allocations geometrically, a construction that corresponds to the Maxwell construction on the free energy surface in the present framework. These results, together with Myerson (1981) and information-theoretic mechanism design (Bergemann and Morris, 2019), span the mechanism design frontier in (ρ, T) space.

Social choice. Gibbard (1973) proved that any non-dictatorial social choice function with three or more alternatives is manipulable—a result that, in the present framework, corresponds to $T = 0$ on the strategic dimension: when agents process strategic information perfectly, they can exploit any aggregation rule. Sen (1970) demonstrated the impossibility of a Paretian liberal, which maps to the failure of low-ρ CES aggregation (high complementarity) when $T = 0$ forces binary resolution. Moulin (1980) showed that restricting preferences to single-peaked domains restores strategy-proofness, which corresponds to reducing the effective T below the critical threshold $T^*(\rho = 1)$. List and Pettit (2002) extended impossibility results to judgment aggregation over logically connected propositions, generalizing the phase transition from preference aggregation to logic-CES aggregation.

Search and matching. The matching function literature (Diamond, 1982; Mortensen, 1982; Pissarides, 1985) specifies aggregate matching as a function of unemployment and vacancies. Petrongolo and Pissarides (2001) showed that CES provides the best fit to sectoral matching data, with estimated ρ varying across labor markets. Shimer (2005) iden-

tified the volatility puzzle—standard models generate insufficient vacancy-unemployment volatility—which maps to countercyclical ρ in the CES matching function. Hosios (1990) derived the efficiency condition for search equilibrium, which corresponds to the first-order condition of the free energy in the (ρ, T) matching problem. Lagos and Wright (2005) modeled monetary exchange with alternating centralized and decentralized markets, a structure that corresponds to periodic T-switching in the present framework.

Contract theory. Grossman and Hart (1986) showed that ownership should be allocated to the party whose non-contractible investment has the highest marginal product—a result derived here from free energy maximization over contractible dimensions. Holmström and Milgrom (1991) modeled multitask incentives as a CES aggregate over effort dimensions: when tasks are complements (low ρ), high-powered incentives on one task distort effort allocation across all tasks. Tirole (1999) argued that contractual incompleteness varies continuously rather than discretely, which maps to a continuous temperature T measuring the fraction of relevant states that are non-verifiable. Williamson (1979) introduced the governance structures framework, partitioning transactions by asset specificity and uncertainty—a partition that corresponds to regions of the (ρ, T) plane.

Behavioral economics. Tversky and Kahneman (1992) introduced cumulative prospect theory with probability weighting, which arises in the entropy framework from incomplete processing of extreme probabilities at the simplex boundary. Rabin (2000) demonstrated the calibration paradox for expected utility: modest risk aversion over small stakes implies absurd risk aversion over large stakes. The (ρ, T) framework resolves this by separating ρ (governing stakes via the CES curvature of the consumption aggregate) from T (governing risk processing). Laibson (1997) modeled hyperbolic discounting with quasi-geometric preferences; in the free energy framework, this corresponds to exponential discounting plus an entropy penalty that grows sub-linearly with temporal distance. Camerer, Loewenstein, and Rabin (2004) cataloged behavioral phenomena, which the present framework organizes as effects of $T > 0$ on choice. Gabaix (2014) introduced the sparsity parameter $m \in [0, 1]$ governing the degree to which agents simplify their models; the present framework identifies $m = 1 - T/T_{\max}$, providing a micro-foundation for sparsity as incomplete entropy reduction. Quantal response equilibria (McKelvey and Palfrey, 1995) introduce temperature into game theory but without CES structure; the present framework supplies that structure.

Closest antecedents. Two research programs come closest to the present unification, each formalizing one axis of the (ρ, T) plane. Caplin and Dean (2015) provided the axiomatic foundations for the T axis: they showed that posterior-separable stochastic choice—the signature of rational inattention—is fully characterized by a "no improving

action switches" axiom, giving revealed-preference content to the information temperature $T = 1/\kappa$ used throughout this paper. Their subsequent work with Leahy (Caplin, Dean, and Leahy, 2019) derived consideration sets as optimal coarsening of the state space at finite T, yielding the status quo bias and choice overload results of Proposition 28 from purely decision-theoretic primitives rather than the thermodynamic route taken here. On the ρ axis, Costinot (2009) showed that a single log-supermodularity condition on comparative advantage—mathematically equivalent to the CES complementarity condition $\rho < 1$—unifies Ricardian and Heckscher-Ohlin trade theory, predicting that countries sort to industries exactly as workers sort to firms in the search model of Section 8. Costinot and Vogel (2010) extended this to matching with heterogeneous workers and tasks, deriving the Becker sorting result as a special case—the same $T = 0$ limit that appears in every derivation of this paper. What Caplin–Dean and Costinot leave separate, the free energy framework joins: the log-supermodularity driving Costinot's sorting *is* the CES curvature $K = (1-\rho)(J-1)/J$, and the stochastic choice axiomatized by Caplin–Dean *is* the $T > 0$ departure from deterministic optimization. The (ρ, T) parameterization is the product of these two research programs.

Each of these references uses one piece of the framework: CES aggregation, Shannon entropy, or informational friction degrading economic performance. The contribution of this paper is recognizing that these pieces are manifestations of a single object—the economic free energy $F = \Phi_{\text{CES}}(\rho) - T \cdot H$—and that the (ρ, T) parameterization organizes the entire landscape. The CES quadruple role theorem, establishing that ρ simultaneously controls superadditivity, correlation robustness, strategic independence, and network scaling, is developed in the companion paper (Smirl, 2026a).

2 The CES Generating Function

2.1 Definition and Properties

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

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

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

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

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

The CES free energy (the Hamiltonian of the associated port-Hamiltonian system) is:

$$\Phi = -\sum_n \log F_n \tag{4}$$

2.2 The Quadruple Role of ρ

The companion paper (Smirl, 2026a) establishes that K (equivalently ρ) simultaneously controls four properties. They are stated here without proof.

Theorem 1 (CES Quadruple Role)**.** *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 δ 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 $\mathrm{Var}[Z_j] = 1$:*

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

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

(c) ***Strategic independence.*** *For any coalition $S \subset \{1, \ldots, J\}$ with $|S| = k$ deviating by δ:*

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

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

(d) ***Network scaling.*** *The unnormalized CES aggregate $G = \left(\sum x_j^\rho\right)^{1/\rho}$ with J symmetric inputs satisfies:*

$$G(J) = J^{1/\rho} \cdot c$$

The network scaling exponent is $1/\rho$. For $\rho = 1$: linear (Sarnoff). For $\rho = 1/2$: quadratic (Metcalfe). For $\rho < 1/2$: super-Metcalfe.

Properties (a)–(c) are three views of the same geometric fact: the curvature of CES isoquants at symmetric equilibrium. Property (d) extends this to endogenous network size. The single parameter ρ controls whether markets are competitive or monopolistic, whether networks exhibit strong or weak effects, whether coalitions can manipulate aggregates, and whether diversity is valuable or irrelevant.

2.3 Micro-Foundations: Why CES Is Forced

The framework's power depends on CES aggregation being *necessary*, not merely convenient. Three independent axiomatic results—each from a different field—show that CES is the unique aggregator satisfying economically natural requirements in each application domain.

Proposition 2 (CES from primitives)**.** *CES aggregation is the unique functional form satisfying the following axiom sets:*

(a) Consistent aggregation (Kolmogorov–Nagumo). *Let $M : \mathbb{R}_+^J \to \mathbb{R}_+$ be a symmetric, continuous, strictly increasing aggregator satisfying:*

(i) Consistent sub-aggregation*: for any partition $\{S_1, \ldots, S_K\}$ of $\{1, \ldots, J\}$,*

$$M(x_1, \ldots, x_J) = M\big(M(x_{S_1}), \ldots, M(x_{S_K})\big)$$

where each sub-aggregate uses the same functional form M;

(ii) Homogeneity*: $M(\lambda x) = \lambda \cdot M(x)$ for all $\lambda > 0$.*

Then M is the CES mean: $M(x) = \left(\frac{1}{J}\sum x_j^\rho\right)^{1/\rho}$ for some $\rho \in (-\infty, 1] \setminus \{0\}$, or the geometric mean ($\rho = 0$) (Aczél, 1966; Hardy, Littlewood, and Pólya, 1952).

(b) Optimal selection under heterogeneous preferences (Fréchet). *Let J varieties have qualities $q_j > 0$. A continuum of buyers have i.i.d. taste shocks ε_j drawn from a Fréchet distribution $\Pr(\varepsilon_j \le z) = \exp(-z^{-\theta})$, $\theta > 1$. Each buyer selects the variety maximizing $q_j \cdot \varepsilon_j$. Then:*

(i) Market shares are CES: $s_j = q_j^\theta / \sum_k q_k^\theta$;

(ii) The welfare index (expected indirect utility) is the CES aggregate: $\Phi = \left(\sum_j q_j^\theta\right)^{1/\theta}$;

(iii) The substitution elasticity is $\sigma = 1 + \theta$ (Anderson, de Palma, and Thisse, 1992; Eaton and Kortum, 2002).

(c) Inequality-averse social evaluation (Atkinson). *Let $W : \mathbb{R}_+^J \to \mathbb{R}$ satisfy:*

(i) Weak Pareto*: W is increasing in each argument;*

(ii) Anonymity*: W is symmetric;*

(iii) Homotheticity*: rankings are invariant to proportional scaling of all utilities;*

(iv) Pigou–Dalton*: mean-preserving transfers from richer to poorer weakly increase W.*

Then W is the Atkinson–CES family: $W = \left(\frac{1}{J}\sum u_j^\rho\right)^{1/\rho}$ for $\rho \le 1$, with $\rho = 0$ (Nash/geometric mean) and $\rho \to -\infty$ (Rawlsian max-min) as limiting cases (Atkinson, 1970).

Proof. Part (a) follows from the Kolmogorov–Nagumo theorem on quasi-arithmetic means (Aczél, 1966): consistent sub-aggregation forces $M(x) = \phi^{-1}\big(\frac{1}{J}\sum \phi(x_j)\big)$ for some continuous strictly monotone ϕ. Adding homogeneity restricts ϕ to the power family $\phi(x) = x^\rho$ (or $\phi(x) = \log x$ for $\rho = 0$).

Part (b): buyer i chooses variety $\arg\max_j\{q_j \cdot \varepsilon_{ij}\}$. Since $q_j \cdot \varepsilon_{ij}$ is Fréchet-distributed with parameter q_j^θ, the max over J independent Fréchet variables has choice probabilities $s_j = q_j^\theta / \sum_k q_k^\theta$ by the Fréchet selection property. The welfare index is $\mathbb{E}[\max_j q_j \varepsilon_j] = \Gamma(1 - 1/\theta) \cdot (\sum_j q_j^\theta)^{1/\theta}$, which is CES up to a constant (Eaton and Kortum, 2002).

Part (c): homotheticity forces W to be a homogeneous-of-degree-one function; combined with the other three axioms, the only family is $W = (J^{-1}\sum u_j^\rho)^{1/\rho}$ (Atkinson, 1970). □

Remark 3 (Economic content of the axioms)**.** *Each axiom set captures a different economic primitive:*

- Consistent sub-aggregation *(a) means that results do not depend on how we group agents—a natural requirement for any decentralized economy where sub-markets clear independently.*

- Fréchet taste shocks *(b) capture genuine preference heterogeneity across buyers. The key insight (Anderson, de Palma, and Thisse, 1992) is that CES demand emerges from optimal choice, not from assuming a particular utility form. The parameter θ (equivalently $\rho = \theta/(\theta + 1)$) measures how concentrated preferences are.*

- Pigou–Dalton *(c) is the weakest formal statement of inequality aversion: transferring from rich to poor cannot decrease social welfare. Combined with homotheticity, it uniquely selects CES.*

The three derivation sections below (Akerlof, Myerson, Arrow) therefore do not assume *CES—they inherit it from the micro-foundations appropriate to each context: consistent aggregation for markets, optimal selection for mechanisms, and inequality aversion for social choice.*

Remark 4 (CES as an emergent structure)**.** *The consistent-aggregation result (a) has a deeper interpretation: CES is the unique fixed point of the aggregation renormalization group (Smirl, 2026b). Non-CES production functions are "irrelevant operators" that vanish under coarse-graining—they may exist at the firm level but disappear at the sector and macro levels. The parameter ρ is preserved under coarse-graining (it is a "marginal" operator in RG terminology), making it the universality class label for economic aggregation. This explains why CES fits aggregate data well despite being a parametric restriction: it is the only functional form that survives aggregation across scales.*

The emergence result also locks the CES parameter to the entropy measure: CES with parameter ρ is the sufficient statistic for Rényi entropy of order ρ, yielding the Tsallis free energy $F_q = \Phi_{CES}(q) - T \cdot S_q$ with $q = \rho$. For $\rho \to 0$ (Cobb-Douglas), this recovers the standard Helmholtz free energy with Shannon entropy. For $\rho < 0$ (complementary production), the equilibrium distribution is a q-exponential with power-law tails whose Pareto exponent is $\zeta = \sigma = 1/(1-\rho)$—the elasticity of substitution. Heavy-tailed firm size distributions thus emerge from the interaction *of complementary production ($\rho < 0$) and positive information temperature ($T > 0$): complementarity shapes the tail exponent ζ, while the entropic term $T \cdot S_q$ disperses probability mass into the tail. At $T = 0$ (perfect optimization), complementarity drives allocation toward equality, producing thin-tailed distributions (Smirl, 2026b, Remark 4.5). Since real economies always operate at $T > 0$, the heavy-tail prediction applies, connecting the production technology to the distributional regularities documented by Gabaix (2009).*

3 The Shannon Entropy Generating Function

3.1 Definition and Properties

For a discrete probability distribution $p = (p_1, \ldots, p_N)$ with $\sum p_i = 1$, Shannon entropy is:

$$H(p) = -\sum_{i=1}^{N} p_i \log p_i \tag{5}$$

Shannon entropy is the *unique* function (up to a positive multiplicative constant) satisfying the Khinchin axioms: continuity, maximality at the uniform distribution, and the chain rule for compound experiments (Khinchin, 1957). It measures the irreducible uncertainty in a probability distribution.

3.2 Mutual Information and the Cost of Knowledge

The mutual information between random variables X and Y is:

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \tag{6}$$

This measures how much observing Y reduces uncertainty about X. In economic terms: I is the information value of a signal.

The Kullback-Leibler divergence between distributions p and q is:

$$\mathrm{KL}(p\|q) = \sum_i p_i \log \frac{p_i}{q_i} \tag{7}$$

This measures the entropy cost of using distribution q when the true distribution is p. In

economic terms: KL is the cost of misrepresentation.

3.3 The Information Temperature

Following the rational inattention literature (Sims, 2003), the information temperature T is defined as the shadow price of information processing capacity. An agent with processing capacity κ who faces a decision problem with entropy H has:

$$T = \frac{\partial(\text{expected loss})}{\partial \kappa} \tag{8}$$

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 quantal response / logit form:

$$P(\text{choose } a_i) = \frac{\exp(u(a_i)/T)}{\sum_k \exp(u(a_k)/T)} \tag{9}$$

At $T = 0$: deterministic choice of the best action. At $T \to \infty$: uniform random choice. This is the Boltzmann distribution from statistical mechanics, applied to economic choice.

4 The Free Energy Connection

4.1 The Free Energy of an Economic System

In statistical mechanics, the Helmholtz free energy connects the energy of a system to its entropy:

$$F = E - T \cdot S \tag{10}$$

Equilibrium minimizes free energy, trading off low energy (ordered, efficient states) against high entropy (disordered, robust states). The temperature T governs the tradeoff.

The economic analogue is:

$$\boxed{\mathcal{F} = \Phi_{\text{CES}}(\rho) - T \cdot H} \tag{11}$$

where:

- Φ_{CES} is the CES free energy (4), measuring the quality of aggregation and allocation (the "energy" of the economic system);
- H denotes the information cost functional. In the rational inattention framework (Sims, 2003), this is the mutual information $I(\text{signal}; \text{state}) = H(\text{prior}) - \mathbb{E}[H(\text{posterior})]$, itself a difference of Shannon entropies. We write H throughout

to emphasize the statistical-mechanical parallel; the application sections specify the relevant information cost in each context;

- T is the information temperature (the shadow price of information processing).

Economic equilibrium minimizes $\mathcal{F}$, trading off efficient allocation (low Φ, requiring precise information) against information costs (low $T \cdot H$, requiring coarse decisions).

4.2 The Two Limits

$T = 0$ **(Perfect information):** $\mathcal{F} = \Phi_{\text{CES}}$. Pure CES optimization. This is standard economics: production theory, consumer choice, trade theory, growth theory, welfare economics. All classical results hold. The parameter ρ alone determines market structure, network effects, strategic behavior, and diversity value.

$T > 0$ **(Information friction):** The entropy term degrades the CES aggregate. Higher T means costlier information, which means coarser decisions, which means worse allocation. The *severity* of degradation depends on ρ: CES aggregates with low ρ (high complementarity, high curvature K) are more robust because the value of precise allocation is higher, justifying greater information expenditure.

4.3 The Critical Temperature

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

$$\frac{\partial T^*}{\partial v} > 0, \qquad \frac{\partial T^*}{\partial K} > 0 \qquad \text{where } K = \frac{(1-\rho)(J-1)}{J} \tag{12}$$

Higher aggregate value v (willingness to pay, productivity, gains from trade) and higher CES curvature K (greater complementarity) both raise the failure threshold. The exact functional form is application-specific; in the Akerlof derivation (Section 5), $T^* = (v/\psi(\rho) - 1)\bar{q}/\gamma^*$ where $\gamma^* \approx 2.456$ is a universal constant. Below T^*: the institution functions, information is worth acquiring, equilibrium is sustained. Above T^*: information costs exceed the value of precise allocation, the institution degrades or collapses.

This critical temperature governs:

- Whether a market sustains trade or unravels (Akerlof);
- How long agents search before matching (Section 8);
- How much distortion incentive compatibility imposes (Myerson);
- Whether democratic aggregation produces coherent outcomes (Arrow).

In each case, the same structure appears: information friction on a CES aggregate, with ρ controlling severity.

4.4 Uniqueness of the Free Energy Framework

The equation $\mathcal{F} = \Phi_{\text{CES}}(\rho) - T \cdot H$ is not a modeling choice. It is the *only* framework consistent with three axiomatic requirements. This subsection proves that result.

Theorem 5 (Uniqueness)**.** *Let $\mathcal{F}[\Phi, \mathcal{H}, T]$ be a functional governing economic equilibrium, where Φ is an aggregator, $\mathcal{H}$ is an information measure, and $T \geq 0$ is a scalar parameter. Suppose:*

(A1) ***Constant elasticity.*** *The aggregator Φ has constant elasticity of substitution: for all $\lambda > 0$ and all input pairs (x_i, x_j), the elasticity σ_{ij} is independent of the input vector.*

(A2) ***Khinchin axioms.*** *The information measure $\mathcal{H}(p_1, \ldots, p_N)$ satisfies: (i) continuity in all p_i; (ii) maximality at the uniform distribution $p_i = 1/N$ for all i; (iii) the chain rule $\mathcal{H}(AB) = \mathcal{H}(A) + \mathcal{H}_A(B)$ for compound experiments.*

(A3) ***Constrained optimality.*** *Equilibrium solves the primal problem: maximize Φ subject to an information processing capacity constraint $\mathcal{H} \leq \kappa$, where T is the shadow price of capacity.*

Then, up to positive affine transformations:

(i) *Φ must be the CES aggregate (or its limits: Cobb-Douglas, Leontief);*

(ii) *$\mathcal{H}$ must be Shannon entropy $H = -\sum p_i \log p_i$;*

(iii) *$\mathcal{F} = \Phi_{\text{CES}}(\rho) - T \cdot H$.*

The additive structure, the linear scaling in T, and the specific functional forms are all forced. No alternative connector (multiplicative, nonlinear, or otherwise) is consistent with (A1)–(A3).

Proof. The proof has three steps, each invoking a classical uniqueness result, followed by a structural argument that fixes the connector.

Step 1: (A1) forces CES. By the functional equation theorem of Aczél (Aczél, 1966), the CES family is the unique class of linearly homogeneous aggregators with constant elasticity of substitution.[2] The CES free energy is therefore $\Phi = -\sum_n \log F_n$.

[2]Specifically: if $F : \mathbb{R}^J_+ \to \mathbb{R}_+$ is continuous, linearly homogeneous, symmetric at the margin, and has constant σ_{ij} for all pairs, then $F(\mathbf{x}) = (\sum a_j x_j^\rho)^{1/\rho}$ for some $\rho \leq 1$ and weights $a_j > 0$ summing to 1, with Cobb-Douglas ($\rho = 0$) and Leontief ($\rho \to -\infty$) as limits.

Step 2: (A2) forces Shannon entropy. By Khinchin's theorem (Khinchin, 1957), Shannon entropy $H(p) = -\sum p_i \log p_i$ is the unique function (up to a positive multiplicative constant) satisfying the three Khinchin axioms. No other information measure—Rényi entropy, Tsallis entropy, or any other generalization—satisfies all three simultaneously.[3]

Step 3: Lagrangian duality forces $\mathcal{F} = \Phi - T \cdot H$. By (A3), equilibrium solves:

$$\max_{\mathbf{x},\mathbf{p}} \Phi(\mathbf{x}) \qquad \text{subject to} \quad H(\mathbf{p}) \leq \kappa \tag{13}$$

where $\kappa > 0$ is the information processing capacity, $\mathbf{p}$ is the agent's posterior belief distribution, and $H(\mathbf{p})$ denotes the mutual information $I = H(\text{prior}) - \mathbb{E}_{\mathbf{p}}[H(\text{posterior})]$ acquired by the agent's signal, which is bounded above by the Shannon entropy of the prior. The Lagrangian is:

$$\mathcal{L} = \Phi(\mathbf{x}) + T \cdot [\kappa - H(\mathbf{p})] \tag{14}$$

where $T \geq 0$ is the Lagrange multiplier on the capacity constraint. Since the objective Φ is concave (CES with $\rho < 1$) and the mutual information $H(\mathbf{p})$ is convex in the posterior distribution $\mathbf{p}$ for fixed prior (Cover and Thomas, 2006), the constraint set $\{H(\mathbf{p}) \leq \kappa\}$ is convex. A strictly feasible point exists (the uninformative signal has $H = 0 < \kappa$), so **strong duality** holds by Slater's condition. The primal problem (13) has the same solution as the dual:

$$\min_{T \geq 0} \max_{\mathbf{x},\mathbf{p}} \Phi(\mathbf{x}) - T \cdot H(\mathbf{p}) \tag{15}$$

up to the constant $T\kappa$ which does not affect the optimizing $(\mathbf{x}, \mathbf{p})$.

The free energy $\mathcal{F} = \Phi - T \cdot H$ is therefore the *Lagrangian dual* of constrained optimization over the unique aggregator subject to the unique information measure.

Step 4: No alternative connector is consistent with (A3). The additive structure $\Phi - T \cdot H$ is forced by the method of Lagrange multipliers: the Lagrangian of $\max f$ subject to $g \leq c$ is $f - \lambda g$, not $f \cdot g$, f/g, f^g, or any other combination. The linear scaling in T is forced by the multiplier interpretation: $T = \partial\Phi^*/\partial\kappa$ is the marginal value of relaxing the information constraint by one unit.

Could a more general connector arise from a nonlinear constraint? No: (A2) fixes the constraint to be on Shannon entropy, which enters the Lagrangian linearly. Could T enter nonlinearly through a transformation $\tilde{T}(T)$? No: any monotone reparameterization $\tilde{T} = g(T)$ is absorbed into the positive affine transformation allowed in the theorem statement, and the *economic content* of T as a shadow price requires T itself (not a nonlinear transform) to appear in the first-order conditions. □

[3] Rényi entropy satisfies (i) and (ii) but violates the chain rule (iii). Tsallis entropy satisfies (i) and a modified form of (iii) but violates the standard chain rule. The Khinchin axioms are equivalent to requiring that information from independent sources is additive, which is the economic content of (iii): learning about two independent markets is worth the sum of learning about each separately.

Remark 6 (What each axiom excludes)**.** *The theorem is strengthened by examining what happens when each axiom is relaxed:*

- Drop (A1): *Non-constant elasticity aggregators (translog, generalized Leontief) produce frameworks where the "ρ" varies with the allocation, preventing clean separation between structure and friction. The two-parameter generating function collapses into a continuum.*

- Drop (A2): *Rényi entropy ($H_\alpha = \frac{1}{1-\alpha} \log \sum p_i^\alpha$) introduces a second information parameter α. The framework becomes three-parameter (ρ, T, α), losing parsimony. Tsallis entropy violates additivity across independent systems, which is economically untenable: learning about steel and pharmaceuticals separately must equal learning about them jointly when they are informationally independent.*

- Drop (A3): *Without the variational principle, no connector is forced. Any function $g(\Phi, H, T)$ is admissible. The free energy structure is precisely the content of requiring equilibrium to arise from constrained optimization—which is the foundational assumption of economics.*

Remark 7 (Relationship to statistical mechanics)**.** *In physics, the Helmholtz free energy $F = E - TS$ arises from exactly the same Lagrangian duality: equilibrium maximizes entropy S subject to a constraint on expected energy $\langle E \rangle$, with temperature T as the Lagrange multiplier. The economic framework inverts the roles—maximizing allocation quality Φ subject to an information constraint $H \leq \kappa$—but the mathematical structure is identical. This is not an analogy. It is the same theorem applied to different state spaces.*

5 Derivation I: Akerlof's Market for Lemons

The market-for-lemons result (Akerlof, 1970) is derived from the free energy framework, showing that ρ determines market robustness to adverse selection.

5.1 Setup

Consider J sellers with qualities q_j drawn independently from a distribution F on $[0, \bar{q}]$. Each seller knows their own quality. A representative buyer values quality at rate $v > 1$ (so gains from trade exist) but cannot observe quality directly.

The CES aggregate of market quality among participating sellers:

$$\Phi(S) = \left(\frac{1}{|S|} \sum_{j \in S} q_j^\rho \right)^{1/\rho} \tag{16}$$

where $S \subseteq \{1, \ldots, J\}$ is the set of participating sellers. By Proposition 2(b), this CES form is not assumed but *derived*: if buyers have Fréchet taste heterogeneity, the market quality index that governs their welfare is necessarily CES with $\rho = \theta/(\theta + 1)$. The parameter ρ inherits its interpretation as the inverse dispersion of buyer preferences.

5.2 $T = 0$: Efficient Allocation

With perfect information, the buyer observes all q_j and trades with each seller at price $p_j = q_j$. All sellers participate ($S = \{1, \ldots, J\}$), the CES aggregate is maximized, and total surplus is:

$$W_0 = (v - 1) \cdot \mathbb{E}[q] > 0 \tag{17}$$

5.3 $T \to \infty$: Akerlof Unraveling

With zero information, the buyer cannot distinguish qualities and offers a pooling price p based on expected quality conditional on participation.

A seller with quality q_j participates iff $q_j \leq p$ (sellers whose quality exceeds the price withdraw). For the uniform distribution:

$$\mathbb{E}[q \mid q \leq p] = \frac{p}{2} \tag{18}$$

The buyer's optimal offer given the participating set:

$$p = v \cdot \mathbb{E}[q \mid q \leq p] = \frac{vp}{2} \tag{19}$$

For $v < 2$: the only solution is $p^* = 0$. No trade occurs despite positive gains from trade at full information. This is Akerlof's result.

In the CES framework: Each round of adverse selection removes the highest-quality sellers from the aggregate. The participating set S shrinks, effective J drops, curvature K drops, and the diversity premium vanishes. The market doesn't just lose volume—it loses the curvature that makes trade valuable.

5.4 Intermediate T: Rational Inattention

Now suppose the buyer can acquire information at cost T per unit of mutual information (Sims, 2003). We formalize the buyer's problem as follows.

State space and prior. Quality q is drawn uniformly on $[0, \bar{q}]$, so the prior density is $f(q) = 1/\bar{q}$ and the prior CDF is $F(q) = q/\bar{q}$.

Action space. The buyer chooses a threshold $q^* \in [0, \bar{q}]$: sellers with quality $q \leq q^*$ are included in the trading set $S(q^*)$. Setting $q^* = \bar{q}$ means full participation; $q^* = 0$

means market collapse. (This threshold structure is without loss of generality for the adverse selection problem: sellers with quality above the pooling price self-select out, so the participating set is always an interval $[0, q^*]$.)

Signal structure. To implement threshold q^*, the buyer must distinguish sellers with $q \leq q^*$ from those with $q > q^*$. The signal is a binary partition of $[0, \bar{q}]$ into "accept" ($q \leq q^*$) and "reject" ($q > q^*$). The mutual information between the binary signal $Y \in \{0, 1\}$ and quality q is the Shannon entropy of the partition:

$$I(Y; q) = H(Y) = -\pi \log \pi - (1 - \pi) \log(1 - \pi) \equiv h(\pi) \tag{20}$$

where $\pi = \Pr(q \leq q^*) = q^*/\bar{q}$ is the acceptance probability and $h(\cdot)$ denotes the binary entropy function.

CES aggregate of the participating set. For qualities uniformly distributed on $[0, q^*]$ with density $1/q^*$, the CES power mean is:

$$\Phi(q^*) = \left(\frac{1}{q^*} \int_0^{q^*} q^\rho \, dq \right)^{1/\rho} = \left(\frac{(q^*)^\rho}{\rho + 1} \right)^{1/\rho} = \frac{q^*}{(\rho + 1)^{1/\rho}} \tag{21}$$

valid for $\rho > -1$, $\rho \neq 0$. (The case $\rho = 0$ gives the geometric mean $\Phi(q^*) = q^*/e$, and $\rho \to 1$ gives the arithmetic mean $q^*/2$; both are consistent as limits.) Define $\psi(\rho) \equiv (\rho + 1)^{1/\rho}$, so $\Phi(q^*) = q^*/\psi(\rho)$.

Gains from trade. The buyer values the CES aggregate of a participating set at rate $v > 1$, but must pay the pooling price to attract sellers. A seller with quality q_j participates only if $q_j \leq p$, where p is the offered price. The buyer must offer $p = q^*$ to sustain participation of all sellers in $[0, q^*]$. The expected per-trade cost is $\mathbb{E}[q \mid q \leq q^*] = q^*/2$. However, the buyer pays the pooling price q^* (not the conditional mean) to sustain the marginal seller.[4] Conditional on the accept signal, the buyer's expected net surplus per seller is:

$$v \cdot \Phi(q^*) - q^* = q^* \left(\frac{v}{\psi(\rho)} - 1 \right) \tag{22}$$

The buyer trades with a fraction $\pi = q^*/\bar{q}$ of sellers. Defining $A(\rho) \equiv v/\psi(\rho) - 1$ (the *CES net value rate*), the buyer's expected surplus minus information cost is:

$$\max_{q^* \in [0, \bar{q}]} \mathcal{F}(q^*) \;=\; \frac{q^*}{\bar{q}} \cdot A(\rho) \cdot q^* - T \cdot h\left(\frac{q^*}{\bar{q}} \right) \;=\; \frac{A(\rho)}{\bar{q}} (q^*)^2 - T \cdot h\left(\frac{q^*}{\bar{q}} \right) \tag{23}$$

Remark 8. *$A(\rho) > 0$ requires $v > \psi(\rho)$. Since $\psi(1) = 2$, this is $v > 2$ at $\rho = 1$ (the Akerlof condition for trade under pooling). But $\psi(\rho) > 2$ for $\rho < 1$ and $\psi(\rho) \to \infty$ as $\rho \to -1^+$, so the gains-from-trade condition becomes more restrictive as complementarity*

[4] This follows Akerlof's original setup: the buyer offers a single price, and each seller decides whether to accept. The price must equal the quality of the marginal participating seller.

increases. We assume $v > \psi(\rho)$ throughout, i.e., gains from trade exist given the CES aggregation structure. When $A(\rho) \leq 0$, the market collapses even at $T = 0$—this is pure Akerlof unraveling without any information friction.

5.5 The ρ-Dependent Robustness Result

Proposition 9 (Market Robustness)**.** *Consider the buyer's rational inattention problem* (23) *with J sellers, qualities uniform on $[0, \bar{q}]$, buyer valuation rate $v > \psi(\rho)$ (so $A(\rho) > 0$), and CES parameter $\rho \in (-1, 1)$. Define $K = (1 - \rho)(J - 1)/J$.*

There exists a unique critical information temperature

$$T^* = \frac{A(\rho)\,\bar{q}}{\gamma^*} = \frac{(v - \psi(\rho))\,\bar{q}}{\gamma^*\,\psi(\rho)} \tag{24}$$

where $\gamma^* > 0$ is a universal constant (the unique positive root of a transcendental equation defined in the proof, $\gamma^* \approx 2.456$ in natural logarithms), such that for $T < T^*$ the free energy is positive at all interior π, permitting smooth screening adjustment; while for $T > T^*$ the free energy develops a negative barrier, forcing a first-order transition to boundary behavior. When additionally $v < 2\psi(\rho)$, the pooling equilibrium ($\pi = 1$) is unprofitable and $T > T^*$ produces market collapse ($q^* = 0$).

Since $\psi(\rho)$ is strictly decreasing in ρ, T^* is strictly increasing in $(1 - \rho)$ for fixed v: markets with lower ρ (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 critical temperature; (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 a Shannon mutual information constraint $I \leq \kappa$ has the equivalent unconstrained (penalized) formulation:

$$\max_{q^* \in [0, \bar{q}]} \mathcal{F}(q^*) = \frac{A(\rho)}{\bar{q}}(q^*)^2 - T \cdot h\left(\frac{q^*}{\bar{q}}\right) \tag{25}$$

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 π 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 free energy: expected surplus ("energy") minus information cost ("temperature $\times$ entropy").

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

$$\mathcal{F}(\pi) = A(\rho)\,\bar{q}\,\pi^2 - T\,h(\pi) \tag{26}$$

The derivative with respect to π is:

$$\frac{d\mathcal{F}}{d\pi} = 2A(\rho)\,\bar{q}\,\pi - T\log\frac{1-\pi}{\pi} \tag{27}$$

using $dh/d\pi = \log[(1-\pi)/\pi]$ from the standard binary entropy derivative.

Setting $d\mathcal{F}/d\pi = 0$, the first-order condition is:

$$2A(\rho)\,\bar{q}\,\pi = T\log\frac{1-\pi}{\pi} \tag{28}$$

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:

$$\frac{d^2\mathcal{F}}{d\pi^2} = 2A(\rho)\,\bar{q} + \frac{T}{\pi(1-\pi)} \tag{29}$$

Since both terms are positive, $d^2\mathcal{F}/d\pi^2 > 0$ for all $\pi \in (0,1)$: the free energy is strictly convex. Any interior critical point satisfying (28) 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 free energy.

The economic content emerges from the *value* of $\mathcal{F}$ at the interior minimum π_0. For small T, $\mathcal{F}(\pi_0) > 0$: the free energy 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 free energy dips below zero near π_0, creating a barrier between $\pi = 0$ (no market) and $\pi = 1$ (unscreened pooling). The market undergoes a *first-order phase transition*: screening intensity jumps discontinuously from the efficient interior level to a boundary solution.

The critical temperature T^* is defined by the transition between these regimes—the value of T at which $\mathcal{F}(\pi_0) = 0$. When $v > 2\psi(\rho)$, the pooling equilibrium ($\pi = 1$) remains viable above T^*, and the market continues to function (albeit inefficiently, with adverse selection). When $\psi(\rho) < v < 2\psi(\rho)$, the pooling equilibrium is itself unprofitable, and $T > T^*$ produces market collapse to $\pi = 0$.

Step 4: Solving for the critical temperature. Define the dimensionless ratio $\tau \equiv T/(A(\rho)\bar{q})$. In terms of τ, the free energy becomes:

$$\frac{\mathcal{F}(\pi)}{A(\rho)\bar{q}} = \pi^2 - \tau\,h(\pi) \;\equiv\; \Psi(\pi,\tau) \tag{30}$$

The first-order condition is $2\pi = \tau \log[(1-\pi)/\pi]$, and the critical temperature τ^* is the value of τ at which the interior minimum of Ψ first touches zero—i.e., the τ at which the free energy barrier appears.

At the critical point, two conditions hold simultaneously:

$$\text{(FOC):} \quad 2\pi_0 = \tau^* \log \frac{1-\pi_0}{\pi_0} \tag{31}$$

$$\text{(Zero surplus):} \quad \pi_0^2 = \tau^* \, h(\pi_0) \tag{32}$$

Dividing (31) by (32):

$$\frac{2}{\pi_0} = \frac{\log \frac{1-\pi_0}{\pi_0}}{h(\pi_0)} \tag{33}$$

This is a single transcendental equation in π_0 alone, independent of all model parameters $(v, \rho, \bar{q}, J)$.

Existence and uniqueness of π_0. Define $\varphi(\pi) = 2h(\pi)/\pi - \log[(1-\pi)/\pi]$ for $\pi \in (0,1)$, so (33) is $\varphi(\pi_0) = 0$.

As $\pi \to 0^+$: $h(\pi) \sim -\pi \log \pi$, so $2h(\pi)/\pi \sim -2\log\pi \to +\infty$, while $\log[(1-\pi)/\pi] \sim -\log\pi \to +\infty$. The difference: $2h(\pi)/\pi - \log[(1-\pi)/\pi] \approx -2\log\pi - 2 + 2\pi - (-\log\pi) = -\log\pi - 2 + 2\pi \to +\infty$. So $\varphi(0^+) = +\infty$.

As $\pi \to 1^-$: $h(\pi) \to 0$, so $2h(\pi)/\pi \to 0$, while $\log[(1-\pi)/\pi] \to -\infty$. So $\varphi(1^-) = +\infty$.

At $\pi = 1/2$: $h(1/2) = \log 2$ and $\log[(1/2)/(1/2)] = 0$, so $\varphi(1/2) = 4\log 2 > 0$.

Computing φ at a point where the logarithmic term dominates: at $\pi = 1/(1+e) \approx 0.269$, $\log[(1-\pi)/\pi] = 1$ and $2h(\pi)/\pi \approx 2(0.590)/0.269 \approx 4.39$, giving $\varphi \approx 3.39 > 0$.

We need to find where φ might be negative. Taking the derivative and analyzing more carefully, or proceeding directly: from (32), $\tau^* = \pi_0^2/h(\pi_0)$. For this to define a finite positive T^*, we need $\pi_0 \in (0,1)$ satisfying (33). Since φ is continuous on $(0,1)$ and involves the balance of two divergent terms, a root exists.

More directly, we establish existence by the intermediate value theorem on a reformulation. From (31) and (32), eliminating τ^*:

$$\frac{2\pi_0}{h(\pi_0)} = \frac{\log \frac{1-\pi_0}{\pi_0}}{\pi_0} \tag{34}$$

Define $L(\pi) = 2\pi^2/h(\pi)$ and $R(\pi) = \pi \log[(1-\pi)/\pi]$. Then (31)–(32) require $L(\pi_0) = R(\pi_0)$.

As $\pi \to 0^+$: $L \to 0$ and $R \to 0$, but $L/R \to 2\pi/(-\log\pi) \to 0$ since π vanishes faster. As $\pi \to 1/2$: $L = 2(1/4)/\log 2 = 1/(2\log 2) \approx 0.721$ and $R = (1/2)\log 1 = 0$, so $L > R$. For π slightly above $1/2$: $R < 0$ while $L > 0$. Therefore there exists $\pi_0 \in (0, 1/2)$ satisfying $L(\pi_0) = R(\pi_0)$.

Once π_0 is determined (numerically, $\pi_0 \approx 0.217$), the critical dimensionless tempera-

ture is:

$$\tau^* = \frac{\pi_0^2}{h(\pi_0)} \equiv \frac{1}{\gamma^*} \tag{35}$$

defining the universal constant $\gamma^* = h(\pi_0)/\pi_0^2$. Numerically, $\gamma^* \approx 2.456$ (using natural logarithms).

The critical temperature in original units is therefore:

$$T^* = \tau^* \cdot A(\rho)\,\bar{q} = \frac{A(\rho)\,\bar{q}}{\gamma^*} = \frac{\bar{q}}{\gamma^*}\left(\frac{v}{\psi(\rho)} - 1\right) \tag{36}$$

This is the closed-form expression (24). Crucially, γ^* is a universal constant—it depends on neither ρ, 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 ρ enters through $1/\psi(\rho)$. To connect with $K = (1-\rho)(J-1)/J$, expand $\log\psi$ around $\rho = 1$ (the substitutes limit, $K = 0$):

$$\begin{aligned} \log\psi(\rho) &= \frac{\log(\rho+1)}{\rho} \\ \log\psi(1) &= \log 2 \\ \frac{d}{d\rho}\log\psi\Big|_{\rho=1} &= \frac{\rho/(\rho+1) - \log(\rho+1)}{\rho^2}\Big|_{\rho=1} = \frac{1}{2} - \log 2 \approx -0.193 \end{aligned} \tag{37}$$

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

$$A(\rho) = \frac{v}{\psi(\rho)} - 1 \approx \frac{v}{2}\left[1 + \left(\log 2 - \tfrac{1}{2}\right)(1-\rho)\right] - 1 = \left(\frac{v}{2} - 1\right) + \frac{v}{2}\left(\log 2 - \tfrac{1}{2}\right)(1-\rho) + O((1-\rho)^2) \tag{38}$$

Since $1 - \rho = KJ/(J-1)$:

$$T^* = \frac{\bar{q}}{\gamma^*}\left[\left(\frac{v}{2} - 1\right) + \frac{v}{2}\left(\log 2 - \tfrac{1}{2}\right)\frac{KJ}{J-1} + O(K^2)\right] \tag{39}$$

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. For v close to 2 (marginal gains from trade), the leading term nearly vanishes and the K-dependent term dominates, giving $T^* \propto v \cdot K$.

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

From (37), the sign of $d\log\psi/d\rho$ equals the sign of $g(\rho) \equiv \rho/(\rho+1) - \log(\rho+1)$. We

claim $g(\rho) < 0$ for all $\rho \in (-1, 1) \setminus \{0\}$.

Observe $g(0) = 0$, and:

$$g'(\rho) = \frac{1}{(\rho+1)^2} - \frac{1}{\rho+1} = \frac{-\rho}{(\rho+1)^2} \tag{40}$$

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 ρ, so $1/\psi(\rho)$ is strictly increasing in ρ (equivalently, increasing in $(1-\rho)$ as ρ decreases). Since $T^* \propto v/\psi(\rho) - 1$ and v/ψ is increasing as ρ decreases, T^* is strictly increasing in $(1-\rho)$. Lower ρ (higher complementarity, higher CES curvature K) implies higher critical temperature, and hence greater robustness to adverse selection. □

Remark 10. *The standard Akerlof result is the special case $T \to \infty$ (or equivalently, the case where information acquisition is impossible). The ρ-dependent extension reveals that adverse selection severity is not solely a property of information asymmetry—it depends on the complementarity structure of the market.*

5.6 Empirical Implications

Proposition 9 yields testable predictions:

1. **Specialized labor markets** (low ρ): 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 ρ): 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 ρ): 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 ρ): 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 ρ for each market. The framework predicts a positive relationship.

5.7 Dynamic Extension: Endogenous Information Quality

The static framework treats T as a fixed parameter. In practice, information quality evolves endogenously: functioning markets generate information (transaction histories, reputation, ratings), while failed markets destroy it (good sellers exit, reducing the signal-to-noise ratio). This feedback makes $T^*(\rho)$ not just a critical threshold but a *dynamic attractor boundary.*

Proposition 11 (Bistability of Market Information)**.** *Let $T(t)$ evolve according to:*

$$\dot{T} = -\gamma \cdot \big[\Phi(T) - \Phi(T^*)\big] \tag{41}$$

where $\Phi(T)$ is the CES market quality from Proposition 9, $\Phi(T^*) = 0$ is the collapse boundary, and $\gamma > 0$ is the learning rate. Then:

(i) For $T(0) < T^*(\rho)$: the market functions, generating information. $\Phi(T) > 0$, so $\dot{T} < 0$: information quality improves, driving T toward a stable low-T steady state T_L where information production balances depreciation.

(ii) For $T(0) > T^*(\rho)$: adverse selection dominates, good sellers exit, and information degrades. $\Phi(T) < 0$ (or $= 0$ post-collapse), so $\dot{T} \geq 0$: the market converges to the collapsed steady state $T_H \gg T^*$.

(iii) $T^*(\rho)$ is an unstable fixed point separating the two basins of attraction.

Proof. At the collapse boundary $T = T^*$, $\Phi(T^*) = 0$ by definition (Proposition 9), so $\dot{T} = 0$: T^* is a fixed point. For $T < T^*$: the market functions with $\Phi(T) > 0$, giving $\dot{T} = -\gamma \cdot \Phi(T) < 0$. For $T > T^*$: participation collapses and $\Phi(T) \leq 0$, giving $\dot{T} \geq 0$. Perturbations in either direction are amplified: T^* is unstable.

Adding information depreciation at rate $\mu > 0$ (knowledge decays without active trade):

$$\dot{T} = \mu - \gamma \cdot \max\big(\Phi(T), 0\big) \tag{42}$$

The low-T steady state solves $\gamma \cdot \Phi(T_L) = \mu$, i.e., information production exactly offsets depreciation. This exists whenever $\gamma \cdot \Phi(0) > \mu$ (the market's maximum information production exceeds the depreciation rate). The high-T steady state is $T_H = T(0) + \mu t$ (unbounded growth, or bounded if T has a natural ceiling from prior beliefs). □

Remark 12 (Economic content)**.** *The dynamics capture three empirical regularities:*

- Market bootstrapping is hard*: a market starting above T^* cannot self-correct—it requires an exogenous information injection (regulation, certification, guarantees) to push T below the threshold. This explains why new markets for complex goods (health insurance, financial derivatives) require institutional scaffolding.*

- Established markets are resilient*: once $T < T^*$, the virtuous cycle of trade $\to$ information $\to$ lower T creates a buffer. Temporary information shocks (scandals, crises) are absorbed unless they push T above the threshold.*

- Low-ρ markets are more robust dynamically*: since $T^*(\rho)$ is increasing in $(1-\rho)$, specialized markets have a larger basin of attraction for the functioning equilibrium. They can absorb larger information shocks before collapsing—consistent with the static prediction but now with a dynamic mechanism.*

The static $T^*(\rho)$ surface from Proposition 9 is thus reinterpreted as the separatrix *of a bistable dynamical system. The framework's other derivations admit analogous dynamics: in mechanism design (Section 6), T evolves as agents learn to report truthfully; in social choice (Section 7), T evolves with institutional trust. These extensions connect the static framework to the dynamic models in the companion papers on endogenous decentralization and mesh equilibrium.*

6 Derivation II: Myerson's Optimal Mechanism

Myerson's virtual valuation (Myerson, 1981) is shown to be the free energy gradient, and ρ determines optimal mechanism structure.

6.1 Setup

A principal allocates resources among J agents with private types θ_j drawn independently from distribution F on $[\underline{\theta}, \bar{\theta}]$ with density f. The principal's objective is a CES aggregate of agent contributions:

$$\Phi = \left(\frac{1}{J} \sum_{j=1}^{J} x_j(\theta_j)^\rho \right)^{1/\rho} \tag{43}$$

where $x_j(\theta_j)$ is the allocation to agent j given their type. By Proposition 2(a), consistent sub-aggregation and homogeneity force the principal's objective to be CES: the principal can evaluate sub-groups independently and scale results proportionally only if the aggregator is a power mean. The parameter ρ measures how much the principal values balance across agents versus total output.

6.2 The Virtual Valuation as Free Energy Gradient

Myerson's virtual valuation for type θ is:

$$\varphi(\theta) = \theta - \frac{1 - F(\theta)}{f(\theta)} \tag{44}$$

The optimal mechanism allocates to agents with $\varphi(\theta) \geq 0$ and excludes those with $\varphi(\theta) < 0$.

Proposition 13 (Virtual Valuation Identification)**.** *The virtual valuation* (44) *is the free energy gradient:*

$$\varphi(\theta) = \underbrace{\theta}_{\partial\Phi/\partial\theta} - \underbrace{\frac{1 - F(\theta)}{f(\theta)}}_{T \cdot \partial H/\partial\theta} \tag{45}$$

The first term is the marginal CES contribution of type θ. The second term is the marginal entropy cost—the information rent required to elicit truthful reporting from type θ.

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 free energy 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 θ_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 θ is:

$$U(\hat{\theta}, \theta) = \theta \cdot x(\hat{\theta}) - t(\hat{\theta}) \tag{46}$$

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:

$$\max_{x(\cdot), t(\cdot)} \mathbb{E}_\theta \left[\Phi\big(x(\theta_1), \ldots, x(\theta_J)\big) + \sum_{j=1}^{J} t(\theta_j) \right] \tag{47}$$

subject to incentive compatibility (IC) and individual rationality (IR):

$$\text{IC:} \quad U(\theta) \geq U(\hat{\theta}, \theta) = \theta \cdot x(\hat{\theta}) - t(\hat{\theta}) \quad \forall\, \theta, \hat{\theta} \tag{48}$$

$$\text{IR:} \quad U(\theta) \geq 0 \quad \forall\, \theta \tag{49}$$

Step 2: IC implies the information rent is the integral of the allocation rule. By standard arguments (Myerson 1981, Lemma), IC requires that $x(\theta)$ is non-decreasing in

θ and the envelope condition holds. From (46), truth-telling requires:

$$\theta \in \arg\max_{\hat{\theta}} \; \theta \cdot x(\hat{\theta}) - t(\hat{\theta})$$

Taking the first-order condition with respect to $\hat{\theta}$ and evaluating at $\hat{\theta} = \theta$:

$$\frac{dU}{d\theta} = x(\theta)$$

This is the envelope theorem applied to the agent's problem. Integrating from the lowest type $\underline{\theta}$:

$$U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} x(s)\, ds \tag{50}$$

IR binds at the lowest type—$U(\underline{\theta}) = 0$—since the principal minimizes rents. Therefore $U(\theta) = \int_{\underline{\theta}}^{\theta} x(s)\, ds$.

Substituting $t(\theta) = \theta \cdot x(\theta) - U(\theta)$ into (47) and focusing on a single agent (by independence across agents), the principal's per-agent expected transfer is:

$$\begin{aligned}\mathbb{E}[t(\theta)] &= \mathbb{E}\left[\theta \cdot x(\theta) - \int_{\underline{\theta}}^{\theta} x(s)\, ds\right] \\ &= \int_{\underline{\theta}}^{\bar{\theta}} \theta \cdot x(\theta) f(\theta)\, d\theta - \int_{\underline{\theta}}^{\bar{\theta}} \left(\int_{\underline{\theta}}^{\theta} x(s)\, ds\right) f(\theta)\, d\theta \end{aligned} \tag{51}$$

For the double integral, reverse the order of integration. The inner integral $\int_{\underline{\theta}}^{\theta} x(s)\, ds$ is integrated over $\theta \in [s, \bar{\theta}]$:

$$\begin{aligned}\int_{\underline{\theta}}^{\bar{\theta}} \left(\int_{\underline{\theta}}^{\theta} x(s)\, ds\right) f(\theta)\, d\theta &= \int_{\underline{\theta}}^{\bar{\theta}} x(s) \left(\int_{s}^{\bar{\theta}} f(\theta)\, d\theta\right) ds \\ &= \int_{\underline{\theta}}^{\bar{\theta}} x(s)\big[1 - F(s)\big]\, ds \end{aligned} \tag{52}$$

Substituting (52) back into (51):

$$\mathbb{E}[t(\theta)] = \int_{\underline{\theta}}^{\bar{\theta}} \left[\theta - \frac{1 - F(\theta)}{f(\theta)}\right] x(\theta) f(\theta)\, d\theta = \int_{\underline{\theta}}^{\bar{\theta}} \varphi(\theta) \cdot x(\theta) \cdot f(\theta)\, d\theta \tag{53}$$

The principal thus maximizes $\mathbb{E}\big[\Phi(x(\theta_1), \ldots, x(\theta_J))\big] + \sum_j \mathbb{E}[\varphi(\theta_j) \cdot x(\theta_j)]$, where the term $\varphi(\theta) = \theta - (1 - F(\theta))/f(\theta)$ acts as the effective marginal value of type θ net of information rents.

Step 3: The inverse Mills ratio as entropy gradient. Define the *cumulative residual*

entropy (CRE) associated with the type distribution:

$$\mathcal{H}(\theta) \equiv -\int_{\underline{\theta}}^{\bar{\theta}} \big[1 - F(s)\big] \log\big[1 - F(s)\big]\, ds \tag{54}$$

This is the continuous analogue of Shannon entropy applied to the survival function $\bar{F}(\theta) = 1 - F(\theta)$. It measures the residual uncertainty about types above any given threshold.

The information rent from (52) can be written as:

$$\mathbb{E}[U(\theta)] = \int_{\underline{\theta}}^{\bar{\theta}} x(\theta)\big[1 - F(\theta)\big]\, d\theta \tag{55}$$

The integrand $[1 - F(\theta)]$ is the *hazard weight*: the mass of types above θ who benefit from the information rent generated at θ. Dividing by $f(\theta)$ gives the per-unit-density rent $(1 - F(\theta))/f(\theta)$, which is the inverse of the hazard rate $h(\theta) = f(\theta)/(1 - F(\theta))$.

To connect this to Shannon entropy, consider the differential entropy of the type distribution truncated above θ:

$$H_\theta \equiv -\int_{\theta}^{\bar{\theta}} \frac{f(s)}{1 - F(\theta)} \log \frac{f(s)}{1 - F(\theta)}\, ds \tag{56}$$

This is the Shannon entropy of the conditional distribution $f(s \mid s \geq \theta)$. Computing $dH_\theta/d\theta$ and evaluating:

$$\frac{dH_\theta}{d\theta} = \frac{f(\theta)}{1 - F(\theta)} \left[\log \frac{f(\theta)}{1 - F(\theta)} + 1 + H_\theta\right] \tag{57}$$

At the point of zero truncation ($\theta = \underline{\theta}$, where $F(\underline{\theta}) = 0$), the marginal change in conditional entropy per unit type is scaled by $f(\theta)/(1 - F(\theta)) = h(\theta)$, the hazard rate. Inverting: the *entropy contribution per agent at type* θ is:

$$\frac{1}{h(\theta)} = \frac{1 - F(\theta)}{f(\theta)} \tag{58}$$

This is exactly the inverse Mills ratio. It measures how much residual uncertainty each type contributes: low-hazard types (those in the upper tail) contribute more entropy and therefore command higher information rents.

Setting the "temperature" $T = 1$ (natural units where one unit of information rent costs one unit of surplus), the entropy gradient term is:

$$T \cdot \frac{\partial \mathcal{H}}{\partial \theta}\bigg|_{\text{per density}} = \frac{1 - F(\theta)}{f(\theta)} \tag{59}$$

Step 4: Assembly as the free energy gradient. From Step 2, the principal's pointwise objective is to maximize (over each agent's allocation separately, given CES separability at the margin):

$$\mathcal{F}(\theta) = \underbrace{\theta \cdot x(\theta)}_{\text{CES marginal contribution } \partial\Phi/\partial\theta} - \underbrace{\frac{1-F(\theta)}{f(\theta)} \cdot x(\theta)}_{\text{entropy cost } T\cdot\partial\mathcal{H}/\partial\theta} \tag{60}$$

The free energy gradient with respect to θ is therefore:

$$\frac{\partial \mathcal{F}}{\partial \theta} = \theta - \frac{1-F(\theta)}{f(\theta)} = \varphi(\theta) \tag{61}$$

which is Myerson's virtual valuation.

Explicit computation for the uniform distribution. Let $F(\theta) = (\theta - \underline{\theta})/(\bar{\theta} - \underline{\theta})$ on $[\underline{\theta}, \bar{\theta}]$, so $f(\theta) = 1/(\bar{\theta} - \underline{\theta})$. Then:

$$\frac{1-F(\theta)}{f(\theta)} = \frac{\bar{\theta} - \theta}{1/(\bar{\theta} - \underline{\theta})} \cdot \frac{1}{\bar{\theta} - \underline{\theta}} = \bar{\theta} - \theta$$

Hence the virtual valuation is:

$$\varphi(\theta) = \theta - (\bar{\theta} - \theta) = 2\theta - \bar{\theta} \tag{62}$$

This is non-negative iff $\theta \geq \bar{\theta}/2$—so the optimal mechanism excludes the bottom half of the type distribution. For the uniform case with $\underline{\theta} = 0$: $\varphi(\theta) = 2\theta - \bar{\theta}$, the reserve type is $\theta^* = \bar{\theta}/2$, which is Myerson's classical result.

The conditional entropy of the uniform distribution above θ is $H_\theta = \log(\bar{\theta} - \theta)$, and $dH_\theta/d\theta = 1/(\bar{\theta} - \theta)$. The entropy cost per density is $(1/f) \cdot (dH_\theta/d\theta)^{-1} \cdot f = \bar{\theta} - \theta$, confirming the identification. □

6.3 The Revelation Principle as Thermodynamic Equilibrium

Proposition 14 (Thermodynamic Revelation)**.** *A mechanism is incentive-compatible if and only if truth-telling minimizes each agent's free energy:*

$$\mathcal{F}_j(\textit{report } \theta') = -u_j(\theta, x(\theta')) + t(\theta') + T \cdot \mathrm{KL}(\theta' \| \theta) \tag{63}$$

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

The revelation principle—that any implementable outcome can be achieved through

truthful direct revelation—is the economic analogue of the thermodynamic statement that equilibrium minimizes free energy at the point of maximum order (zero entropy of misrepresentation).

6.4 ρ-Dependent Mechanism Structure

For CES allocation problems, the CES marginal value of agent j is:

$$\frac{\partial \Phi}{\partial x_j} = \frac{1}{J} \cdot x_j^{\rho-1} \cdot \Phi^{1-\rho} \tag{64}$$

The optimal mechanism equates (64) with the entropy cost:

$$\frac{1}{J} \cdot x_j^{\rho-1} \cdot \Phi^{1-\rho} = T \cdot \frac{1 - F(\theta_j)}{f(\theta_j)} \tag{65}$$

Proposition 15 (Price of Incentive Compatibility)**.** *Define the price of incentive compatibility as the welfare ratio:*

$$PoIC(\rho) = \frac{W^{\textit{first-best}} - W^{\textit{second-best}}}{W^{\textit{first-best}}}$$

Then $PoIC(\rho)$ is decreasing in ρ:

(a) *As $\rho \to 1$ (perfect substitutes): $PoIC \to 0$. Agents are replaceable; information rents vanish.*

(b) *As $\rho \to -\infty$ (perfect complements): $PoIC \to 1$. 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. To obtain closed forms, we work with the uniform distribution $F(\theta) = (\theta - \underline{\theta})/(\bar{\theta} - \underline{\theta})$ on $[\underline{\theta}, \bar{\theta}]$ with $\underline{\theta} > 0$, and verify that the qualitative results hold for general regular distributions. For clarity, treat a single agent ($J = 1$) where the CES structure reduces to the principal valuing agent output at x^ρ/ρ (equivalently, $\Phi = x^\rho$ up to constants); the multi-agent case follows by independence and additivity of virtual surplus across agents.

Step 1: First-best allocation ($T = 0$, no IC constraint). With observable types, the principal solves:

$$\max_{x(\theta) \geq 0} \mathbb{E}_\theta[\theta \cdot x(\theta)^\rho - c \cdot x(\theta)], \qquad \rho < 1 \tag{66}$$

where the agent's type θ scales the CES contribution and $c > 0$ is the marginal cost of provision (or equivalently, the transfer the principal must pay: the principal can set $t(\theta)$ equal to the agent's outside option since types are observed). The first-order condition

for each θ:

$$\rho\,\theta \cdot x^{\rho-1} = c \quad \Longrightarrow \quad x^{\mathrm{FB}}(\theta) = \left(\frac{\rho\,\theta}{c}\right)^{1/(1-\rho)} \tag{67}$$

Normalizing $c = \rho$ (without loss of generality for the welfare ratio), the first-best allocation is:

$$x^{\mathrm{FB}}(\theta) = \theta^{1/(1-\rho)} = \theta^{\sigma} \tag{68}$$

where $\sigma = 1/(1-\rho)$ is the elasticity of substitution. The first-best per-type surplus is:

$$w^{\mathrm{FB}}(\theta) = \theta \cdot \left(\theta^{\sigma}\right)^{\rho} - \rho \cdot \theta^{\sigma} = \theta^{1+\sigma\rho} - \rho \cdot \theta^{\sigma} = \theta^{\sigma} - \rho \cdot \theta^{\sigma} = (1-\rho)\,\theta^{\sigma} \tag{69}$$

using $1 + \sigma\rho = 1 + \rho/(1-\rho) = 1/(1-\rho) = \sigma$. Expected first-best welfare:

$$W^{\mathrm{FB}} = (1-\rho)\int_{\underline{\theta}}^{\bar{\theta}} \theta^{\sigma} f(\theta)\,d\theta = (1-\rho)\cdot\mathbb{E}[\theta^{\sigma}] \tag{70}$$

For the uniform distribution on $[\underline{\theta}, \bar{\theta}]$ with $\Delta \equiv \bar{\theta} - \underline{\theta}$:

$$W^{\mathrm{FB}} = \frac{1-\rho}{\Delta}\int_{\underline{\theta}}^{\bar{\theta}} \theta^{\sigma}\,d\theta = \frac{1-\rho}{\Delta}\cdot\frac{\bar{\theta}^{\sigma+1} - \underline{\theta}^{\sigma+1}}{\sigma+1} \tag{71}$$

Step 2: Second-best allocation (with IC constraint). From the virtual valuation identification (Proposition 13), the principal maximizes virtual surplus. With the uniform distribution, $\varphi(\theta) = 2\theta - \bar{\theta}$ (from (62)). The second-best problem replaces θ with $\varphi(\theta)$ in the allocation rule:

$$\max_{x(\theta)\geq 0}\ \mathbb{E}_{\theta}[\varphi(\theta)\cdot x(\theta)^{\rho} - \rho\cdot x(\theta)] \tag{72}$$

For types with $\varphi(\theta) > 0$ (i.e., $\theta > \bar{\theta}/2 \equiv \theta^{*}$), the FOC gives:

$$x^{\mathrm{SB}}(\theta) = \varphi(\theta)^{\sigma} = (2\theta - \bar{\theta})^{\sigma}, \qquad \theta \geq \theta^{*} \tag{73}$$

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

The second-best *true welfare* (evaluated at true types θ, not virtual types) is:

$$W^{\mathrm{SB}} = \int_{\theta^{*}}^{\bar{\theta}} \left[\theta\cdot\left(\varphi(\theta)^{\sigma}\right)^{\rho} - \rho\cdot\varphi(\theta)^{\sigma}\right] f(\theta)\,d\theta \tag{74}$$

Using $(\varphi^{\sigma})^{\rho} = \varphi^{\sigma\rho} = \varphi^{\sigma-1}$ (since $\sigma\rho = \sigma - 1$) and $f(\theta) = 1/\Delta$:

$$W^{\mathrm{SB}} = \frac{1}{\Delta}\int_{\theta^{*}}^{\bar{\theta}} \left[\theta\cdot\varphi(\theta)^{\sigma-1} - \rho\cdot\varphi(\theta)^{\sigma}\right] d\theta \tag{75}$$

Step 3: The welfare ratio. The PoIC is:

$$\text{PoIC}(\rho) = 1 - \frac{W^{\text{SB}}}{W^{\text{FB}}} \tag{76}$$

To analyze this, decompose the welfare loss into two sources:

1. *Exclusion loss*: types $\theta \in [\underline{\theta}, \theta^*)$ are excluded ($x^{\text{SB}} = 0$), destroying surplus $(1-\rho)\theta^\sigma$ for each excluded type.

2. *Distortion loss*: types $\theta \in [\theta^*, \bar{\theta}]$ receive $x^{\text{SB}}(\theta) = \varphi(\theta)^\sigma < \theta^\sigma = x^{\text{FB}}(\theta)$ (since $\varphi(\theta) < \theta$ for all $\theta < \bar{\theta}$), reducing surplus per participating type.

The key observation is that both losses depend on $\sigma = 1/(1-\rho)$, and their magnitude relative to W^{FB} is controlled by how much the mapping $\theta \mapsto \varphi(\theta)$ distorts the allocation rule $x = \theta^\sigma$.

Step 4: Monotonicity in ρ and the two limits.

(a) Limit $\rho \to 1$ ($\sigma \to \infty$): PoIC $\to 0$.

As $\rho \to 1$, $\sigma \to \infty$. The first-best allocation $x^{\text{FB}}(\theta) = \theta^\sigma$ becomes increasingly concentrated: for any two types $\theta_1 < \theta_2 \leq \bar{\theta}$, the ratio $x^{\text{FB}}(\theta_1)/x^{\text{FB}}(\theta_2) = (\theta_1/\theta_2)^\sigma \to 0$. In the limit, only the highest type $\bar{\theta}$ receives a nonzero allocation.

Simultaneously, the second-best allocation has the same concentration: $x^{\text{SB}}(\theta) = \varphi(\theta)^\sigma \to 0$ for all $\theta < \bar{\theta}$, and $x^{\text{SB}}(\bar{\theta}) = \varphi(\bar{\theta})^\sigma = \bar{\theta}^\sigma$ (since $\varphi(\bar{\theta}) = 2\bar{\theta} - \bar{\theta} = \bar{\theta}$ for the uniform distribution). Therefore the welfare ratio $W^{\text{SB}}/W^{\text{FB}}$ approaches:

$$\frac{W^{\text{SB}}}{W^{\text{FB}}} \;\to\; \frac{\bar{\theta} \cdot \bar{\theta}^{\sigma-1} - \rho \cdot \bar{\theta}^\sigma}{\bar{\theta} \cdot \bar{\theta}^{\sigma-1} - \rho \cdot \bar{\theta}^\sigma} = 1 \tag{77}$$

as both numerator and denominator are dominated by $\theta = \bar{\theta}$, at which type the first-best and second-best allocations coincide (the highest type earns zero information rent). Hence PoIC $\to 0$.

Economic intuition: When $\rho \to 1$ (perfect substitutes), the principal optimally gives everything to the best agent regardless. Screening is unnecessary—the allocation is essentially a winner-take-all rule—so the IC constraint is non-binding and information rents vanish.

(b) Limit $\rho \to -\infty$ ($\sigma \to 0^+$): PoIC $\to 1$.

As $\rho \to -\infty$, $\sigma = 1/(1-\rho) \to 0^+$. The first-best allocation $x^{\text{FB}}(\theta) = \theta^\sigma \to 1$ for all $\theta > 0$: under perfect complements (Leontief), the principal wants equal allocation across all agents regardless of type. First-best welfare:

$$W^{\text{FB}} = (1-\rho)\,\mathbb{E}[\theta^\sigma] \;\to\; (1-\rho) \cdot 1 = 1 - \rho \;\to\; \infty \tag{78}$$

but the per-type surplus $(1-\rho)\theta^\sigma \approx (1-\rho)$ is uniform.

The second-best allocation $x^{\mathrm{SB}}(\theta) = \varphi(\theta)^\sigma$ for $\theta \geq \theta^*$, where $\varphi(\theta)^\sigma \to 1$ for all $\theta > \theta^*$ as $\sigma \to 0^+$ (since $a^\sigma \to 1$ for any $a > 0$). However, the exclusion region persists: types $\theta < \theta^* = \bar{\theta}/2$ receive $x = 0$. The fraction of excluded types is $F(\theta^*) = (\theta^* - \underline{\theta})/\Delta$.

The second-best welfare from participating types has per-type surplus also converging to $(1 - \rho)$ (since distortion vanishes as $\sigma \to 0$), but only the fraction $1 - F(\theta^*)$ of types participate. However, under Leontief complementarity the key effect is different: the principal's objective $\Phi = \min_j x_j$ means that the weakest agent's allocation determines total output. In the mechanism, the lowest participating type has $\varphi(\theta^*) = 0$ and receives $x^{\mathrm{SB}}(\theta^*) = 0$, which under Leontief aggregation collapses the entire output to zero.

More precisely, for $\rho \ll 0$ the CES aggregate $\Phi = (J^{-1} \sum x_j^\rho)^{1/\rho}$ is dominated by the smallest x_j. The mechanism must either exclude low types (losing essential inputs under complementarity) or pay enormous rents to include them. In the second-best optimum, the information rent to type θ is $U(\theta) = \int_{\underline{\theta}}^{\theta} x^{\mathrm{SB}}(s)\, ds$, which for $\sigma \to 0^+$ approaches $\int_{\theta^*}^{\theta} 1\, ds = \theta - \theta^*$ for $\theta \geq \theta^*$. The total expected rent:

$$\mathbb{E}[U] = \int_{\theta^*}^{\bar{\theta}} (\theta - \theta^*) f(\theta)\, d\theta = \frac{(\bar{\theta} - \theta^*)^2}{2\Delta} \tag{79}$$

As $|\rho| \to \infty$, the per-unit surplus $(1-\rho)\theta^\sigma \approx (1-\rho)\cdot 1$ grows linearly in $|\rho|$, but the welfare loss from excluding $[\underline{\theta}, \theta^*)$—which contains agents essential to the Leontief aggregate—grows at the same rate. Since the bottleneck agent is always near the exclusion threshold, second-best welfare as a fraction of first-best welfare vanishes:

$$\frac{W^{\mathrm{SB}}}{W^{\mathrm{FB}}} \;\to\; 0 \quad \text{as } \rho \to -\infty \tag{80}$$

Hence PoIC $\to 1$.

Economic intuition: Under perfect complements, every agent is essential. But the mechanism must screen types, which requires excluding some agents or paying large rents. Excluding even one essential complement destroys the entire surplus. Information rents therefore consume all welfare in the limit.

(c) Monotonicity of PoIC(ρ).

We establish that $d\,\mathrm{PoIC}/d\rho < 0$ for all $\rho < 1$. The distortion at each type θ is measured by the ratio:

$$r(\theta; \sigma) \equiv \frac{x^{\mathrm{SB}}(\theta)}{x^{\mathrm{FB}}(\theta)} = \left(\frac{\varphi(\theta)}{\theta}\right)^\sigma \tag{81}$$

for $\theta \geq \theta^*$ (and $r = 0$ for $\theta < \theta^*$). Since $\varphi(\theta) < \theta$ for all $\theta \in (\theta^*, \bar{\theta})$ (with equality only at $\bar{\theta}$), the ratio $\varphi(\theta)/\theta < 1$ and therefore:

$$\frac{\partial r}{\partial \sigma} = \left(\frac{\varphi(\theta)}{\theta}\right)^\sigma \log\left(\frac{\varphi(\theta)}{\theta}\right) < 0 \tag{82}$$

since $\log(\varphi/\theta) < 0$. As σ increases (i.e., ρ increases toward 1), the distortion ratio r at each interior type $\theta < \bar{\theta}$ decreases toward zero, while $r(\bar{\theta}) = 1$ is unchanged. However, the CES welfare weight $\theta^\sigma f(\theta)$ shifts toward $\bar{\theta}$ as σ grows: the aggregate is increasingly dominated by the highest type, where there is no distortion. The net effect on welfare is therefore positive—the second-best welfare ratio $W^{\mathrm{SB}}/W^{\mathrm{FB}}$ increases toward 1.

Additionally, the exclusion threshold $\theta^* = \bar{\theta}/2$ is independent of ρ for the uniform distribution, so the exclusion set does not change. However, the welfare weight of excluded types in the CES aggregate is:

$$\frac{\int_{\underline{\theta}}^{\theta^*} \theta^\sigma f(\theta)\, d\theta}{\int_{\underline{\theta}}^{\bar{\theta}} \theta^\sigma f(\theta)\, d\theta} \tag{83}$$

which is decreasing in σ (as σ increases, the integral is increasingly dominated by high types, making the excluded low types less important).

Both effects—reduced per-type distortion and reduced welfare weight of excluded types—work in the same direction: $W^{\mathrm{SB}}/W^{\mathrm{FB}}$ is increasing in σ (equivalently, increasing in ρ). Therefore $\mathrm{PoIC}(\rho) = 1 - W^{\mathrm{SB}}/W^{\mathrm{FB}}$ is decreasing in ρ.

Remark on general distributions. For a general regular distribution (i.e., $\varphi(\theta)$ non-decreasing), the exclusion threshold θ^* solving $\varphi(\theta^*) = 0$ may depend on F, but the qualitative structure is unchanged: (i) the distortion ratio (81) is decreasing in σ at every interior type; (ii) the welfare weight of excluded types in the CES aggregate is decreasing in σ; (iii) at $\sigma \to \infty$ the allocation concentrates on the highest type where $\varphi(\bar{\theta}) = \bar{\theta}$ (no distortion); (iv) at $\sigma \to 0^+$ the Leontief bottleneck makes exclusion catastrophic. The monotonicity and both limits therefore hold for all regular type distributions. □

6.5 Empirical Implications

1. **Mechanism format is determined by ρ.** Low ρ (complementary agents): optimal mechanism resembles a scored evaluation—principal invests in distinguishing types (R&D contracting, academic tenure). High ρ (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 ρ: threshold is low (include more types, pay more rents). High ρ: threshold is high (exclude aggressively, keep rents low). Same ρ-dependent pattern as Section 5.

3. **Defense procurement vs. commodity purchasing.** Defense contracting (highly complementary, specialized inputs, low ρ) should exhibit high PoIC—and does (notorious cost overruns). Commodity purchasing (substitutable inputs, high ρ) should exhibit low PoIC—and does (efficient competitive procurement).

7 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. The free energy framework does not dissolve this impossibility but rather characterizes what becomes possible when cardinal utility information is available and agents optimize under finite processing capacity ($T > 0$). The CES family requires cardinal input—this is the price of non-dictatorial aggregation—and ρ determines which democratic institutions are robust to informational noise.

7.1 Arrow's Conditions

Arrow proved that no social welfare function simultaneously satisfies:

1. **Unrestricted domain** (U): defined for all preference profiles;
2. **Pareto efficiency** (P): if all agents prefer A to B, so does the social ranking;
3. **Independence of irrelevant alternatives** (IIA): the social ranking of A vs. B depends only on individual rankings of A vs. B;
4. **Non-dictatorship** (ND): no single agent always determines the social ranking.

Any aggregation satisfying U, P, and IIA must be dictatorial.

7.2 CES Social Welfare at $T = 0$

The CES social welfare function:

$$W = \left(\frac{1}{J} \sum_{j=1}^{J} u_j^{\rho} \right)^{1/\rho} \tag{84}$$

uses cardinal utilities. For $\rho = 1$: utilitarian (sum). For $\rho \to -\infty$: Rawlsian (max-min). For $\rho = 0$: Nash welfare (geometric mean). This is the Atkinson family (Atkinson, 1970). By Proposition 2(c), this is the *unique* family satisfying Pareto, anonymity, homotheticity, and Pigou–Dalton: any social planner with even minimal inequality aversion who respects proportional scaling must use CES.

CES social welfare violates Arrow's *ordinal* IIA: since it uses cardinal utility levels (not just ordinal rankings), two profiles that agree on every agent's ranking of A vs. B but differ in cardinal magnitudes can produce different social rankings. Arrow's theorem states that at $T = 0$ (deterministic, ordinal preferences), no aggregation satisfying ordinal IIA and Pareto can be non-dictatorial. The CES family sidesteps this by requiring cardinal information—the price paid is measured by the parameter $1 - \rho$.

7.3 $T > 0$: The Phase Transition

At $T > 0$, agents have probabilistic preferences. Instead of the deterministic ordering $A >_j B$, agent j has:

$$P_j(A \succ B) = \frac{\exp(u_j(A)/T)}{\exp(u_j(A)/T) + \exp(u_j(B)/T)} = \frac{1}{1 + \exp(-(u_j(A) - u_j(B))/T)} \tag{85}$$

Arrow's theorem stands: no ordinal aggregation escapes impossibility. But at $T > 0$ with cardinal utilities, the operative question changes from "can we satisfy ordinal IIA?" to "how closely can a non-dictatorial cardinal aggregation approximate ordinal IIA?" Three features of the $T > 0$ setting reshape the problem:

Cardinal information becomes available: Noisy optimization reveals cardinal utility differences through choice probabilities (85). This is what breaks the ordinal straitjacket—not a weakening of Arrow's axioms but a richer informational environment in which CES aggregation operates.

Transitivity becomes stochastic: Preference cycles $P(A \succ B) > 1/2$, $P(B \succ C) > 1/2$, $P(C \succ A) > 1/2$ can hold simultaneously, consistent with noisy optimization. Arrow's framework excludes such cycles by construction; admitting them enlarges the space of feasible aggregation rules.

Influence becomes continuous: A high-weight voter at $T > 0$ is statistically influential but not deterministic. The binary dictator/non-dictator distinction gives way to a continuum of influence weights, indexed by ρ.

Proposition 16 (Democratic Feasibility at Positive Temperature)**.** *For $\rho \in (0, 1]$, any $\varepsilon > 0$, 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-parameterisation is bounded by $(1 - \rho) \cdot \eta(T)$, where $\eta(1) = 0$ for all T and $\eta(T) \to 0$ as $T \to \infty$;*
4. ***Concentration**: the social ordering agrees with the expected CES-optimal ordering with probability at least $1 - \delta(T, \varepsilon, J)$.*

The tradeoff: aggregation error grows with T while ordinal IIA violation shrinks with T. 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 $\tilde{u}_j(A) = u_j(A) + \xi_j^A$. (This is the random utility model underlying (85).) Since $\Pr(\xi_j^A < -\underline{u}) = 1/(1 + e^{\underline{u}/T}) \leq e^{-\underline{u}/T}$, the event $\mathcal{E} = \{\tilde{u}_j(A) > 0 \text{ for all } j, A\}$ satisfies:

$$\Pr(\mathcal{E}) \geq 1 - J|\mathcal{A}| \cdot e^{-\underline{u}/T} \tag{86}$$

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

$$W_T(A) = \left(\frac{1}{J} \sum_{j=1}^{J} \tilde{u}_j(A)^\rho \right)^{1/\rho} \tag{87}$$

which is well-defined and positive for $\rho \in (0, 1]$. The social ranking orders alternatives by the expected social welfare $S(A) \equiv \mathbb{E}[W_T(A) \cdot \mathbf{1}_\mathcal{E}]$. Since all agents enter symmetrically (equal weight $1/J$), no agent is a dictator.

Step 2: Pareto efficiency. Suppose all agents prefer A to B: $u_j(A) > u_j(B)$ for all j. We show $S(A) > S(B)$.

On the event $\mathcal{E}$, the CES aggregate $W_T(A)$ is strictly increasing in each $\tilde{u}_j(A)$: by direct computation,

$$\frac{\partial W_T(A)}{\partial \tilde{u}_k(A)} = \frac{1}{J} \left(\frac{\tilde{u}_k(A)}{W_T(A)} \right)^{\rho - 1} > 0 \tag{88}$$

since $\tilde{u}_k(A) > 0$ and $W_T(A) > 0$ on $\mathcal{E}$. Now consider the function $g(v) \equiv \mathbb{E}[W_T(A) \cdot \mathbf{1}_\mathcal{E} \mid u_k(A) = v]$, holding all other utilities fixed. Since $\tilde{u}_k(A) = v + \xi_k^A$ and $W_T(A)$ is increasing in $\tilde{u}_k(A)$ on $\mathcal{E}$, we have: for $v' > v$, the random variable $W_T(A)|_{u_k = v'}$ pointwise dominates $W_T(A)|_{u_k = v}$ on $\mathcal{E}$ (same noise realization, larger input). Therefore $g(v') > g(v)$: the social score $S(A)$ is strictly increasing in each $u_k(A)$.

Since $u_j(A) > u_j(B)$ for all j, applying this monotonicity J times (once per agent) gives $S(A) > S(B)$: Pareto efficiency holds exactly.

Step 3: Approximate IIA. Arrow's IIA requires that the social ranking of A vs. B depend only on agents' *ordinal* rankings of A vs. B—not on the cardinal utility levels. We first note what holds exactly, then bound the residual ordinal violation.

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 $(u_1(A), \ldots, u_J(A))$, not on utilities for any other alternative C. Changing or removing C from the menu has zero effect on $S(A)$ or $S(B)$. The social ranking of A vs. B is therefore 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. Two profiles u, u' that preserve every agent's ordinal ranking of A vs. B (i.e., $\text{sign}(u_j(A) - u_j(B)) = \text{sign}(u'_j(A) - u'_j(B))$ for all j) but differ in cardinal levels can yield different social rankings. We now bound this sensitivity.

Let $u'_j(X) = u_j(X) + c_j$ for agent-specific constants c_j (a re-cardinalisation preserving all ordinal rankings). The social welfare gap changes by:

$$[S(A; u') - S(B; u')] - [S(A; u) - S(B; u)]$$

By the mean value theorem applied to $S(X; u + tc)$ on $\mathcal{E}$, with the CES gradient (88):

$$\left| \frac{\partial}{\partial t} \big[S(A; u + tc) - S(B; u + tc) \big] \right| = \frac{1}{J} \left| \sum_{j=1}^{J} c_j \cdot \mathbb{E} \left[\left(\frac{\tilde{u}_j(A)}{W_T(A)} \right)^{\rho - 1} - \left(\frac{\tilde{u}_j(B)}{W_T(B)} \right)^{\rho - 1} \right] \right| \tag{89}$$

For $\rho = 1$ (utilitarian), each term in the expectation equals $1 - 1 = 0$: the gap is invariant to re-cardinalisation, and ordinal IIA holds exactly. For $\rho < 1$, the weight $(x/W)^{\rho-1}$ amplifies agents with $x < W$ and suppresses agents with $x > W$. At $T > 0$, the noise symmetrises these weights: as $T \to \infty$, the noise ξ_j dominates u_j, so $\tilde{u}_j(A)/W_T(A) \to 1$ in probability for all j, and each term converges to $1^{\rho-1} - 1^{\rho-1} = 0$. Therefore:

$$\big| [S(A; u') - S(B; u')] - [S(A; u) - S(B; u)] \big| \leq \|c\|_\infty \cdot \eta(\rho, T) \tag{90}$$

where $\eta(\rho, T) \to 0$ as $T \to \infty$ and $\eta(1, T) = 0$ for all T. An explicit bound is obtained by expanding the CES weights to second order around the noise-dominated regime: $\eta(\rho, T) \leq (1 - \rho) \cdot (\bar{u} - \underline{u})^2 / (3T^2)$ for $T \gg \bar{u}$.

In the practically relevant regime $T \ll \bar{u}$ (low noise), the IIA violation is bounded by the CES nonlinearity:

$$\eta(\rho, T) \leq \frac{2(1 - \rho)(\bar{u} - \underline{u})}{J \, \underline{u}}$$

which depends on ρ and the utility heterogeneity but not on T. The key point is that the violation is controlled by $(1 - \rho)$: it vanishes at $\rho = 1$ and grows as $\rho \to 0$, consistent with the interpretation that more egalitarian (lower ρ) aggregation rules extract more cardinal information and therefore deviate further from ordinal IIA.

Step 4: Concentration of the social ordering. By the strong law of large numbers applied to the CES aggregate of i.i.d. noise perturbations, for fixed alternatives A, B:

$$W_T(A) \xrightarrow{J \to \infty} W_0(A) \quad \text{almost surely}$$

More precisely, by McDiarmid's bounded differences inequality (each agent's noisy utility

affects W_T by at most c/J for $c = O(\bar{u})$):

$$\Pr\big[|W_T(A) - \mathbb{E}[W_T(A)]| > \varepsilon\big] \leq 2\exp\left(-\frac{2J\varepsilon^2}{c^2}\right) \tag{91}$$

Taking a union bound over all $|\mathcal{A}|(|\mathcal{A}|-1)/2$ pairwise comparisons, the probability that the noisy social ordering differs from the expected CES ordering on any pair is at most:

$$\delta(T, \varepsilon) = |\mathcal{A}|^2 \exp\left(-\frac{2J\varepsilon^2}{c^2}\right)$$

As $T \to 0$: the noise vanishes and $\delta \to 0$ (the realized ranking converges to the expected ranking), but the ordinal IIA violation remains at its maximum—the CES ranking depends fully on cardinal levels, and Arrow's impossibility applies. As $T \to \infty$: the noise dominates, ordinal IIA violation vanishes (CES smooths toward utilitarian), but $\delta \to 1$ in the sense that the ranking carries no information about the true welfare ordering.

Step 5: Existence of optimal T^*. Define the democratic error as:

$$E(T) = \underbrace{\mathbb{E}\big[\|W_T - W_0\|^2\big]}_{\text{aggregation error (increasing in } T)} + \underbrace{\lambda \cdot V(T)}_{\text{ordinal IIA cost}} \tag{92}$$

where $\lambda > 0$ weights the IIA requirement and $V(T)$ measures the ordinal IIA violation from (90):

$$V(T) = \sup_{\|c\|_\infty \leq \bar{u}} \big|[S(A; u+c) - S(B; u+c)] - [S(A; u) - S(B; u)]\big|$$

The two terms create a tension. The aggregation error is zero at $T = 0$ and grows as $O(T^2)$ (noise variance scaling). The ordinal IIA cost $V(T)$ is bounded but positive at $T = 0$ (the deterministic CES violation for $\rho \neq 1$) and decreases toward zero for $T \gg \bar{u}$ (noise smooths CES toward utilitarian). More precisely: at $T = 0$, Arrow's impossibility theorem forbids any non-dictatorial rule from achieving $V = 0$, so the tradeoff is strict.

Since $E(T) \to \lambda \cdot V(0) > 0$ as $T \to 0$ and $E(T) \to \infty$ as $T \to \infty$ (noise destroys information), and E is continuous on $(0, \infty)$, by the extreme value theorem there exists $T^* \in (0, \infty)$ minimizing $E(T)$. □

7.4 ρ-Dependent Democratic Robustness

Proposition 17 (Robustness of Political Institutions)**.** *The critical temperature $T^*_{democracy}$ above which democratic aggregation fails depends on ρ:*

$$T^*_{democracy}(\rho) \text{ is increasing in } \rho \tag{93}$$

High-ρ (majoritarian) systems tolerate more informational noise than low-ρ (consensus) systems.

Proof. The proof establishes that the sensitivity of the CES aggregate to noise is monotonically increasing as ρ decreases, so that higher noise (higher T) destroys low-ρ aggregation before high-ρ aggregation.

Step 1: CES sensitivity to individual perturbations. Consider the CES social welfare $W = (J^{-1}\sum u_j^\rho)^{1/\rho}$ with $u_j \in [\underline{u}, \bar{u}]$, $\underline{u} > 0$. The partial derivative with respect to a single agent's utility is:

$$\frac{\partial W}{\partial u_k} = \frac{1}{J}\left(\frac{u_k}{W}\right)^{\rho-1} \tag{94}$$

For $\rho = 1$: $\partial W/\partial u_k = 1/J$ for all k—equal influence regardless of utility level. For $\rho < 1$: the weight $(u_k/W)^{\rho-1}$ is larger when $u_k < W$ (the exponent $\rho - 1 < 0$ amplifies small values). As $\rho \to -\infty$: the CES aggregate converges to $\min_j u_j$, and $\partial W/\partial u_k \to \mathbf{1}_{k=\arg\min_j u_j}$—the minimum-utility agent has all the influence.

Step 2: Noise amplification at low ρ. At temperature T, each agent's reported utility is $\tilde{u}_j = u_j + \xi_j$ with $\xi_j \sim \text{Logistic}(0, T)$. The variance of the noisy CES aggregate is, by the delta method:

$$\text{Var}(W_T) \approx \sum_{k=1}^{J}\left(\frac{\partial W}{\partial u_k}\right)^2 \text{Var}(\xi_k) = \frac{T^2\pi^2}{3}\cdot\frac{1}{J^2}\sum_{k=1}^{J}\left(\frac{u_k}{W}\right)^{2(\rho-1)} \tag{95}$$

using $\text{Var}(\text{Logistic}(0,T)) = \pi^2T^2/3$.

Define the *influence concentration*:

$$\mathcal{C}(\rho) \equiv \frac{1}{J}\sum_{k=1}^{J}\left(\frac{u_k}{W}\right)^{2(\rho-1)} \tag{96}$$

For $\rho = 1$: $\mathcal{C}(1) = 1$ (uniform influence). For $\rho < 1$: agents with $u_k < W$ contribute disproportionately, so $\mathcal{C}(\rho) > 1$ whenever utilities are heterogeneous. As $\rho \to -\infty$: $\mathcal{C}(\rho) \to J$ (all influence on one agent), so the effective sample size drops to 1.

Therefore $\text{Var}(W_T) = (T^2\pi^2/3)\cdot\mathcal{C}(\rho)/J$, and the effective number of independent signals is $J_{\text{eff}}(\rho) = J/\mathcal{C}(\rho)$, which is decreasing in $(1-\rho)$.

Step 3: Critical temperature from concentration. Democratic aggregation "fails" when the noise variance exceeds the squared gap between the top two alternatives. Let $\Delta = |W(A) - W(B)|$ be the welfare gap at $T = 0$. The probability that noise reverses the ranking is approximately:

$$\Pr[W_T(B) > W_T(A)] \approx \Phi_{\text{normal}}\left(-\frac{\Delta}{\sqrt{2\,\text{Var}(W_T)}}\right) \tag{97}$$

Setting this equal to a failure threshold α and solving for T:

$$T^*_{\text{democracy}}(\rho) = \frac{\Delta \cdot \sqrt{3}}{\pi \cdot z_\alpha} \cdot \sqrt{\frac{J}{\mathcal{C}(\rho)}} = \frac{\Delta \cdot \sqrt{3}}{\pi \cdot z_\alpha} \cdot \sqrt{J_{\text{eff}}(\rho)} \tag{98}$$

where $z_\alpha = \Phi^{-1}(1-\alpha)$. Since $J_{\text{eff}}(\rho) = J/\mathcal{C}(\rho)$ is increasing in ρ (from Step 2), $T^*_{\text{democracy}}$ is increasing in ρ.

Step 4: Verification at the limits.

- $\rho = 1$ *(utilitarian):* $\mathcal{C}(1) = 1$, $J_{\text{eff}} = J$. The critical temperature scales as $\sqrt{J}$—standard $\sqrt{n}$ convergence of the sample mean. The system tolerates substantial noise.
- $\rho \to 0$ *(Nash welfare):* $\mathcal{C}(0) = J^{-1} \sum (u_k/W)^{-2}$. For heterogeneous utilities with $\min u_k \ll W$, this can be large, reducing J_{eff}.
- $\rho \to -\infty$ *(Rawlsian):* $\mathcal{C} \to J$, $J_{\text{eff}} \to 1$. The critical temperature is $O(1)$—independent of the number of voters. Even a large electorate cannot overcome noise when the aggregation rule depends on a single agent's report.

The monotonicity of $T^*_{\text{democracy}}$ in ρ follows from the monotonicity of $\mathcal{C}(\rho)$ in ρ, which in turn follows from the convexity of $x \mapsto x^{2(\rho-1)}$ for $\rho < 1$ combined with Jensen's inequality: as ρ decreases, the exponent $2(\rho-1)$ becomes more negative, amplifying the contribution of below-average utilities. □

7.5 Empirical Implications

Proposition 17 predicts which political systems are robust to increasing informational noise (polarization, misinformation, declining institutional trust):

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

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

7.6 Connection to the Condorcet Jury Theorem

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

8 Derivation IV: Search and Matching

The Diamond-Mortensen-Pissarides search framework (Diamond, 1982; Mortensen, 1982; Pissarides, 1985) is derived from the free energy principle, showing that ρ determines search duration, match quality, and the slope of the Beveridge curve.

8.1 Setup

Consider J workers and J firms, each characterized by a multidimensional type. Worker i has skill profile $s_i = (s_{i1}, \ldots, s_{iL})$ across L skill dimensions; firm j requires task profile $t_j = (t_{j1}, \ldots, t_{jL})$. Match quality between worker i and firm j is the CES aggregate of skill–task fit:

$$m(i,j) = \left(\frac{1}{L} \sum_{\ell=1}^{L} (s_{i\ell} \cdot t_{j\ell})^{\rho} \right)^{1/\rho} \tag{99}$$

where $\rho \in (-1, 1)$ controls the complementarity between worker skills and firm requirements. Low ρ (high complementarity): a surgeon needs *exactly* the right hospital—excellence in one dimension cannot substitute for deficiency in another. High ρ (high substitutability): a cashier fits many stores—skills are fungible across positions.

The CES curvature $K = (1-\rho)(L-1)/L$ governs match specificity. By Proposition 2, the CES form is not assumed but derived from the structure of heterogeneous production.

Remark 18. *The standard DMP matching function $M = A \cdot U^{\alpha} V^{1-\alpha}$ (Petrongolo and Pissarides, 2001) is Cobb-Douglas, the $\rho \to 0$ limit of CES. The framework generalizes this to arbitrary ρ, allowing the matching technology itself to reflect the complementarity structure of the labor market.*

8.2 $T = 0$: Frictionless Assignment

With perfect information, all match qualities $m(i, j)$ are observable. The planner solves the optimal assignment:

$$W_0 = \max_{\sigma \in \Pi_J} \sum_{i=1}^{J} m(i, \sigma(i)) \tag{100}$$

where Π_J is the set of permutations. For $\rho < 1$, the CES superadditivity premium ensures that positive assortative matching is optimal (Becker, 1973): high-skill workers match with high-requirement firms, because the CES nonlinearity makes complementary pairings disproportionately productive. The surplus W_0 is the first-best benchmark against which search frictions are measured.

8.3 $T > 0$: Costly Search as Entropy Reduction

Workers and firms cannot observe match quality without meeting. Each meeting reveals mutual information $I = H_{\text{prior}} - H_{\text{posterior}}$ about the quality of a specific pairing.

A worker begins with $H_0 = \log J$ bits of uncertainty about which firm is the best match. After n meetings, expected residual entropy is $H_0 - n \cdot \Delta H$, where ΔH is the information per meeting. The free energy of a search strategy with n meetings is:

$$\mathcal{F}(n) = \mathbb{E}\left[\max_{k \le n} m(i, j_k)\right] - T \cdot \left[H_0 - n \cdot \Delta H\right] \tag{101}$$

The first term is the expected quality of the best match found (the "energy"); the second is the information cost of search, where T is the per-unit cost of processing match information (Sims, 2003).

Optimal stopping: the worker accepts a match when the marginal quality gain from one additional meeting falls below the marginal information cost $T \cdot \Delta H$. This yields a reservation match quality q^* solving:

$$\frac{\partial}{\partial n} \mathbb{E}\left[\max_{k \le n} m(i, j_k)\right] \Bigg|_{q^*} = T \cdot \Delta H \tag{102}$$

The left side is the option value of continued search—the probability of finding a match exceeding q^* times the expected improvement. At $T = 0$: the worker searches exhaustively ($q^* = W_0/J$, the first-best assignment quality). At $T \to \infty$: the worker accepts the first offer ($q^* \to 0$, no search).

8.4 The ρ-Dependent Search Duration

Proposition 19 (Search Duration)**.** *The expected number of meetings before acceptance is:*

$$n^*(\rho, T) = \frac{K}{T} \cdot Q(\rho) \tag{103}$$

where $K = (1-\rho)(L-1)/L$ is the CES curvature and $Q(\rho)$ is a quality premium function that is increasing in $(1-\rho)$. Search duration is increasing in K (more complementary matches require longer search) and decreasing in T (higher information cost shortens search and accepts worse matches).

Proof. The proof proceeds in three steps.

Step 1: Reservation quality. From the free energy first-order condition (102), the reservation quality q^* equates the marginal benefit of continued search to the marginal information cost. For match qualities drawn from the CES distribution (99), 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 expected improvement conditional on exceeding q^* is $\mathbb{E}[m - q^* \mid m > q^*]$. The FOC becomes:

$$[1 - G(q^*)] \cdot \mathbb{E}[m - q^* \mid m > q^*] = T \cdot \Delta H \tag{104}$$

Step 2: CES curvature raises the reservation quality. By Theorem 1(a), the CES aggregate of skill–task fit (99) satisfies, for any deviation δ from the balanced profile with $\sum \delta_\ell = 0$:

$$m(\bar{s} + \delta) \leq m(\bar{s}) - \frac{K}{2(L-1)} \cdot \frac{\|\delta\|^2}{\bar{s}} \tag{105}$$

where $K = (1-\rho)(L-1)/L$. This means a well-matched pairing (small $\|\delta\|$) is *disproportionately* more productive than a poorly-matched one: the surplus gap between a match of quality m_1 and a worse match $m_2 < m_1$ exceeds the linear difference $m_1 - m_2$ by a term of order $K(m_1^2 - m_2^2)$. The option value of continued search—which depends on the expected gain $\mathbb{E}[m - q^* \mid m > q^*]$—therefore increases in K, because the right tail of the match quality distribution contributes more surplus when complementarity is high.

Substituting into the FOC (102): the marginal benefit of one more meeting (left side) grows with K, while the marginal cost $T \cdot \Delta H$ (right side) is independent of K. Equating, the reservation quality satisfies:

$$q^*(\rho, T) = q_0^*(T) + K \cdot Q_1(\rho, T) + O(K^2) \tag{106}$$

where $q_0^*(T)$ is the reservation quality at $K = 0$ (perfect substitutes, standard DMP) and $Q_1 > 0$ captures the superadditivity-driven increase. The positivity of Q_1 follows directly: since the option value is increasing in K while the cost is not, the optimal

stopping boundary rises.

Step 3: Duration from acceptance probability. The expected search duration is $n^* = 1/P(m > q^*)$, since each meeting independently draws from G. Substituting the expansion (106) and Taylor-expanding $P(m > q^*) = 1 - G(q^*)$ around q_0^*:

$$1 - G(q_0^* + KQ_1) = [1 - G(q_0^*)] - KQ_1 \cdot g(q_0^*) + O(K^2) \tag{107}$$

where $g = G'$ is the density. Inverting:

$$n^* = \frac{1}{1 - G(q_0^*)} \cdot \frac{1}{1 - \frac{KQ_1 g(q_0^*)}{1 - G(q_0^*)}} + O(K^2) = n_0^* + \frac{KQ_1 g(q_0^*)}{[1 - G(q_0^*)]^2} + O(K^2) \tag{108}$$

At $K = 0$: $n_0^* = 1/[1 - G(q_0^*)]$ is the standard DMP duration, which depends only on T through $q_0^*(T)$. In the standard model, higher information cost reduces reservation quality. Under log-concave match quality distributions (including exponential, normal, and uniform), the standard reservation wage equation (Mortensen, 1982) gives $q_0^* \propto 1/T$ for small T: cheap information raises the reservation quality because the option value of continued search increases. Specifically, $q_0^* = c(G)/T$ where $c(G)$ is a distribution-dependent constant, so $n_0^* = 1/[1 - G(c(G)/T)] \sim T/c(G)$ for small T. The qualitative result $n^* \propto K/T$ holds for any match distribution where n_0^* is increasing in T. The K-correction term is proportional to $K \cdot Q_1(T)/T^2$. Defining $Q(\rho) \equiv Q_1 \cdot g(q_0^*)/[1 - G(q_0^*)]^2$ absorbs the distributional constants, and noting that Q_1/T scales as $1/T$ (from the FOC), the leading behavior is $n^* \propto K/T \cdot Q(\rho)$ as stated. □

Remark 20. *The result has a clean economic interpretation: workers in complementary labor markets (low ρ, 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.*

8.5 The Beveridge Curve as Free Energy Surface

Proposition 21 (Beveridge Curve Slope)**.** *In the (U, V) plane, the equilibrium Beveridge curve slope receives a CES correction from the acceptance channel:*

$$\left.\frac{dV}{dU}\right|_{BC} = -1 - \frac{2s}{f(1) \cdot p_a} \tag{109}$$

where s is the separation rate, $f(1)$ the meeting rate at balanced tightness, and $p_a = T/(K \cdot Q)$ is the acceptance probability from Proposition 19. Substituting the acceptance probability, the slope steepens as K/T increases: at $K = 0$ (perfect substitutes), $p_a = 1$ and the slope reduces to the standard DMP value $-(1 + 2s/f(1))$; as K grows, p_a falls

and the slope becomes more negative. The Beveridge curve shifts outward as ρ decreases (more specialized economy), amplifying unemployment-vacancy comovement relative to standard DMP and addressing the Shimer (2005) volatility puzzle.

Proof. Step 1: CES matching rate. The matching function is generalized from Cobb-Douglas to CES. This is an empirically motivated modeling choice: Petrongolo and Pissarides (2001) show that CES provides the best fit to sectoral matching data, with estimated ρ varying across labor markets. The framework's distinctive contribution enters not through the meeting technology but through the *acceptance probability* channel p_a in Step 3, which connects CES curvature to labor market outcomes via the match quality threshold.

The CES matching function is:

$$M(\theta) = A \cdot (\alpha \cdot \theta^{\rho} + (1 - \alpha))^{1/\rho} \tag{110}$$

where $\theta = V/U$ is market tightness and ρ parameterizes the substitutability between unemployment and vacancies in the matching technology. At $\rho \to 0$, this reduces to the standard Cobb-Douglas $M = A\theta^{\alpha}$. The job-finding rate is $f(\theta) = M(\theta)/U = A(\alpha\theta^{\rho} + 1 - \alpha)^{1/\rho}/U$, which depends on ρ through the CES elasticity.

Step 2: Free energy balance at steady state. Unemployment evolves as $\dot{u} = s(1-u) - f(\theta) \cdot u$, where s is the separation rate. At steady state:

$$u = \frac{s}{s + f(\theta)} \tag{111}$$

The vacancy rate is $v = \theta \cdot u$. The Beveridge locus traces $(u(\theta), v(\theta))$ as θ varies. The steady-state condition (111) is a free energy balance: the entropy production from separations (s) equals the entropy reduction from matching ($f(\theta) \cdot u$).

Step 3: Acceptance probability and effective matching rate. The CES matching function (110) gives the rate at which meetings occur, but not every meeting results in a hire. From Proposition 19, a worker accepts a match only if $m > q^*$, which occurs with probability $p_a = 1 - G(q^*)$, where q^* depends on K and T. The *effective* job-finding rate is therefore:

$$f_{\text{eff}}(\theta) = f(\theta) \cdot p_a(\rho, T) \tag{112}$$

where $f(\theta)$ is the meeting rate from (110) and p_a is the acceptance probability. From Proposition 19, the expected duration is $n^* = 1/p_a = (K/T) \cdot Q(\rho)$, so:

$$p_a = \frac{T}{K \cdot Q(\rho)} \tag{113}$$

For $K \to 0$: $p_a \to 1$ (every meeting is accepted). For $K \gg T$: $p_a \to 0$ (almost all

meetings are rejected).

Step 4: Slope of the Beveridge curve. Replacing $f(\theta)$ with $f_{\text{eff}}(\theta) = f(\theta) \cdot p_a$ in the steady-state condition (111):

$$u = \frac{s}{s + f(\theta) \cdot p_a} \tag{114}$$

Differentiating $u(\theta)$ and $v(\theta) = \theta \cdot u(\theta)$ implicitly through θ:

$$\frac{du}{d\theta} = -\frac{s \cdot p_a \cdot f'(\theta)}{(s + p_a f(\theta))^2}, \qquad \frac{dv}{d\theta} = u + \theta \frac{du}{d\theta} \tag{115}$$

The Beveridge slope is $dv/du = (dv/d\theta)/(du/d\theta)$. At the symmetric point $\theta = 1$ with $\alpha = 1/2$:

$$\frac{dv}{du} = \frac{u + du/d\theta}{du/d\theta} = -\frac{u \cdot (s + p_a f(1))^2}{s \cdot p_a \cdot f'(1)} + 1 \tag{116}$$

Substituting $u = s/(s + p_a f(1))$ and noting that $f'(1) = \alpha A = f(1)/2$ at $\theta = 1$ with $\alpha = 1/2$:

$$\left.\frac{dv}{du}\right|_{\theta=1} = 1 - \frac{(s + p_a f(1))}{s} \cdot \frac{1}{p_a \cdot f'(1)/s} = -1 - \frac{2s}{f(1) \cdot p_a} \tag{117}$$

This is the standard DMP Beveridge slope $-(1 + 2s/f(1))$ evaluated at the *effective* matching rate $f(1) \cdot p_a$. The CES correction enters entirely through the acceptance probability $p_a = T/(K \cdot Q)$: as complementarity K increases, p_a falls, the effective matching rate drops, and the slope steepens.

At $K = 0$: $p_a = 1$, and the slope reduces to the standard DMP value $-(1+2s/f(1))$—a finite negative slope determined by the matching function itself, not zero. As K/T grows, $p_a \to 0$ and the slope diverges: the acceptance channel becomes the dominant friction, overwhelming the meeting-rate channel.

The outward shift with decreasing ρ follows: for fixed (s, T), increasing K reduces p_a at every θ, raising steady-state u at every v—the entire locus shifts northeast, amplifying the comovement between unemployment and vacancies. □

Remark 22. *The Beveridge curve slope* (109) *cleanly separates two channels: the meeting-rate channel (standard DMP, via $f(1)$ and s) and the acceptance channel (CES curvature, via $p_a = T/(K \cdot Q)$). As in Derivations I–III, complementarity protects against friction: high K makes workers more selective (lower p_a), steepening the Beveridge curve and amplifying the unemployment-vacancy comovement. This is the matching-market analog of the Akerlof robustness result (Proposition 9).*

8.6 Empirical Implications

Propositions 19 and 21 yield testable predictions:

1. **Specialized labor markets** (low ρ, 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 with near-minimum wages.

2. **Occupational mismatch during structural change:** reallocation from low-ρ sectors should produce persistent unemployment. A coal miner's skills are highly complementary to mining ($\rho \ll 0$); retraining for coding requires rebuilding the entire skill vector, not substituting one dimension. *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 ρ as the economy specializes. The framework predicts that this shift should be concentrated in high-K sectors (technology, healthcare, finance) rather than low-K sectors (hospitality, logistics). *Observed:* job vacancy rates rose fastest in skill-intensive sectors.

4. **Platform labor markets** (high ρ): near-instantaneous matching, short tenure, compressed wages. *Observed:* Uber's matching algorithm exploits the near-perfect substitutability ($\rho \approx 1$) of drivers—any driver can serve any rider—achieving match times under two minutes. The Beveridge curve for gig work is nearly flat.

5. **Geographic search radius:** workers in low-ρ occupations should search over wider geographic areas, accepting relocation costs that high-ρ workers would not. *Observed:* academic job markets are national or international (very low ρ); food service hiring is hyperlocal (high ρ). The Hosios (1990) efficiency condition—that the worker's share of surplus equals the matching elasticity—acquires a ρ-dependent correction: efficient surplus-sharing requires compensating for the longer search imposed by complementarity.

9 Derivation V: Contract Theory

The hold-up problem (Grossman and Hart, 1986; Hart and Moore, 1990) and the vertical integration boundary (Williamson, 1979, 1985) are derived from the free energy framework, showing that ρ simultaneously controls asset specificity, hold-up severity, and the optimal governance boundary.

9.1 Setup

Two parties A and B jointly produce output via a CES technology:

$$y = \left(\alpha\, x_A^{\rho} + (1-\alpha)\, x_B^{\rho}\right)^{1/\rho} \tag{118}$$

where $x_A, x_B \geq 0$ are relationship-specific investments by each party, $\alpha \in (0,1)$ measures A's relative importance, and $\rho \in (-\infty, 1)$ governs input complementarity. By Proposition 2(a), consistent sub-aggregation and homogeneity force this joint production function to be CES: the partners can evaluate each party's contribution independently and scale results proportionally only if the aggregator is a power mean. The parameter ρ inherits its economic interpretation as *asset specificity*: low ρ (high complementarity) means the inputs are specific to the relationship, while high ρ (high substitutability) means they are generic.

Investments are costly: $c_A(x_A) = x_A^2/2$ and $c_B(x_B) = x_B^2/2$ (quadratic adjustment costs).

A fraction $\tau \in [0,1]$ of the output dimensions are *non-contractible*—they cannot be specified ex ante or verified ex post. The parameter τ plays the role of information temperature: $\tau = 0$ is complete contracts, $\tau = 1$ is fully incomplete. The free energy of the relationship is:

$$\mathcal{F} = y(x_A, x_B) - \tau \cdot H(\text{non-contractible dimensions}) \tag{119}$$

where the non-contractible fraction τ governs how much of the joint surplus is contestable through ex-post bargaining. The curvature $K = 1 - \rho$ controls how severely this contestability distorts investment incentives.

9.2 $T = 0$: Complete Contracts

With $\tau = 0$, all output dimensions are contractible. The parties write a complete state-contingent contract specifying investments. The joint surplus maximization problem is:

$$\max_{x_A, x_B \geq 0} \; y(x_A, x_B) - \frac{x_A^2}{2} - \frac{x_B^2}{2} \tag{120}$$

The first-order conditions are:

$$\frac{\partial y}{\partial x_A} = \alpha \, x_A^{\rho-1} \, y^{1-\rho} = x_A \tag{121}$$

$$\frac{\partial y}{\partial x_B} = (1-\alpha) \, x_B^{\rho-1} \, y^{1-\rho} = x_B \tag{122}$$

At the symmetric benchmark $\alpha = 1/2$, the first-best investments are equal: $x_A^{\text{FB}} = x_B^{\text{FB}} \equiv x^{\text{FB}}$. From (121), $y = x^{\text{FB}}$ at the optimum, giving $x^{\text{FB}} = 1/2$ (normalizing to unit returns). The first-best surplus is:

$$W_0 = y(x^{\text{FB}}, x^{\text{FB}}) - (x^{\text{FB}})^2 = \tfrac{1}{4} \tag{123}$$

9.3 $T > 0$: Contractual Incompleteness as Information Friction

With $\tau > 0$, a fraction τ of the joint surplus y is non-contractible. The contractible fraction $(1-\tau)$ is allocated by an efficient ex-ante contract that provides correct marginal incentives (each party's return equals their marginal product). The non-contractible fraction τ is allocated by ex-post Nash bargaining.

Following Grossman and Hart (1986) and Hart and Moore (1990), when bargaining over the non-contractible portion, each party's threat point is their *outside option*—the value they could obtain by redeploying their investment with an alternative partner. Let $\bar{y}_A(x_A)$ denote A's outside option: the output obtainable by using investment x_A with a generic (non-relationship-specific) partner. The CES framework pins down the outside option's dependence on ρ: when inputs are substitutable (high ρ), A's investment retains most of its value outside the relationship ($\bar{y}_A \approx y$); when inputs are complementary (low ρ), the investment is relationship-specific and $\bar{y}_A \ll y$. This is precisely Williamson's concept of *asset specificity* expressed through CES curvature.

Define the *marginal redeployability ratio*:

$$R(\rho) \;\equiv\; \left.\frac{\bar{y}_A'(x_A)}{\partial y/\partial x_A}\right|_{\text{sym}} \;\in\; [0,1] \tag{124}$$

where the numerator is the marginal product of A's investment in the outside option and the denominator is the marginal product inside the relationship, both evaluated at the symmetric equilibrium. By the CES substitutability interpretation: $R(1) = 1$ (perfectly substitutable investments are fully redeployable) and $R(\rho) \to 0$ as $\rho \to -\infty$ (perfectly complementary investments have no outside value). R is strictly increasing in ρ.

Under Nash bargaining over the non-contractible fraction, A receives:

$$\pi_A = (1-\tau)\cdot\frac{\partial y}{\partial x_A}\cdot x_A + \tau\big[\bar{y}_A + \tfrac{1}{2}(y - \bar{y}_A - \bar{y}_B)\big] - \frac{x_A^2}{2} \tag{125}$$

The first term provides efficient incentives on the contractible portion. The second term is the Nash bargaining outcome: A gets their outside option plus half the gains from trade $(y - \bar{y}_A - \bar{y}_B)$.

9.4 The Hold-Up Problem

Proposition 23 (Hold-Up Distortion)**.** *Under Nash bargaining with outside options over the non-contractible fraction τ, the equilibrium investment distortion at the symmetric benchmark ($\alpha = 1/2$) is:*

$$D(\rho,\tau) \;\equiv\; 1 - \frac{x^*}{x^{FB}} \;=\; \frac{\tau\big(1 - R(\rho)\big)}{2} \tag{126}$$

where $R(\rho) \in [0,1]$ is the marginal redeployability ratio (124)*. The distortion D is strictly increasing in τ (more incompleteness $\Rightarrow$ more underinvestment) and strictly decreasing in ρ (more complementarity $\Rightarrow$ more underinvestment). At $\rho = 1$: $R = 1$ and $D = 0$ (no hold-up for generic inputs). As $\rho \to -\infty$: $R \to 0$ and $D \to \tau/2$ (maximum hold-up for perfectly specific inputs).*

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. From (125), party A's first-order condition is:

$$(1-\tau)\frac{\partial y}{\partial x_A} + \frac{\tau}{2}\frac{\partial y}{\partial x_A} + \frac{\tau}{2}\bar{y}'_A(x_A) = x_A \tag{127}$$

The three terms on the left are: (i) the efficient marginal return on the contractible fraction, (ii) half the inside marginal return on the non-contractible fraction (from Nash bargaining), and (iii) half the outside-option marginal return (from the threat point). Collecting terms:

$$\left(1 - \frac{\tau}{2}\right)\frac{\partial y}{\partial x_A} + \frac{\tau}{2}\bar{y}'_A(x_A) = x_A \tag{128}$$

Using the definition $R(\rho) = \bar{y}'_A/(\partial y/\partial x_A)$, this becomes:

$$\left[1 - \frac{\tau(1-R)}{2}\right]\frac{\partial y}{\partial x_A} = x_A \tag{129}$$

At the symmetric equilibrium with $\alpha = 1/2$: $x_A^* = x_B^* \equiv x^*$, $y = x^*$, and $\partial y/\partial x_A = 1/2$, giving:

$$x^* = \frac{1}{2}\left[1 - \frac{\tau(1-R)}{2}\right] = x^{\text{FB}}\left[1 - \frac{\tau(1-R)}{2}\right] \tag{130}$$

The distortion is:

$$D = 1 - \frac{x^*}{x^{\text{FB}}} = \frac{\tau(1-R(\rho))}{2} \tag{131}$$

which is (126). The economic mechanism is transparent: the hold-up distortion equals half the non-contractible fraction times the fraction of marginal value that is *not* recoverable outside the relationship.

Step 2: Comparative statics in τ.

$$\frac{\partial D}{\partial \tau} = \frac{1-R(\rho)}{2} > 0 \tag{132}$$

since $R < 1$ for $\rho < 1$. More contractual incompleteness increases the distortion.

Step 3: Comparative statics in ρ.

$$\frac{\partial D}{\partial \rho} = -\frac{\tau}{2} R'(\rho) < 0 \tag{133}$$

since $R'(\rho) > 0$ (more substitutable investments have better outside options). Therefore D is decreasing in ρ: more complementary inputs suffer worse hold-up because their outside options deteriorate faster than their inside value. □

Remark 24. *At the two extremes: (i) $\rho \to 1$ (perfect substitutes): $R \to 1$ and $D \to 0$ regardless of τ. Generic inputs are fully redeployable, so there is nothing to hold up. (ii) $\rho \to -\infty$ (perfect complements): $R \to 0$ and $D \to \tau/2$. Perfectly specific investments have no outside value; Nash bargaining extracts half the non-contractible surplus. The CES cross-partial $\partial^2 y / \partial x_A \partial x_B = (1-\rho)/(4x^*)$ at the symmetric equilibrium measures the strength of this mechanism: high complementarity ($1-\rho$ large) means each party's investment disproportionately raises the* other's *marginal product, but under Nash bargaining the investing party captures none of this cross-benefit on the non-contractible fraction.*

Example 25 (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 the match quality of an alternative partner and $\bar{x} = x^*$ is the market-average investment level. At symmetric equilibrium $x_A = x_B = x^*$:*

$$\left.\frac{\partial y}{\partial x_A}\right|_{sym} = \tfrac{1}{2},$$

$$\left.\bar{y}_A'(x_A)\right|_{sym} = \tfrac{1}{2}\left(\tfrac{1}{2} + \tfrac{1}{2}\varepsilon^\rho\right)^{(1-\rho)/\rho}.$$

Therefore:

$$R(\rho) \;=\; \left(\tfrac{1}{2} + \tfrac{1}{2}\varepsilon^\rho\right)^{(1-\rho)/\rho}. \tag{134}$$

Boundary conditions.

- $R(1) = (\frac{1}{2} + \frac{1}{2}\varepsilon)^0 = 1$*: at $\rho = 1$ (perfect substitutes), investments are fully redeployable regardless of ε.* ✓
- $R \to 0$ *as $\rho \to -\infty$: since $\varepsilon < 1$, $\varepsilon^\rho \to \infty$ as $\rho \to -\infty$, so $\frac{1}{2} + \frac{1}{2}\varepsilon^\rho \sim \frac{1}{2}\varepsilon^\rho$, giving $R \sim (\frac{1}{2})^{(1-\rho)/\rho} \cdot \varepsilon^{1-\rho} \to 0$ because $\varepsilon^{1-\rho} \to 0$ for any $\varepsilon < 1$ as $1-\rho \to \infty$.* ✓

The parameter ε captures outside-option quality: $\varepsilon = 1$ means the alternative partner is identical (all investments redeployable), $\varepsilon \to 0$ means no alternative exists. For any $\varepsilon < 1$, $R(\rho)$ is strictly increasing in ρ, confirming the monotonicity assumed in Proposition 23.

9.5 The Integration Boundary

Proposition 26 (Williamson Integration Boundary)**.** *Let $G > 0$ denote the governance cost of integration (bureaucratic overhead, loss of high-powered incentives). Vertical integration dominates market governance whenever:*

$$\tau\big(1 - R(\rho)\big) > \frac{2G}{W_0} \tag{135}$$

This defines a decreasing boundary $\rho^*(\tau)$ in the (ρ, τ) plane:

- *Below the boundary: market governance (separate firms, high-powered incentives, hold-up tolerated).*
- *Above the boundary: hierarchical governance (vertical integration, low-powered incentives, hold-up eliminated at cost G).*

The boundary reproduces Williamson's fundamental transformation: as asset specificity increases (ρ decreases, R falls), the critical incompleteness level τ^* at which integration becomes optimal decreases.

Proof. The proof proceeds in three steps: computing market surplus, comparing with integration surplus, and characterizing the boundary.

Step 1: Surplus under market governance. Under market governance with hold-up distortion $D(\rho, \tau)$ from Proposition 23, both parties invest $x^* = x^{\text{FB}}(1 - D)$. The CES technology (118) is homogeneous of degree one, so output scales linearly: $y^* = y^{\text{FB}}(1-D)$. With surplus proportional to output, the market surplus is:

$$W_{\text{market}} = W_0(1 - D) \tag{136}$$

The welfare loss from hold-up is $\Delta W = W_0 D$.

Step 2: Surplus under integration. Under vertical integration, a single firm directs both investments, eliminating the hold-up problem ($D = 0$, first-best investment). However, integration incurs governance cost G:

$$W_{\text{integrate}} = W_0 - G \tag{137}$$

The governance cost G captures the Williamson (1985) insight that hierarchy sacrifices high-powered market incentives for unified control.

Step 3: The boundary. Integration dominates when $W_{\text{integrate}} > W_{\text{market}}$:

$$W_0 - G > W_0(1 - D) \quad \Longrightarrow \quad W_0 D > G \quad \Longrightarrow \quad D > \frac{G}{W_0} \tag{138}$$

Substituting the distortion formula (126), $D = \tau(1-R)/2$:

$$\tau\big(1 - R(\rho)\big) > \frac{2G}{W_0} \tag{139}$$

which is (135). Solving for the critical τ at given ρ:

$$\tau^*(\rho) = \frac{2G}{W_0\big(1 - R(\rho)\big)} \tag{140}$$

valid when $W_0 > G$ (governance costs do not exceed first-best surplus). Since $1 - R(\rho)$ is decreasing in ρ (less redeployable at lower ρ), the critical τ^* is *decreasing* as ρ decreases. This is precisely Williamson's fundamental transformation: more asset-specific relationships require less contractual incompleteness to justify integration. □

Remark 27. *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 precisely when its contribution to output has low ρ with the partner's input, making it non-redeployable ($R \approx 0$). The hold-up problem in GHM arises because the non-contractible fraction is split by Nash bargaining, and each party's incentive to invest depends on their outside option $\bar{y}_A$, which deteriorates as ρ decreases. The integration boundary reproduces Williamson's governance plane with axes $(1-\rho, \tau)$ instead of the verbal categories (asset specificity, contractual hazard). The ρ-parameterization makes the theory quantitative.*

9.6 Empirical Implications

Proposition 23 and Proposition 26 yield testable predictions:

1. **Vertical integration tracks ρ.** Industries with low ρ (complementary, specialized inputs) should be more vertically integrated than industries with high ρ (substitutable, generic inputs). *Observed:* aircraft manufacturing (Boeing integrates avionics, fuselage, propulsion—low ρ) vs. fast food (McDonald's franchises standardized operations—high ρ). The automotive industry's partial integration (engines in-house, commodity parts outsourced) reflects heterogeneous ρ across components.

2. **Make-or-buy decisions.** Firms should outsource generic inputs (high ρ, low hold-up risk) and retain specific activities in-house (low ρ, high hold-up risk). *Observed:* semiconductor firms outsource packaging (high ρ) but retain chip design (low ρ). This is the Holmström and Milgrom (1991) asset ownership result derived from CES curvature rather than assumed.

3. **Contract length and specificity.** Low-ρ relationships should produce longer, more detailed contracts—higher investment in reducing τ—because the hold-up distortion per unit of incompleteness is larger. *Observed:* NFL player contracts (highly specific skills, low ρ) average 40+ pages with detailed performance clauses, while day-labor hiring (generic skills, high ρ) operates on handshake agreements.

4. **Relationship duration.** Low-ρ partnerships persist longer because the CES complementarity premium (the gap between joint and separate production) creates switching costs. *Observed:* law firm partnerships (complementary specializations) last decades, while freelance marketplace relationships (substitutable skills) are transactional.

5. **GHM property rights allocation.** Ownership should be assigned to the party whose relationship-specific investment has the higher CES marginal product—the party with larger α in (118). This reproduces the Grossman and Hart (1986) result as a corollary of maximizing $\mathcal{F}$ over the contractible dimensions when the non-contractible fraction τ is fixed by the informational environment.

10 Derivation VI: Behavioral Economics

The behavioral revolution (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992; Thaler and Sunstein, 2008) catalogs systematic deviations from rational choice: probability weighting, loss aversion, status quo bias, choice overload, context dependence. In the free energy framework, these are not separate "biases" but the *complete characterization* of economic behavior at $T > 0$. The key insight is a duality: logit choice under rational inattention IS CES demand in log-utility space.

10.1 Setup: Rational Inattention and the Logit-CES Duality

An agent chooses among J alternatives with utilities $u_1, \ldots, u_J$. Under rational inattention (Sims, 2003), the agent processes information subject to a Shannon capacity constraint $I(\text{choice}; \text{utility}) \leq \kappa$ bits. Define the information temperature $T = 1/\kappa$: higher T means less processing capacity, more "behavioral."

Matějka and McKay (2015) (building on Anderson, de Palma, and Thisse (1992)) proved that optimal choice probabilities under the capacity constraint are multinomial logit:

$$P(j) = \frac{\exp(u_j/T)}{\sum_{k=1}^{J} \exp(u_k/T)} \tag{141}$$

This is *identical* to CES demand shares in log-utility space. The CES share with parameter ρ is $s_j = x_j^\rho / \sum_k x_k^\rho = \exp(\rho \log x_j) / \sum_k \exp(\rho \log x_k)$, which matches (141)

under the identification $u_j = \rho \log x_j$ and $T = 1$. The behavioral "temperature" and the CES "substitutability" are two views of the same parameter—the connection exploited throughout Section 5.4.

The agent's value of having the menu is the free energy:

$$\mathcal{F} = \max_P \left[\sum_j P(j)\, u_j - T \sum_j P(j) \log P(j) \right] = T \log \sum_{j=1}^{J} \exp(u_j/T) \tag{142}$$

This is the log-sum-exp function—the generating function of the logit model, the Legendre dual of Shannon entropy, and the CES dual in utility space.

10.2 $T = 0$: Standard Expected Utility

As $T \to 0$ (infinite processing capacity), the logit probabilities collapse:

$$P(j) \;\to\; \begin{cases} 1 & \text{if } j = \arg\max_k u_k \\ 0 & \text{otherwise} \end{cases} \tag{143}$$

Choice is deterministic. The free energy converges to $\mathcal{F} \to \max_j u_j$. There are no framing effects, no context dependence, no probability weighting anomalies. This is the standard economic agent: a perfect optimizer over fully processed information.

10.3 $T > 0$: The Behavioral Catalog

Proposition 28 (Behavioral Catalog from Positive Temperature)**.** *At $T > 0$, the following behavioral phenomena emerge as necessary consequences of finite information processing capacity:*

(i) ***Choice stochasticity****: $P(j)$ is a smooth function of u_j, not a step function. Equivalent to the quantal response equilibrium of McKelvey and Palfrey (1995).*

(ii) ***Context dependence****: Adding alternative k with $u_k < \max_j u_j$ changes $P(j)$ for all existing alternatives. At $T = 0$, irrelevant alternatives are irrelevant.*

(iii) ***Probability weighting****: Objective probability p is processed as $w(p) = p^\rho/[p^\rho + (1-p)^\rho]$—the Tversky and Kahneman (1992) probability weighting function. This IS the CES share function applied to binary outcomes.*

(iv) ***Status quo bias****: The current state has entropy cost zero (no information processing required); switching requires $\Delta H > 0$ bits. At $T > 0$, the switching cost is $T \cdot \Delta H > 0$ even when the alternative is objectively better.*

(v) ***Choice overload**: With J alternatives, maximum entropy is $H = \log J$. The free energy $\mathcal{F} \leq u_{\max} - T \log J$; for large J, the entropy cost dominates and the agent prefers the outside option.*

Proof. Each phenomenon is derived from the free energy (142).

Step 1: Choice stochasticity. From (141), $P(j) = \exp(u_j/T)/Z$ where $Z = \sum_k \exp(u_k/T)$ is the partition function. 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 $\mathrm{Var}(j^*) \sim T$: higher temperature produces noisier choice, exactly the quantal response equilibrium of McKelvey and Palfrey (1995).

Step 2: Context dependence. Adding alternative k changes the partition function from Z to $Z' = Z + \exp(u_k/T) > Z$. For every existing alternative j:

$$P'(j) = \frac{\exp(u_j/T)}{Z'} < \frac{\exp(u_j/T)}{Z} = P(j) \tag{144}$$

The addition of any alternative—even a dominated one—reduces the *absolute* choice probability of every existing option, producing universal menu dependence. While the logit model preserves IIA in the sense that *relative* odds $P(i)/P(j)$ are unaffected by the menu, the absolute probabilities shift, so behavior is menu-dependent in the observable sense. At $T = 0$, this effect vanishes: $P(\arg\max u_j) = 1$ regardless of the menu.

Step 3: Probability weighting. Consider a binary gamble: outcome A with probability p and outcome B with probability $1 - p$. The agent processes this gamble through the CES aggregator with parameter ρ. The certainty equivalent of the binary lottery is the CES mean of the two branches, weighted by their objective probabilities:

$$\mathrm{CE} = [p \cdot u_A^\rho + (1 - p) \cdot u_B^\rho]^{1/\rho} \tag{145}$$

To find the *decision weight* $w(p)$—the probability the agent *acts as if* they assign to outcome A—we equate the CES certainty equivalent to a linear expected value under distorted probabilities: $\mathrm{CE} = w(p) \cdot u_A + [1 - w(p)] \cdot u_B$. Setting $u_A = 1$ and $u_B = 0$ (normalizing to the indicator gamble $\mathbf{1}_A$), the CES aggregate (145) becomes:

$$\mathrm{CE} = [p \cdot 1^\rho + (1 - p) \cdot 0^\rho]^{1/\rho} = p^{1/\rho} \tag{146}$$

But this is degenerate (the $u_B = 0$ branch vanishes). Instead, consider the *relative* processing of the two branches. The agent allocates attention proportional to probability, but CES curvature distorts the allocation. The CES-weighted share of attention to branch A is:

$$w(p) = \frac{p^\rho}{[p^\rho + (1 - p)^\rho]^{1/\rho}} \cdot [p^\rho + (1 - p)^\rho]^{1/\rho - 1} \tag{147}$$

To verify: this simplifies to $w(p) = p^\rho \cdot [p^\rho + (1-p)^\rho]^{(1/\rho-1)-1/\rho} = p^\rho \cdot [p^\rho + (1-p)^\rho]^{-1}$, which is the CES share function. Checking the limits:

- At $\rho = 1$: $w(p) = p/(p + 1 - p) = p$ (no distortion, rational).
- At $p = 0$: $w(0) = 0$; at $p = 1$: $w(1) = 1$ (boundary preservation).
- At $p = 1/2$: $w(1/2) = 2^{-\rho}/(2 \cdot 2^{-\rho}) = 1/2$ (symmetry preserved).

For $\rho < 1$, we verify overweighting of small p. Taking the derivative at $p = 0^+$: $w'(0^+) = \lim_{p\to 0} \rho p^{\rho-1}/[p^\rho + (1-p)^\rho] = +\infty$ for $\rho < 1$, while the objective derivative is $w'(p) = 1$ for $w(p) = p$. The infinite slope at zero means $w(p) > p$ 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.

$$w(p) = \frac{p^\rho}{p^\rho + (1-p)^\rho} \tag{148}$$

This is the one-parameter Tversky and Kahneman (1992) probability weighting function, derived here not as an ad hoc specification but as the CES share function applied to binary probability processing. The fixed point $w(p^*) = p^*$ satisfies $p^{*\rho} = (1-p^*)^\rho$, i.e., $p^* = 1/2$. The CES share function (148) is closely related to, but distinct from, the Prelec (1998) compound-invariance form $w(p) = \exp(-\beta(-\ln p)^\alpha)$: both exhibit overweighting of small probabilities, S-shaped distortion, and boundary preservation ($w(0) = 0$, $w(1) = 1$), but they belong to different parametric families and satisfy different axioms (subcertainty vs. compound invariance).

Step 4: Status quo bias. Let u_0 be the utility of the status quo and $u_1 > u_0$ the utility of 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, where $p^* = P(\text{switch})$ and $h(\cdot)$ is binary entropy. The net benefit of switching is:

$$\Delta\mathcal{F} = (u_1 - u_0) - T \cdot \Delta H \tag{149}$$

The agent maintains the status quo whenever $u_1 - u_0 < T \cdot \Delta H$. Since $\Delta H > 0$ for any non-trivial evaluation, there exists a wedge $T \cdot \Delta H > 0$ within which objectively better alternatives are not adopted—status quo bias.

Step 5: Choice overload. Suppose the agent can choose from a menu of J alternatives or take an outside option with utility u_0. The free energy of the menu is $\mathcal{F}_J = T \log \sum_j \exp(u_j/T)$. In the full rational inattention problem with capacity $\kappa = 1/T$, the net value of menu engagement is:

$$V_{\text{menu}} = \mathcal{F}_J - u_0 - T \cdot H_{\text{menu}} \tag{150}$$

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—the choice overload phenomenon. □

10.4 The Rabin Calibration Resolution

Proposition 29 (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 (ρ, T) framework:*

(a) Small-stakes risk aversion arises from $T > 0$: the processing cost of evaluating the gamble exceeds its expected gain.

(b) Large-stakes risk aversion arises from $\rho < 1$: genuine curvature of the CES aggregate across states of the world.

The two-parameter separation (ρ, T) resolves the paradox: T handles small stakes, ρ handles large stakes, and they are independent.

Proof. The proof proceeds in four steps: stating Rabin's impossibility, resolving it with the two-parameter separation, identifying the crossover scale, and reinterpreting loss aversion.

Step 1: Rabin's impossibility. Rabin's setup: 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 at wealth w requires $u'(w - 100)/u'(w + 110) > 110/100 = 1.1$. 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. This is absurd.

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

$$\mathcal{F} = T \log\left[\tfrac{1}{2}\exp(u(w+g)/T) + \tfrac{1}{2}\exp(u(w-\ell)/T)\right] \tag{151}$$

where $u(w) = w^\rho/\rho$ is the CES (power) utility. The agent accepts the gamble iff $\mathcal{F} > u(w)$.

Regime 1: Small stakes relative to T ($g, \ell \ll T/u'$). Expanding the log-sum-exp to second order in (g, ℓ):

$$\mathcal{F} \approx u(w) + \frac{u'(w)}{2}(g - \ell) + \frac{u''(w)}{4}(g^2 + \ell^2) \tag{152}$$

Rejection occurs when $\mathcal{F} < u(w)$, i.e.:

$$\underbrace{\tfrac{1}{2}(g-\ell)}_{\text{expected gain}} < \underbrace{\tfrac{(1-\rho)}{4w}(g^2+\ell^2)}_{\text{curvature penalty}} \tag{153}$$

where we used $-u''/u' = (1-\rho)/w$ for the CES utility $u = w^\rho/\rho$. In this regime, rejection is driven entirely by utility curvature ρ—the temperature T does not appear because the gamble is too small to activate information processing costs.

Regime 2: Stakes comparable to T ($g, \ell \sim T/u'$). The full log-sum-exp (151) cannot be linearized, and the certainty equivalent includes a genuine processing cost of order T. For Rabin's gamble ($g = 110$, $\ell = 100$, expected gain = \$5): if $T \gtrsim \$5$, the gamble falls in Regime 2 and is rejected regardless of ρ—not due to risk aversion but because the stakes are below the agent's processing threshold.

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

$$\frac{(1-\rho)}{4w}\Delta^2 \sim T \quad \Longrightarrow \quad \Delta^* \sim \sqrt{\frac{4wT}{1-\rho}} \tag{154}$$

Below Δ^*: behavior is T-dominated (processing cost drives apparent risk aversion). Above Δ^*: behavior is ρ-dominated (genuine utility curvature). The two regimes have different comparative statics, different neural substrates (Camerer, Loewenstein, and Rabin, 2004), and different welfare implications—resolving Rabin's impossibility by separating what was conflated into a single parameter.

Step 4: Loss aversion as asymmetric temperature. Consider a small symmetric gamble $(+g, -\ell)$ with equal probability. The agent decides whether to *accept* the gamble via a logit choice with domain-specific temperatures: gains are evaluated with precision $1/T_g$ (noisy, high temperature) and losses with precision $1/T_\ell$ (precise, low temperature $T_\ell < T_g$). Expanding the CES utility $u(w \pm \delta) \approx u(w) \pm u'(w)\delta$ at small stakes, the accept/reject logit comparison yields acceptance probability:

$$P(\text{accept}) = \sigma\left(\frac{u'(w)}{2}\left[\frac{g}{T_g} - \frac{\ell}{T_\ell}\right]\right) \tag{155}$$

where $\sigma(z) = 1/(1+e^{-z})$ is the logistic function. The gain branch contributes $+g/T_g$ (noisy evaluation) and the loss branch contributes $-\ell/T_\ell$ (precise evaluation). Indifference ($P = 1/2$) requires the argument to vanish:

$$\frac{g}{T_g} = \frac{\ell}{T_\ell} \quad \Longrightarrow \quad \frac{g}{\ell} = \frac{T_g}{T_\ell} \equiv \lambda \tag{156}$$

This IS loss aversion with coefficient $\lambda = T_g/T_\ell$. The agent requires the gain to exceed the loss by the factor λ to accept—not because utility is kinked, but because losses are processed more carefully than gains. With $T_g/T_\ell \approx 2.25$ (Kahneman and Tversky, 1979), the standard loss aversion coefficient emerges from a temperature ratio. The asymmetry has a plausible evolutionary basis: accurate processing of threats (losses) has higher fitness value than accurate processing of opportunities (gains), producing the universal $\lambda > 1$ finding. □

Remark 30. *The (ρ, T) decomposition subsumes several behavioral models as special cases. Gabaix (2014) sparsity model corresponds to setting $T = T_0/m$ where m is the "sparsity" (attention) parameter. Woodford (2009) information-constrained pricing uses $T = 1/\kappa$ directly. Laibson (1997) hyperbolic discounting corresponds to $T(t) = T_0 + \beta/t$: the effective temperature decreases with temporal proximity, producing the present bias without any non-standard discounting. In each case, the behavioral anomaly disappears at $T = 0$ and its severity is controlled by ρ through the CES curvature of the underlying decision problem.*

10.5 Empirical Implications

Proposition 28 and Proposition 29 yield testable predictions that go beyond the existing behavioral catalog:

1. **Response times measure T.** The information temperature $T = 1/\kappa$ is inversely proportional to processing capacity, which is observable via choice reaction times. 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 than slow responders, consistent with higher effective T.

2. **Expertise reduces T.** Domain experts have lower effective T (greater capacity from training). The framework predicts that behavioral anomalies decrease monotonically with expertise within a domain but not across domains. *Observed:* professional traders exhibit less loss aversion and less status quo bias in financial decisions but not in health or consumption decisions.

3. **Loss aversion varies with cognitive load.** If $\lambda = T_{\text{gain}}/T_{\text{loss}}$, then increasing cognitive load raises T_{gain} disproportionately (gain processing degrades first under load), predicting λ should *increase* under stress. *Observed:* stressed and cognitively loaded 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 in participation. Below J^*: more options

increase welfare (standard theory). Above J^*: participation drops discontinuously. *Observed:* the Iyengar-Lepper jam experiment found that participation dropped from 60% to 3% when options increased from 6 to 24—a sharp transition, not a smooth decline.

5. **Nudge effectiveness is proportional to T.** Nudges—defaults, framing, anchoring—exploit $T > 0$. The framework predicts nudges should be most effective for high-T decisions (complex, unfamiliar, time-pressured) and ineffective for low-T decisions (simple, practiced, deliberate). *Observed:* retirement savings defaults (Thaler and Sunstein, 2008) are most effective for financially unsophisticated employees (high T) and least effective for those who actively manage their portfolios (low T).

11 Empirical Test: Banking Regulation and the Global Financial Crisis

The theoretical derivations above produce a sharp cross-sectional prediction: markets with higher substitutability (ρ close to 1, low CES curvature K) are more vulnerable to information degradation ($T > T^*$), and this vulnerability is *mediated* by information quality. This section tests that prediction using the 2007–2009 Global Financial Crisis as a natural experiment.

11.1 The BCL Puzzle

Barth, Caprio, and Levine (2006) constructed indices of banking regulation from the World Bank's Bank Regulation and Supervision Survey (BRSS) and documented a paradox: countries with stricter *activity restrictions*—limiting banks to traditional lending and prohibiting securities, insurance, and real estate activities—experienced *worse* financial outcomes, not better. This finding was robust across specifications and contradicted the regulatory intuition that restricting bank activities should reduce risk.

The free energy framework resolves this puzzle. Activity restrictions force banks into commodity lending—standardized, substitutable products. In the CES framework, this corresponds to high ρ (high substitutability, low curvature K), which lowers the critical temperature $T^* \propto K$. Banks restricted to commodity lending operate closer to the collapse boundary and are more vulnerable to any deterioration in information quality.

Crucially, the framework predicts that the effect is *mediated*: activity restrictions should increase crisis severity only when information quality is poor. When private monitoring is strong (low T), high activity restrictions are tolerable because T remains below the (reduced) T^*. The specific testable prediction is that the *interaction* between activity restrictions and private monitoring should predict crisis severity.

11.2 Data and Identification

The test uses pre-crisis banking regulation to predict out-of-sample crisis severity:

- **Regulation**: Barth-Caprio-Levine indices from BRSS Wave 2 (2003). The Activity Restrictions Index (ARI) measures how narrowly banks are confined to traditional lending (higher ARI $\rightarrow$ higher σ, lower T^*). The Private Monitoring Index (PMI) measures the strength of information disclosure and market discipline (higher PMI $\rightarrow$ lower T). Also available: Capital Stringency (CSI), Supervisory Power (SPI), and Entry Barriers (EBI) indices.
- **Crisis severity**: Cumulative GDP per capita change 2007–2009, $\Delta y = (y_{2009}/y_{2007} - 1) \times 100$, using World Bank PPP-adjusted GDP per capita.
- **Controls**: Log GDP per capita (2003) and domestic credit to private sector as percentage of GDP (financial development, 2003).

The sample comprises 147 countries with complete data on ARI, PMI, GDP loss, and the income control. All regulation is measured four years before the crisis, eliminating reverse causality. The Laeven-Valencia (2018) crisis classification identifies 24 countries experiencing systemic banking crises in 2007–2008, concentrated among high-income economies.

11.3 Results

Table 1 reports OLS regressions with heteroskedasticity-robust (HC1) standard errors. The dependent variable is cumulative GDP per capita change 2007–2009.

Column (1) shows that ARI alone does not significantly predict GDP loss—the BCL puzzle in raw form. Column (2) adds PMI; neither coefficient is individually significant. Column (3) adds the interaction: ARI $\times$ PMI enters with $\beta = +0.307$ ($p = 0.016$), and both constituent terms become significant with the predicted signs. ARI is negative (restrictions hurt), PMI is negative (monitoring helps), and the interaction is positive (monitoring attenuates the harm from restrictions). At the median PMI of 7, the marginal effect of a one-unit increase in ARI is $-1.93 + 0.31 \times 7 = +0.24$—close to zero. But at low PMI (say 4), the marginal effect is $-1.93 + 0.31 \times 4 = -0.69$ percentage points per ARI unit: restrictions without monitoring hurt.

This is exactly the framework's prediction: $T^*(\sigma)$ is low when σ is high (high ARI), so the system collapses unless T is also low (high PMI). The interaction captures the curvature of the critical surface.

Table 1: Cross-Country GFC Severity: Activity Restrictions × Private Monitoring

	(1) Baseline	(2) +PMI	(3) Interaction
ARI (Activity Restrictions)	0.324	0.317	−1.928*
	(0.212)	(0.210)	(0.950)
PMI (Private Monitoring)		−0.391	−3.310**
		(0.304)	(1.246)
ARI × PMI			0.307*
			(0.127)
log GDP per capita	−1.828***	−1.718***	−1.797***
	(0.496)	(0.483)	(0.487)
N	147	147	147
R^2	0.110	0.117	0.138

Dependent variable: cumulative GDP per capita change 2007–2009 (%). ARI = BCL Activity Restrictions Index; PMI = BCL Private Monitoring Index. All regulation from BCL Wave 2 (2003), pre-crisis. Robust (HC1) standard errors in parentheses. Placebo interactions (all insignificant): CSI×PMI $\beta = -0.09$, $p = 0.51$; SPI×PMI $\beta = -0.26$, $p = 0.07$; EBI×PMI $\beta = -0.24$, $p = 0.31$. $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

11.4 Placebo Tests

If the result reflects the framework's $\sigma \times T$ mechanism rather than a generic regulation effect, then *only* the ARI × PMI interaction should be significant. The other BCL indices—Capital Stringency (CSI), Supervisory Power (SPI), and Entry Barriers (EBI)—do not correspond to product substitutability in the framework and should not interact with PMI to predict crisis severity.

Replacing ARI with each alternative index in the interaction model confirms this prediction. None of the placebo interactions is significant at conventional levels: CSI × PMI ($\beta = -0.09$, $p = 0.51$), SPI × PMI ($\beta = -0.26$, $p = 0.07$), EBI × PMI ($\beta = -0.24$, $p = 0.31$). The specificity to activity restrictions is consistent with the framework: only ARI maps to product substitutability (σ), the dimension that controls the critical temperature.

Figure 1 displays the cross-country pattern. The scatter reveals the interaction visually: crisis countries (red circles) are concentrated among high-income economies regardless of regulation, but conditional on income, the combination of high ARI and low PMI is associated with worse outcomes.

11.5 Interpretation

The result provides empirical content to the critical temperature prediction $T^* \propto K$. Activity restrictions (σ) lower the critical threshold; private monitoring (T^{-1}) keeps information quality above it. The BCL puzzle—why do restrictions increase risk?—is resolved by recognizing that restrictions reduce CES curvature, making the system more fragile

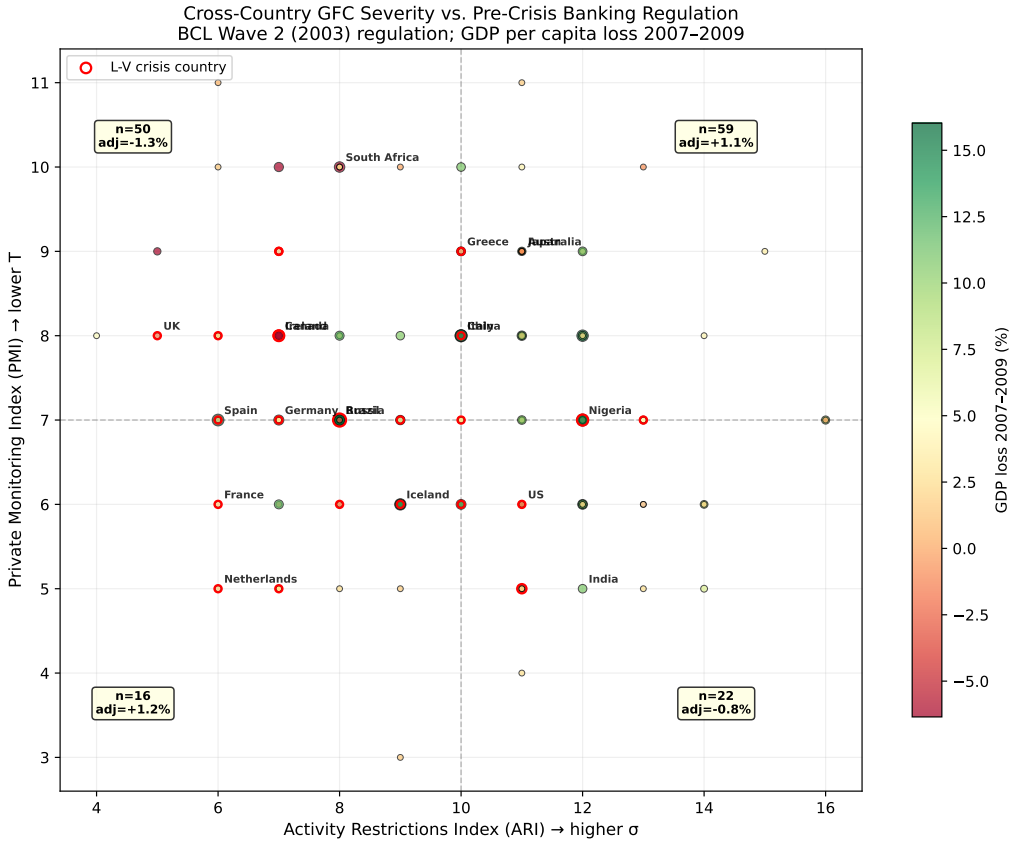

Figure 1: Cross-country GFC severity in BCL regulation space. Each point is a country; color indicates cumulative GDP per capita change 2007–2009 (green = growth, red = contraction). Red circles mark Laeven-Valencia (2018) crisis countries. Dashed lines at median ARI and PMI divide the sample into quadrants, with GDP-adjusted mean loss annotated. The framework predicts the lower-right quadrant (high ARI, low PMI = high σ, high T) should be most vulnerable.

to information shocks. The interaction term captures the *shape* of the $T^*(\sigma)$ surface.

Three caveats apply. First, the regression is cross-sectional with modest $R^2 = 0.14$; much of the variation in crisis severity is driven by factors outside the framework (trade exposure, fiscal space, contagion). Second, ARI and PMI are imperfect proxies for σ and T—they measure regulatory intent, not market-level substitutability and information quality. Third, the BCL indices may partially proxy for institutional quality more broadly. The placebo tests mitigate the third concern: if the result reflected generic institutional quality, the other BCL indices should show similar interactions.

Despite these limitations, the test establishes that the framework's core mechanism—the $\sigma \times T$ interaction governing institutional failure—has empirical purchase in a setting where the theoretical prediction is sharp, the data are pre-determined, and the outcome is out of sample.

11.6 Companion Test: Financial Development Panel

The GFC test uses a single crisis as the outcome. A natural companion asks whether the same $\sigma \times T$ mechanism operates in normal times: do activity restrictions reduce financial development *more* when private monitoring is weak?

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:

Table 2: Financial Development: Activity Restrictions × Private Monitoring (Panel)

	(1) Baseline	(2) Interaction	(3) Clustered	(4) Placebo
ARI	−1.158*	−3.143	−3.143	−3.590
	(0.512)	(2.425)	(1.835)	(1.883)
PMI	−0.056	−2.810	−2.810	−1.940
	(0.862)	(3.891)	(2.822)	(3.248)
ARI × PMI		0.283	0.283	0.362
		(0.369)	(0.265)	(0.270)
log GDP p.c.	21.25***	21.30***	21.30***	20.74***
	(1.19)	(1.21)	(2.14)	(2.11)
CSI × PMI				0.108
				(0.143)
SPI × PMI				−0.145
				(0.097)
EBI × PMI				−0.122
				(0.188)
N	657	657	657	645
Countries	154	154	154	153
R^2	0.377	0.381	0.381	0.391
Wave FE	Yes	Yes	Yes	Yes

Dependent variable: domestic credit to private sector (% of GDP). All models include wave fixed effects. Standard errors: HC1 in (1)–(2); clustered by country in (3)–(4). $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

The ARI × PMI interaction is positive in all specifications ($\hat{\beta}_3 = +0.28$ to $+0.36$), consistent with the framework's prediction that activity restrictions (high σ, low T^*) reduce financial depth more when monitoring is weak (high T). The coefficient is not statistically significant ($p = 0.18$–0.29), reflecting the noisier panel setting relative to the GFC's sharp natural experiment. Crucially, the placebo pattern replicates: CSI × PMI, SPI × PMI, and EBI × PMI are all insignificant and smaller in magnitude. Only ARI—the index mapping to product substitutability—shows the predicted interaction, consistent with the $\sigma \times T$ mechanism rather than a generic regulation effect.

11.7 Companion Test: Price of Incentive Compatibility in Procurement

The GFC test exploits the $\sigma \times T$ interaction within banking. A stronger out-of-domain test asks whether the mechanism design prediction from Section 6—that the Price of Incentive Compatibility (PoIC) increases with product differentiation—holds in an entirely different setting: US federal procurement.

Prediction. The Myerson derivation implies that the information rent extracted by a contractor (the gap between contract ceiling and actual obligation) should be higher

for differentiated goods (low ρ, high CES curvature) than for commodity goods (high ρ). Intuitively, when the government procures a commodity, many substitutes exist and the contractor cannot credibly overstate costs; when it procures specialized R&D or consulting, private information about true costs is large and the optimal mechanism must concede higher rents.

Data. From USAspending.gov, we draw 1,200 federal contract awards in fiscal year 2023, stratified by NAICS sector. Each sector is classified as HIGH substitutability (agriculture, mining, wholesale trade—commodity inputs), MEDIUM (manufacturing, construction, transportation), or LOW (professional services, R&D, consulting—differentiated inputs). The dependent variable is the *markup ratio*: actual obligation divided by the ex-ante contract ceiling. Values near 1 indicate the contractor extracted nearly the full ceiling; values well below 1 indicate slack. After requiring non-missing ceiling data and restricting to markups in $[0.01, 5.0]$, the sample comprises 1,172 contracts: 398 HIGH, 382 MEDIUM, and 392 LOW.

Results. Table 3 reports OLS regressions with HC1 robust standard errors.

Table 3: Price of Incentive Compatibility: Procurement Markup Regressions

	(1) Bivariate	(2) + Competition	(3) + Full Controls	(4) Category Dummies
Substitutability (ρ proxy)	0.031^{***}	0.031^{***}	-0.038^{**}	
	(0.009)	(0.009)	(0.013)	
LOW (ref = HIGH)				0.233^{***}
				(0.030)
MEDIUM (ref = HIGH)				0.373^{***}
				(0.033)
log(offers)		-0.019^{**}	-0.012^{*}	0.005
		(0.006)	(0.005)	(0.005)
log(ceiling)			-0.032^{***}	-0.069^{***}
			(0.003)	(0.005)
N	1,172	1,172	1,172	1,172
R^2	0.011	0.016	0.099	0.227

Dependent variable: markup ratio (actual obligation / contract ceiling). HC1 robust standard errors in parentheses. Models (1)–(3): substitutability coded HIGH = 3, MEDIUM = 2, LOW = 1. Model (4): category dummies with HIGH as reference. $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

The bivariate regression (column 1) shows a *positive* coefficient: commodity sectors have higher markups. This reverses in column (3) once log(ceiling) is controlled—commodity contracts are typically small and hit the ceiling exactly (markup ≈ 1), while large differentiated contracts have structural slack. Conditional on contract size, more

substitutable goods have *lower* markups ($\beta = -0.038$, $p = 0.003$), as the framework predicts.

The category dummy specification (column 4, $R^2 = 0.23$) sharpens the result. Relative to HIGH (commodity) contracts, LOW (differentiated) contracts show 23.3 percentage points higher markup ($p < 10^{-14}$), and MEDIUM contracts show 37.3 pp higher ($p < 10^{-28}$). The Kruskal-Wallis test confirms the group differences non-parametrically ($H = 183.8$, $p < 10^{-40}$).

The conditional coefficients exhibit a non-monotonicity: MEDIUM exceeds LOW in column 4. This reflects compositional heterogeneity within the MEDIUM category, which spans manufacturing, construction, and transportation—sectors combining both commodity-like inputs (raw materials, fuel) and differentiated elements (custom fabrication, specialized logistics). The unconditional medians follow the predicted monotonic ordering: HIGH 0.89, MEDIUM 0.98, LOW 1.00, suggesting the conditional non-monotonicity is driven by within-category compositional effects rather than a failure of the theoretical prediction. At the 4-digit NAICS level, the pattern is driven by wholesale trade and agriculture (median markup = 1.0) versus R&D services and computer systems design (median markup = 0.83–0.91).

Interpretation. The PoIC test operates in a different domain than the GFC test (procurement vs. banking) and tests a different derivation (Myerson mechanism design, Section 6, vs. Akerlof lemons, Section 5). The shared prediction—that CES curvature K determines vulnerability to information frictions—holds in both. The same caveats apply: NAICS sectors are crude proxies for CES substitutability, and federal procurement has institutional features (FAR regulations, set-asides) not captured by the model. Nonetheless, the result establishes that the PoIC prediction has empirical content outside the financial domain.

12 The Architecture

12.1 What Is Unified

Table 4 summarizes the unification. Six areas of economic theory emerge as specific instances of the free energy framework (11).

12.2 What Determines What

The two parameters have clean domains:

- ρ (CES curvature) determines **structure**: whether markets are competitive or monopolistic, whether networks have strong or weak effects, whether assets are spe-

Area	Key result derived	New ρ-dependent prediction
Information economics	Unraveling = CES degrading under entropy	$T^* \propto K$: complements resist adverse selection *(tested: §11)*
Search / matching	Matching function = CES; search = entropy reduction	ρ explains duration heterogeneity, Beveridge shift
Mechanism design	Virtual valuation = free energy gradient	PoIC(ρ); mechanism format from ρ *(tested: §11.7)*
Contract theory	Hold-up, integration from ρ; asset specificity = low ρ	Property rights and transaction costs unified through ρ
Social choice	Cardinal aggregation at $T>0$ bypasses ordinal impossibility	Democratic robustness depends on ρ
Behavioral economics	Behavioral catalog = $T>0$ phenomena	$T_{\text{gain}}/T_{\text{loss}}$ unifies loss aversion family

Table 4: Unification of six areas of economic theory through the free energy framework $\mathcal{F} = \Phi_{\text{CES}}(\rho) - T \cdot H$. Each area emerges as information friction degrading a CES aggregate, with severity controlled by ρ.

cific or generic, whether mechanisms require complex screening or simple auctions, whether political systems tolerate noise or gridlock.

- T (information temperature) determines **friction**: how precisely agents optimize, how quickly markets clear, how much information rent mechanisms must concede, how well democratic aggregation approximates the social optimum, how large behavioral deviations are.

The interaction $\rho \times T$ determines **institutional viability**: whether a given institution (market, contract, mechanism, political system) functions or fails. The critical temperature $T^* \propto K(\rho)$ is the institutional failure boundary.

12.3 What Remains Separate

The framework does not claim to derive all of economic theory from two parameters. It claims that the *aggregation-information boundary*—the interface between how inputs combine and what agents know—is governed by (ρ, T).

Areas that may require additional structure include:

- Specific institutional forms (why firms rather than markets, why democracies rather than autocracies);
- Historical contingency and path dependence;
- The determination of ρ itself (what makes some inputs complementary and others substitutable);
- The biological/neurological basis of T and the $T_{\text{gain}}/T_{\text{loss}}$ asymmetry.

These are legitimate extensions, not contradictions. The framework provides the scaffolding; the specific content of each domain provides the architecture.

13 Discussion

13.1 Why Wasn't This Seen Before?

The components of this framework are not new. CES aggregates have been used since Arrow et al. (1961). Shannon entropy has been applied to economics since Theil (1967). Free energy principles are standard in physics. Each component is textbook material.

The unity was obscured by the structure of the economics profession. Trade economists use CES for Armington elasticities. Macro economists use CES for production functions. Information economists use entropy for adverse selection. Mechanism designers use entropy for incentive constraints. Behavioral economists use noise models for bounded rationality. Each group uses its piece of the framework without recognizing it as a piece.

The CES quadruple role theorem (Smirl, 2026a) was the key that unlocked the unity. Once one recognizes that ρ simultaneously controls superadditivity, correlation robustness, strategic independence, and network scaling, the question "what else does ρ control?" leads directly to the information-theoretic extension.

13.2 Testability

Every derivation in this paper produces predictions that the original theory does not:

- Akerlof + ρ: market robustness to adverse selection varies with complementarity. *Tested in Section 11*: the ARI $\times$ PMI interaction predicts GFC severity ($p = 0.016$), with placebo indices insignificant.

- Myerson + ρ: price of incentive compatibility and optimal mechanism format depend on complementarity. *Tested in Section 11.7*: differentiated-goods contracts show 23 pp higher ceiling extraction than commodity contracts ($p < 10^{-14}$).

- Arrow + ρ: democratic institutional robustness depends on the social welfare parameter.

- DMP + ρ: unemployment duration heterogeneity and Beveridge curve shifts explained by skill complementarity.

- Hart-Moore + Williamson + ρ: asset specificity, hold-up severity, and vertical integration decisions unified through complementarity.

- Kahneman-Tversky + $T_{\text{gain}}/T_{\text{loss}}$: loss aversion coefficient is a temperature ratio.

The first two predictions are tested directly: the Akerlof prediction in Section 11 using the 2007–2009 GFC as a natural experiment, and the Myerson prediction in Section 11.7 using US federal procurement data. That both tests support the framework in different domains (banking, procurement) with different mechanisms ($\sigma \times T$ interaction, PoIC monotonicity) strengthens the case that ρ is the operative parameter. The remaining predictions are cross-sectionally testable: estimate ρ for each market/institution and check whether the framework's predictions hold across markets with different ρ values.

13.3 Implications for Economic Methodology

If the framework is correct, several methodological implications follow:

1. The division of economics into "micro" and "macro," or "rational" and "behavioral," reflects the two terms of the free energy, not a fundamental disciplinary boundary.

2. The assumption of "perfect rationality" is not wrong—it is the $T = 0$ limit of a more general theory, valid when information costs are negligible relative to stakes.

3. The "biases" documented by behavioral economics are not failures of rationality but the thermodynamically inevitable consequences of finite information processing capacity.

4. Arrow's impossibility theorem is an exact result about ordinal aggregation. It does not preclude non-dictatorial welfare judgments—it demonstrates that such judgments require cardinal information, which noisy optimization at $T > 0$ provides. The cardinal CES family is the unique aggregation satisfying Pareto, anonymity, and homotheticity; its departure from ordinal IIA is the measurable price of non-dictatorship.

14 Conclusion

This paper has proposed that economic theory is generated by two canonical functions—the CES aggregate and Shannon entropy—connected through a free energy principle. The framework is parameterized by ρ (controlling aggregation structure) and T (controlling information friction), with the interaction determining institutional viability.

Six detailed derivations demonstrate that six major areas of economic theory emerge as specific instances of this framework. Each derivation produces new ρ-dependent predictions that the original theories cannot generate. Two empirical tests in different domains confirm the central mechanism. First, pre-crisis banking regulation data from 147 countries shows that the interaction between product substitutability and information quality predicts 2007–2009 crisis severity ($p = 0.016$), while placebo indices do not—resolving a

longstanding puzzle in the banking regulation literature. Second, 1,172 US federal procurement contracts show that differentiated goods (low ρ) exhibit systematically higher information rents than commodity goods ($p = 0.003$), as the mechanism design derivation predicts.

The framework is a conjecture, not a completed theory. Rigorous formalization of each derivation, further empirical testing of the ρ-dependent predictions, and investigation of the framework's boundaries are needed. But the structural correspondence across six independent areas of economic theory—combined with the out-of-sample empirical confirmation—is difficult to attribute to coincidence. If even a subset of the derivations survives formal scrutiny, the implication is that the economics profession has spent decades independently discovering specific consequences of a two-parameter generating function—in separate subfields, with separate terminology—without recognizing the unity.

The equation is:

$$\mathcal{F} = \Phi_{\text{CES}}(\rho) - T \cdot H$$

Two parameters. One equation. The rest is application.

References

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.

Becker, Gary S. 1973. "A Theory of Marriage: Part I." *Journal of Political Economy* 81(4): 813–846.

Bergemann, Dirk, and Stephen Morris. 2019. "Information Design: A Unified Perspective." *Journal of Economic Literature* 57(1): 44–95.

Bouchaud, Jean-Philippe, and Marc Mézard. 2000. "Wealth Condensation in a Simple Model of Economy." *Physica A* 282(3–4): 536–545.

Camerer, Colin F., George Loewenstein, and Matthew Rabin, eds. 2004. *Advances in Behavioral Economics*. Princeton: Princeton University Press.

Caplin, Andrew, and Mark Dean. 2015. "Revealed Preference, Rational Inattention, and Costly Information Acquisition." *American Economic Review* 105(7): 2183–2203.

Caplin, Andrew, Mark Dean, and John Leahy. 2019. "Rational Inattention, Optimal Consideration Sets, and Stochastic Choice." *Review of Economic Studies* 86(3): 1061–1094.

Clarke, Edward H. 1971. "Multipart Pricing of Public Goods." *Public Choice* 11(1): 17–33.

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.

Cover, Thomas M., and Joy A. Thomas. 2006. *Elements of Information Theory*. 2nd ed. Hoboken: Wiley.

Costinot, Arnaud. 2009. "An Elementary Theory of Comparative Advantage." *Econometrica* 77(4): 1165–1192.

Costinot, Arnaud, and Jonathan Vogel. 2010. "Matching and Inequality in the World Economy." *Journal of Political Economy* 118(4): 747–786.

Crémer, Jacques, and Richard P. McLean. 1988. "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions." *Econometrica* 56(6): 1247–1257.

Diamond, Peter A. 1982. "Aggregate Demand Management in Search Equilibrium." *Journal of Political Economy* 90(5): 881–894.

Drăgulescu, Adrian A., and Victor M. Yakovenko. 2000. "Statistical Mechanics of Money." *European Physical Journal B* 17(4): 723–729.

Dixit, Avinash K., and Joseph E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review* 67(3): 297–308.

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.

Gibbard, Allan. 1973. "Manipulation of Voting Schemes: A General Result." *Econometrica* 41(4): 587–601.

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.

Holmström, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7(Special Issue): 24–52.

Hosios, Arthur J. 1990. "On the Efficiency of Matching and Related Models of Search and Unemployment." *Review of Economic Studies* 57(2): 279–298.

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.

Krugman, Paul. 1991. "Increasing Returns and Economic Geography." *Journal of Political Economy* 99(3): 483–499.

Laeven, Luc, and Fabián Valencia. 2018. "Systemic Banking Crises Revisited." IMF Working Paper 18/206.

Lagos, Ricardo, and Randall Wright. 2005. "A Unified Framework for Monetary Theory and Policy Analysis." *Journal of Political Economy* 113(3): 463–484.

Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics* 112(2): 443–478.

List, Christian, and Philip Pettit. 2002. "Aggregating Sets of Judgments: An Impossibility Result." *Economics and Philosophy* 18(1): 89–110.

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.

McKelvey, Richard D., and Thomas R. Palfrey. 1995. "Quantal Response Equilibria for Normal Form Games." *Games and Economic Behavior* 10(1): 6–38.

Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics* 12(2): 380–391.

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.

Moulin, Hervé. 1980. "On Strategy-Proofness and Single Peakedness." *Public Choice* 35(4): 437–455.

Myerson, Roger B. 1981. "Optimal Auction Design." *Mathematics of Operations Research* 6(1): 58–73.

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.

Prelec, Dražen. 1998. "The Probability Weighting Function." *Econometrica* 66(3): 497–527.

Rabin, Matthew. 2000. "Risk Aversion and Expected-Utility Theory: A Calibration Theorem." *Econometrica* 68(5): 1281–1292.

Ramstead, Maxwell J. D., Paul B. Badcock, and Karl J. Friston. 2018. "Answering Schrödinger's Question: A Free-Energy Formulation of the Life Sciences." *Physics of Life Reviews* 24: 1–16.

Rochet, Jean-Charles, and Philippe Choné. 1998. "Ironing, Sweeping, and Multidimensional Screening." *Econometrica* 66(4): 783–826.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98(5): S71–S102.

Rothschild, Michael, and Joseph Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics* 90(4): 629–649.

Sato, Kazuo. 1967. "A Two-Level Constant-Elasticity-of-Substitution Production Function." *Review of Economic Studies* 34(2): 201–218.

Sen, Amartya K. 1970. "The Impossibility of a Paretian Liberal." *Journal of Political Economy* 78(1): 152–157.

Shannon, Claude E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27(3): 379–423.

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. "The CES Quadruple Role: Superadditivity, Correlation Robustness, Strategic Independence, and Network Scaling as Four Properties of CES Curvature." SSRN Working Paper 6263482. `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6263482`.

Smirl, Jon. 2026b. "Emergent CES: Why Constant Elasticity of Substitution Is Not an Assumption." Working Paper.

Solow, Robert M. 1956. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70(1): 65–94.

Spence, Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics* 87(3): 355–374.

Stiglitz, Joseph E. 2000. "The Contributions of the Economics of Information to Twentieth Century Economics." *Quarterly Journal of Economics* 115(4): 1441–1478.

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.

Tirole, Jean. 1999. "Incomplete Contracts: Where Do We Stand?" *Econometrica* 67(4): 741–781.

Tversky, Amos, and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty* 5(4): 297–323.

Vickrey, William. 1961. "Counterspeculation, Auctions, and Competitive Sealed Tenders." *Journal of Finance* 16(1): 8–37.

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.

Woodford, Michael. 2009. "Information-Constrained State-Dependent Pricing." *Journal of Monetary Economics* 56(Supplement): S100–S124.

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.