

Predictive Market Theory: A Bayesian–Probabilistic Foundation for Markets as Prediction Engines

Part II of: Markets as Distributed Bayesian Inference Engines

Matthew Long

*The YonedaAI Collaboration
YonedaAI Research Collective
Chicago, IL*

matthew@yonedaai.com · <https://yonedaai.com>

March 4, 2026

Abstract

We introduce *Predictive Market Theory* (PMT), a new axiomatic framework that reconceptualizes financial markets as probabilistic prediction engines whose fundamental output is not price *per se* but rather a continuously updated posterior probability measure over future states of the world. Building on the distributed Bayesian inference perspective of Part I, we elevate the predictive function from an emergent property to the foundational primitive: rather than deriving prediction as a consequence of equilibrium pricing, we derive pricing as a consequence of collective prediction. The theory is erected on five axioms—*Predictive Completeness*, *Bayesian Coherence*, *Informational Irreversibility*, *Predictive Sufficiency of Price*, and *Thermodynamic Dissipation*—from which we derive the standard results of asset pricing theory (martingale pricing, risk-neutral valuation, the fundamental theorem of asset pricing) as corollaries of a deeper predictive structure. We introduce the *predictive measure* $\mathcal{P}$, a probability measure on outcome space that is distinct from both the physical measure $\mathbb{P}$ and the risk-neutral measure $\mathbb{Q}$, and show that it encodes the market’s best collective forecast. We prove that the predictive measure satisfies a variational principle: it minimizes the expected free energy of prediction over all belief-consistent measures, unifying Bayesian inference, information theory, and thermodynamics within a single market-theoretic framework. The theory resolves several longstanding puzzles—the equity premium puzzle, excess volatility, the value of active management—through the lens of predictive inefficiency and the cost of maintaining the market’s predictive function. We establish formal connections to the theory of prediction markets, Kelly criterion portfolio theory, and de Finetti’s operational subjective probability, and we prove a *Predictive Market Completeness Theorem* showing that a sufficiently liquid market generates a complete probability measure over any measurable partition of future states. Finally, we derive novel results on the *predictive resolution* of markets—a quantitative measure of forecasting precision—and show it is governed by a fundamental uncertainty relation linking temporal resolution to state-space resolution.

Keywords: predictive market theory, Bayesian probability, market prediction, predictive measure, variational inference, free energy principle, equity premium puzzle, market completeness, Kelly criterion, de Finetti, uncertainty relations

Contents

1	Introduction	4
1.1	From Inference to Prediction	4
1.2	Historical and Intellectual Context	4
1.3	Overview of Contributions	5
2	The Axioms of Predictive Market Theory	5
2.1	Discussion of the Axioms	6
2.1.1	Axiom 1: Predictive Completeness	6
2.1.2	Axiom 2: Bayesian Coherence	7
2.1.3	Axiom 3: Informational Irreversibility	7
2.1.4	Axiom 4: Predictive Sufficiency of Price	7
2.1.5	Axiom 5: Thermodynamic Dissipation	7
3	The Predictive Measure	8
3.1	Three Measures in Finance	8
3.2	Relationship Between $\mathbb{P}$, $\mathbb{Q}$, and $\mathcal{P}$	8
3.3	Extracting $\mathcal{P}$ from Market Data	9
3.3.1	Method 1: Calibration via Prediction Markets	9
3.3.2	Method 2: Power Utility Inversion	9
3.4	The Predictive Measure as a Martingale Measure	9
4	The Variational Prediction Principle	10
4.1	Free Energy of Prediction	10
4.2	The Market as a Variational Inference Machine	11
4.3	Connection to the Free Energy Principle	12
5	Deriving Price from Prediction	12
5.1	The Pricing Functional	12
5.2	Recovering Classical Results	13
6	The Fundamental Theorem of Predictive Markets	13
6.1	Predictive Completeness	13
6.2	Implications for Market Design	14
7	Predictive Resolution and Uncertainty Relations	15
7.1	Quantifying Forecasting Precision	15
7.2	The Predictive Uncertainty Relation	15
7.3	Consequences of the Uncertainty Relation	16
8	Resolution of Classical Puzzles	16
8.1	The Equity Premium Puzzle	16
8.2	The Excess Volatility Puzzle	17
8.3	The Value of Active Management	17
9	Connections to Kelly Theory and de Finetti	18
9.1	The Kelly Criterion as Optimal Prediction Exploitation	18
9.2	De Finetti's Theorem and Market Coherence	18
9.3	Scoring Rules and the Market as a Mechanism	19

10 Predictive Dynamics: The Stochastic Evolution of $\mathcal{P}_t$	19
10.1 The Predictive Process	19
10.2 The Stochastic Differential Equation for $\mathcal{P}_t$	20
10.3 Information Arrival Rate and Predictive Volatility	20
11 Quantifying Market Prediction Accuracy	21
11.1 Calibration and Discrimination	21
11.2 The Predictive Efficiency Ratio	21
11.3 Comparing Predictive Efficiency Across Markets	22
12 Predictive Market Microstructure	22
12.1 The Predictive Interpretation of Order Flow	22
12.2 The Bid–Ask Spread as Predictive Insurance	23
12.3 Market Making as Predictive Mediation	23
13 Empirical Predictions and Experimental Design	23
13.1 Core Empirical Predictions	23
13.2 Experimental Design: A Controlled Test of PMT	24
13.3 Data Requirements and Feasibility	24
14 Conclusion	25

1 Introduction

1.1 From Inference to Prediction

In Part I of this series (19), we established that financial markets function as distributed Bayesian inference engines, aggregating heterogeneous private information into equilibrium prices through a process formally equivalent to posterior computation in probabilistic graphical models. The present paper takes the next logical step: if markets perform inference, what exactly do they infer? And can we build a complete theory of markets by taking the *predictive function* as the foundational primitive?

The classical answer—that markets estimate fundamental value—is unsatisfying for several reasons. “Fundamental value” is itself a derived concept, defined as the expected discounted sum of future cashflows. But expected with respect to *which* probability measure? Discounted at *which* rate? And who decides? These circularities suggest that the conventional approach has the logical order backwards: it attempts to define value first and derive prediction as a byproduct, when in fact prediction is the more primitive operation.

We propose a radical inversion. In *Predictive Market Theory* (PMT), the fundamental output of a market is not a price but a *prediction*—specifically, a probability measure over the space of possible future states of the world. Price is then derived from this predictive measure through a pricing functional that reflects agents’ risk preferences. This inversion is not merely philosophical: it leads to different mathematics, different theorems, and different empirical predictions.

1.2 Historical and Intellectual Context

The idea that markets are prediction machines has deep roots. The earliest formal prediction markets date to the Iowa Electronic Markets established in 1988 for forecasting election outcomes. Arrow (2) recognized that competitive equilibria in complete markets are isomorphic to probability measures over states of nature. De Finetti’s foundational work on subjective probability (9) established that coherent betting behavior implies—and is implied by—a probability measure, a result that applies directly to market prices of contingent claims.

More recently, the literature on prediction markets (25; 14; 3) has explicitly studied markets designed for forecasting. The key insight from this literature is that market prices for binary contracts (“Will event E occur?”) can be interpreted directly as probability estimates. What has been missing is a unified theoretical framework that:

- (i) Extends the prediction market insight from binary contracts to the full complexity of financial markets.
- (ii) Provides axiomatic foundations for the market’s predictive function.
- (iii) Derives standard asset pricing results as consequences of the predictive framework.
- (iv) Quantifies the precision and limitations of market prediction.
- (v) Connects to the thermodynamic and information-theoretic structures identified in Part I.

Predictive Market Theory accomplishes all five.

1.3 Overview of Contributions

The principal contributions of this paper are:

1. **Axiomatic foundations** (Section 2): Five axioms from which the entire theory follows.
2. **The Predictive Measure $\mathcal{P}$** (Section 3): A new probability measure on outcome space, distinct from both $\mathbb{P}$ and $\mathbb{Q}$, encoding the market’s collective forecast.
3. **Variational Prediction Principle** (Section 4): The predictive measure minimizes expected predictive free energy.
4. **Price as a Functional of Prediction** (Section 5): Standard pricing results derived from the predictive framework.
5. **The Fundamental Theorem of Predictive Markets** (Section 6): A market-completeness result for predictive measures.
6. **Predictive Resolution and Uncertainty Relations** (Section 7): Fundamental limits on market forecasting precision.
7. **Resolution of Classical Puzzles** (Section 8): New explanations for the equity premium puzzle and excess volatility.
8. **Connections to Kelly Theory and de Finetti** (Section 9): Unifying threads with optimal betting and operational subjective probability.
9. **Predictive Dynamics** (Section 10): The stochastic evolution of the predictive measure.
10. **Empirical predictions and experimental design** (Section 13).

2 The Axioms of Predictive Market Theory

We now state the five axioms on which Predictive Market Theory is erected. Throughout, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space where Ω is the space of possible future states of the world, $\mathcal{F}$ is a σ -algebra of events, and $\mathbb{P}$ is the physical (objective) probability measure. Let $\{P_t^{(k)}\}_{k \in \mathcal{K}}$ denote the collection of prices of all traded assets at time t , where $\mathcal{K}$ is the asset index set. Let $\mathcal{F}_t$ denote the market filtration at time t —the information available from observing market activity up to time t .

Axiom 1 (Predictive Completeness). *For every event $A \in \mathcal{F}$ with $\mathbb{P}(A) \in (0, 1)$, there exists (in principle) a portfolio of traded assets whose market value at time t encodes the market’s conditional probability estimate $\mathcal{P}_t(A) = \mathcal{P}(A \mid \mathcal{F}_t)$.*

Axiom 2 (Bayesian Coherence). *The predictive measure $\mathcal{P}_t$ is a proper probability measure on $(\Omega, \mathcal{F})$ for all t , and its evolution satisfies Bayes' rule: for any event A and any new information $\mathcal{I}_{t+1}$ arriving at time $t + 1$:*

$$\mathcal{P}_{t+1}(A) = \frac{\mathcal{P}_t(A \mid \mathcal{I}_{t+1}) \cdot \mathcal{P}_t(\mathcal{I}_{t+1})}{\mathcal{P}_t(\mathcal{I}_{t+1})} = \mathcal{P}_t(A \mid \mathcal{I}_{t+1}) \quad (1)$$

Moreover, if no new information arrives, the predictive measure does not change:

$$\mathcal{I}_{t+1} = \emptyset \implies \mathcal{P}_{t+1} = \mathcal{P}_t \quad (2)$$

Axiom 3 (Informational Irreversibility). *The evolution of the predictive measure is informationally irreversible: the total information content of the market (as measured by the negative entropy of $\mathcal{P}_t$) is non-decreasing:*

$$H[\mathcal{P}_t] \leq H[\mathcal{P}_s] \quad \text{for all } t > s \quad (3)$$

where $H[\mathcal{P}] = -\int_{\Omega} \frac{d\mathcal{P}}{d\lambda} \log \frac{d\mathcal{P}}{d\lambda} d\lambda$ is the differential entropy relative to a reference measure λ . Equivalently, the predictive measure becomes progressively more concentrated as information accumulates.

Axiom 4 (Predictive Sufficiency of Price). *The vector of market prices $\mathbf{P}_t = \{P_t^{(k)}\}_{k \in \mathcal{K}}$ is a sufficient statistic for the predictive measure $\mathcal{P}_t$ in the sense that:*

$$\mathcal{P}_t(\cdot) = \mathcal{P}(\cdot \mid \mathbf{P}_t) \quad (4)$$

The predictive measure is uniquely determined by, and recoverable from, market prices.

Axiom 5 (Thermodynamic Dissipation). *Every update of the predictive measure dissipates predictive free energy. Specifically, the change in the predictive free energy:*

$$\mathcal{F}[\mathcal{P}_t] = \mathbb{E}_{\mathcal{P}_t}[-\log \mathbb{P}(\omega)] + \mathbb{E}_{\mathcal{P}_t}[\log \mathcal{P}_t(\omega)] \quad (5)$$

satisfies:

$$\Delta \mathcal{F} = \mathcal{F}[\mathcal{P}_{t+1}] - \mathcal{F}[\mathcal{P}_t] \leq 0 \quad (6)$$

with equality if and only if no genuinely new information is incorporated. The predictive free energy is a Lyapunov function for the market dynamics.

2.1 Discussion of the Axioms

2.1.1 Axiom 1: Predictive Completeness

This is the predictive analogue of market completeness in the Arrow–Debreu sense (1; 8). In the Arrow–Debreu framework, a complete market has a traded security for every state of nature. Predictive Completeness is weaker: it requires only that a *portfolio* (possibly synthetic) exists whose price encodes the probability of any event. This is achievable whenever the market has a sufficient variety of derivative instruments, particularly options at different strike prices.

The Breeden–Litzenberger result (6) provides the classical construction: the price of a European call option $C(K)$ with strike K encodes the tail probability of the underlying exceeding K at expiration. Formally:

$$\mathcal{P}_t(S_T > K) = -e^{r(T-t)} \frac{\partial C(K)}{\partial K} \quad (7)$$

and the full predictive density is:

$$\frac{d\mathcal{P}_t}{dK}(K) = e^{r(T-t)} \frac{\partial^2 C(K)}{\partial K^2} \quad (8)$$

We emphasize that our $\mathcal{P}$ is *not* the risk-neutral measure $\mathbb{Q}$. The relationship between $\mathcal{P}$, $\mathbb{Q}$, and $\mathbb{P}$ is developed in Section 3.

2.1.2 Axiom 2: Bayesian Coherence

This axiom asserts that the market, viewed as a collective prediction device, obeys the laws of probability. It is the market-level analogue of individual rationality: just as a rational agent’s beliefs satisfy the axioms of probability, so too do the market’s collective beliefs. The key empirical content is the dynamic updating rule—the market does not simply assign static probabilities but revises them in response to new information in a manner consistent with Bayes’ theorem.

This axiom rules out certain pathological market behaviors:

- **Dutch books:** If $\mathcal{P}_t$ were not a proper probability measure, an arbitrageur could construct a Dutch book—a combination of trades that guarantees a profit regardless of the outcome (9).
- **Non-Bayesian updating:** If the market updated in a manner inconsistent with Bayes’ rule (e.g., overreacting systematically to new information), the resulting $\mathcal{P}$ would fail the coherence axiom.

2.1.3 Axiom 3: Informational Irreversibility

This axiom encodes a form of the second law of thermodynamics for markets. As information is incorporated, the predictive distribution sharpens—uncertainty decreases, and this process is irreversible. The market cannot “unlearn” information that has been publicly reflected in prices.

Note that this does not preclude increased *price volatility*. Volatility can increase even as entropy decreases: a bimodal distribution (e.g., “the company will either succeed spectacularly or fail completely”) can have lower entropy than a diffuse unimodal distribution, yet exhibit higher variance. The axiom constrains entropy, not variance.

2.1.4 Axiom 4: Predictive Sufficiency of Price

This is the strongest axiom and the one most subject to empirical challenge. It asserts that prices contain all the information needed to reconstruct the predictive measure. In its strong form, this is equivalent to the strong-form Efficient Market Hypothesis. We will later weaken it to an approximate version, introducing the concept of *predictive residuals*.

2.1.5 Axiom 5: Thermodynamic Dissipation

This axiom connects the predictive theory to the thermodynamic framework of Part I. The predictive free energy combines two terms: the expected surprise of outcomes under the predictive measure (an accuracy term) and the KL divergence from the

predictive measure to the physical measure (a complexity or confidence penalty). The axiom states that the market’s prediction improves over time in the free-energy sense.

3 The Predictive Measure

3.1 Three Measures in Finance

Standard mathematical finance operates with two probability measures:

- (i) The *physical measure* $\mathbb{P}$: the objective probability distribution governing actual outcomes. Under $\mathbb{P}$, stocks have positive expected returns (the equity premium), and the expected return on any asset reflects its systematic risk.
- (ii) The *risk-neutral measure* $\mathbb{Q}$: a mathematical construct under which all assets have the same expected return (the risk-free rate). The risk-neutral measure is not a genuine probability forecast—it is a pricing convenience that encodes risk preferences alongside probability assessments.

We introduce a third measure:

Definition 3.1 (The Predictive Measure). The *predictive measure* $\mathcal{P}_t$ is the probability measure on $(\Omega, \mathcal{F})$ that represents the market’s collective best estimate of the likelihood of future states, purged of risk preference distortions. Formally, $\mathcal{P}_t$ is defined by the requirement that for any event $A \in \mathcal{F}$:

$$\mathcal{P}_t(A) = \text{“the market’s probability forecast for } A \text{ at time } t\text{”} \quad (9)$$

subject to the constraint that $\mathcal{P}_t$ is coherent (Axiom 2) and informationally consistent with market prices (Axiom 4).

3.2 Relationship Between $\mathbb{P}$, $\mathbb{Q}$, and $\mathcal{P}$

The three measures are related through Radon–Nikodym derivatives. Under regularity conditions (absolute continuity), we can write:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{d\mathbb{Q}}{d\mathcal{P}} \cdot \frac{d\mathcal{P}}{d\mathbb{P}} \quad (10)$$

The Radon–Nikodym derivative $d\mathbb{Q}/d\mathbb{P}$ is the classical *pricing kernel* or *stochastic discount factor* M :

$$M(\omega) = \frac{d\mathbb{Q}}{d\mathbb{P}}(\omega) \quad (11)$$

We decompose this into two components:

Definition 3.2 (Predictive Accuracy and Risk Distortion). The pricing kernel admits a multiplicative decomposition:

$$M(\omega) = A(\omega) \cdot R(\omega) \quad (12)$$

where:

$$A(\omega) = \frac{d\mathcal{P}}{d\mathbb{P}}(\omega) \quad (\text{predictive accuracy kernel}) \quad (13)$$

$$R(\omega) = \frac{d\mathbb{Q}}{d\mathcal{P}}(\omega) \quad (\text{risk distortion kernel}) \quad (14)$$

The predictive accuracy kernel $A(\omega)$ measures how well the market’s collective forecast matches the true probabilities: if $A(\omega) = 1$ for all ω , the market is a perfect forecaster. The risk distortion kernel $R(\omega)$ captures pure risk preferences: even a perfectly forecasting market will price risky assets differently from risk-neutral valuation because agents demand compensation for bearing risk.

Proposition 3.3 (Measure Separation). *The three measures $\mathbb{P}$, $\mathcal{P}$, $\mathbb{Q}$ satisfy the following chain of relationships:*

$$\mathbb{P} \xrightarrow{\text{market prediction}} \mathcal{P} \xrightarrow{\text{risk adjustment}} \mathbb{Q} \quad (15)$$

where:

- (i) $\mathcal{P}$ is the market’s probabilistic forecast, which may differ from $\mathbb{P}$ due to informational limitations and behavioral biases.
- (ii) $\mathbb{Q}$ adjusts $\mathcal{P}$ for aggregate risk preferences, overweighting bad states (where marginal utility is high) and underweighting good states.

The standard “ $\mathbb{P}$ to $\mathbb{Q}$ directly” approach conflates these two distinct transformations.

3.3 Extracting $\mathcal{P}$ from Market Data

In practice, how do we disentangle the predictive measure $\mathcal{P}$ from the risk-neutral measure $\mathbb{Q}$ observed in option prices? We provide two methods.

3.3.1 Method 1: Calibration via Prediction Markets

If a prediction market exists for the same event, the prediction market price directly estimates $\mathcal{P}(A)$, while the option market provides $\mathbb{Q}(A)$. The ratio yields the risk distortion kernel evaluated on the event:

$$R(A) = \frac{\mathbb{Q}(A)}{\mathcal{P}(A)} \quad (16)$$

3.3.2 Method 2: Power Utility Inversion

Under the assumption of power utility with relative risk aversion γ , the predictive density can be recovered from the risk-neutral density by:

$$\frac{d\mathcal{P}}{dx}(x) \propto \frac{d\mathbb{Q}}{dx}(x) \cdot x^\gamma \quad (17)$$

where $x = S_T/S_t$ is the gross return. This follows from the first-order condition of a representative agent with power utility.

3.4 The Predictive Measure as a Martingale Measure

A fundamental structural property of the predictive measure is that it generates a martingale—but not on prices. Rather, it is the *predictive probabilities themselves* that form a martingale:

Theorem 3.4 (Predictive Martingale Property). *Under Axioms 2–4, for any event $A \in \mathcal{F}$:*

$$\mathbb{E}_{\mathcal{P}_t}[\mathcal{P}_{t+1}(A) \mid \mathcal{F}_t] = \mathcal{P}_t(A) \quad (18)$$

That is, the market’s probability forecasts are martingales under the predictive measure itself.

Proof. By the tower property of conditional expectation and the Bayesian coherence axiom:

$$\mathbb{E}_{\mathcal{P}_t}[\mathcal{P}_{t+1}(A) \mid \mathcal{F}_t] = \mathbb{E}_{\mathcal{P}_t}[\mathcal{P}(A \mid \mathcal{F}_{t+1}) \mid \mathcal{F}_t] \quad (19)$$

$$= \mathcal{P}(A \mid \mathcal{F}_t) \quad (20)$$

$$= \mathcal{P}_t(A) \quad (21)$$

The second equality uses the tower property, which holds for any probability measure. $\square$

Corollary 3.5 (Forecast Unbiasedness). *If $\mathcal{P} = \mathbb{P}$ (the market is a perfect forecaster), then the predictive probabilities are unbiased estimators of the true probabilities:*

$$\mathbb{E}_{\mathbb{P}}[\mathcal{P}_t(A)] = \mathbb{P}(A) \quad \text{for all } A, t \quad (22)$$

This is a strong testable prediction: realized frequencies of events should match their market-implied probabilities, on average.

4 The Variational Prediction Principle

4.1 Free Energy of Prediction

We now show that the predictive measure is characterized by a variational principle: it minimizes the expected *predictive free energy* over all measures consistent with the market’s information.

Definition 4.1 (Predictive Free Energy). The predictive free energy of a candidate measure μ relative to the market information $\mathcal{F}_t$ is:

$$\mathcal{F}[\mu; \mathcal{F}_t] = \underbrace{\mathbb{E}_{\mu}[-\log p(\mathcal{D}_t \mid \omega)]}_{\text{prediction error}} + \underbrace{\text{KL}(\mu \parallel \pi_0)}_{\text{complexity penalty}} \quad (23)$$

where $\mathcal{D}_t$ denotes the data (market observations) available at time t , $p(\mathcal{D}_t \mid \omega)$ is the likelihood of the data given state ω , and π_0 is the prior measure.

The predictive free energy has two competing terms:

- The **prediction error** term penalizes measures that assign low likelihood to the observed data. A measure that “predicted poorly” in light of what has been observed receives a large penalty.
- The **complexity penalty** (KL divergence from the prior) penalizes measures that deviate far from the prior, implementing a form of Occam’s razor: among measures that explain the data equally well, prefer the simplest (closest to the prior).

Theorem 4.2 (Variational Characterization of $\mathcal{P}$). *The predictive measure $\mathcal{P}_t$ is the unique minimizer of the predictive free energy:*

$$\mathcal{P}_t = \arg \min_{\mu \in \mathcal{M}(\Omega)} \mathcal{F}[\mu; \mathcal{F}_t] \quad (24)$$

where $\mathcal{M}(\Omega)$ is the space of probability measures on $(\Omega, \mathcal{F})$. Moreover, the minimum value of the free energy is:

$$\mathcal{F}[\mathcal{P}_t; \mathcal{F}_t] = -\log Z_t \quad (25)$$

where $Z_t = \int p(\mathcal{D}_t \mid \omega) d\pi_0(\omega)$ is the marginal likelihood (evidence).

Proof. Expand the free energy:

$$\mathcal{F}[\mu; \mathcal{F}_t] = \mathbb{E}_\mu[-\log p(\mathcal{D}_t \mid \omega)] + \int \log \frac{d\mu}{d\pi_0} d\mu \quad (26)$$

$$= \int \left(-\log p(\mathcal{D}_t \mid \omega) + \log \frac{d\mu}{d\pi_0}(\omega) \right) d\mu(\omega) \quad (27)$$

$$= \int \log \frac{d\mu(\omega)}{p(\mathcal{D}_t \mid \omega) d\pi_0(\omega)} d\mu(\omega) \quad (28)$$

$$= \text{KL} \left(\mu \parallel \frac{p(\mathcal{D}_t \mid \cdot) \pi_0}{Z_t} \right) - \log Z_t \quad (29)$$

The first term is a KL divergence, which is non-negative and equals zero if and only if:

$$\mu(\omega) = \frac{p(\mathcal{D}_t \mid \omega) \pi_0(\omega)}{Z_t} \quad (30)$$

This is precisely the Bayesian posterior, which we identify with $\mathcal{P}_t$. The minimum free energy value is $-\log Z_t$. $\square$

4.2 The Market as a Variational Inference Machine

Theorem 4.2 reveals that the market performs *variational inference*—the same computational framework used in modern machine learning (5). The market searches over the space of probability measures to find the one that best balances explanatory power (fitting observed data) against complexity (staying close to the prior).

This has profound implications:

- (i) **Bounded rationality as approximate inference:** When the market cannot compute the exact posterior (due to bounded rationality, limited information processing, or insufficient liquidity), it effectively performs *approximate* variational inference, searching over a restricted family of measures $\mathcal{M}_0 \subset \mathcal{M}(\Omega)$. The quality of the approximation is measured by the variational gap:

$$\Delta \mathcal{F} = \mathcal{F}[\mathcal{P}_t^{\text{approx}}; \mathcal{F}_t] - \mathcal{F}[\mathcal{P}_t^*; \mathcal{F}_t] = \text{KL}(\mathcal{P}_t^{\text{approx}} \parallel \mathcal{P}_t^*) \geq 0 \quad (31)$$

- (ii) **Model selection:** The marginal likelihood Z_t provides an automatic model selection criterion. Markets that can entertain more complex models (richer asset universes, more derivative instruments) achieve lower free energy, corresponding to better predictions. This provides a principled explanation for financial innovation: new instruments are valuable precisely because they expand the space of measures over which the market can optimize.

- (iii) **The prior as market structure:** The prior π_0 encodes structural assumptions about the economy—the base rate expectations, the background model. Changes in the prior (e.g., regime changes, new economic paradigms) correspond to fundamental reconfigurations of the market’s predictive framework.

4.3 Connection to the Free Energy Principle

The variational formulation connects directly to Friston’s *Free Energy Principle* (11) from computational neuroscience, which posits that biological agents minimize variational free energy to maintain a model of their environment. Under PMT, the market is a collective agent that minimizes predictive free energy, treating the economy as its “environment.” The parallel is exact:

Free Energy Principle (Brain)	PMT (Market)
Organism	Market
Sensory data	Market data, news, signals
Generative model	Collective belief structure
Free energy	Predictive free energy $\mathcal{F}$
Perception	Price discovery
Action	Trading

This is more than analogy. Both the brain and the market face the same fundamental computational problem—maintaining an accurate internal model of an uncertain, changing environment—and both solve it by minimizing variational free energy.

5 Deriving Price from Prediction

5.1 The Pricing Functional

In PMT, price is not a primitive—it is derived from the predictive measure through a *pricing functional* that encodes aggregate risk preferences.

Definition 5.1 (Predictive Pricing Functional). The price of an asset with payoff $X(\omega)$ at time T is:

$$P_t(X) = e^{-r(T-t)} \int_{\Omega} X(\omega) \cdot R(\omega) d\mathcal{P}_t(\omega) \quad (32)$$

where r is the risk-free rate and $R(\omega) = d\mathbb{Q}/d\mathcal{P}(\omega)$ is the risk distortion kernel. Equivalently:

$$P_t(X) = e^{-r(T-t)} [\mathbb{E}_{\mathcal{P}_t}[X] + \text{Cov}_{\mathcal{P}_t}(X, R)] \quad (33)$$

where $\mathbb{E}_{\mathcal{P}_t}[X]$ is the market’s prediction of the expected payoff and $\text{Cov}_{\mathcal{P}_t}(X, R)$ is the risk premium.

Equation (33) makes the decomposition transparent: the asset price equals the discounted predicted expected payoff, adjusted by a risk premium that depends on the covariance of the payoff with the risk distortion kernel.

5.2 Recovering Classical Results

We now show that the standard results of asset pricing theory follow as corollaries of the predictive framework.

Corollary 5.2 (Fundamental Theorem of Asset Pricing). *Under Axioms 1–4, the absence of arbitrage is equivalent to the existence of a predictive measure $\mathcal{P}$ and a risk distortion kernel $R > 0$ such that all asset prices satisfy (32).*

Proof. By Axiom 4, prices determine $\mathcal{P}$. Define $\mathbb{Q}$ by $d\mathbb{Q} = R d\mathcal{P}$. The pricing formula (32) becomes $P_t(X) = e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}[X]$, the standard risk-neutral pricing formula. The existence of $\mathbb{Q} \sim \mathcal{P} \sim \mathbb{P}$ with $\mathbb{Q}$ being an equivalent martingale measure is exactly the first fundamental theorem of asset pricing (15). $\square$

Corollary 5.3 (Capital Asset Pricing Model). *If the risk distortion kernel is an affine function of the market return, $R(\omega) = a + b \cdot R_m(\omega)$, then the CAPM follows:*

$$\mathbb{E}_{\mathcal{P}}[R_i] - r = \beta_i \cdot (\mathbb{E}_{\mathcal{P}}[R_m] - r) \quad (34)$$

where $\beta_i = \text{Cov}_{\mathcal{P}}(R_i, R_m) / \text{Var}_{\mathcal{P}}(R_m)$.

Corollary 5.4 (Black–Scholes as Degenerate Prediction). *The Black–Scholes option pricing formula arises when the predictive measure is log-normal:*

$$\mathcal{P}_t \left(\log \frac{S_T}{S_t} \leq x \right) = \Phi \left(\frac{x - \mu_{\mathcal{P}}(T-t)}{\sigma \sqrt{T-t}} \right) \quad (35)$$

and the risk distortion kernel is exponential-affine in the log-return. In this case, the risk-neutral measure $\mathbb{Q}$ is also log-normal (with drift r instead of $\mu_{\mathcal{P}}$), and the Black–Scholes formula follows.

The PMT perspective reveals that the Black–Scholes model is not merely a pricing formula but a statement about the *form of the market's prediction*: it assumes the market predicts log-normal returns. Deviations from Black–Scholes (volatility smiles, skew) are therefore deviations of the predictive measure from log-normality, not just deviations of the pricing kernel.

6 The Fundamental Theorem of Predictive Markets

We now state and prove the central result of the theory: the Predictive Market Completeness Theorem.

6.1 Predictive Completeness

Definition 6.1 (Predictive Resolution). The *predictive resolution* of a market with respect to a partition $\{A_1, \dots, A_n\}$ of Ω is the vector of predictive probabilities:

$$\boldsymbol{\pi}_t = (\mathcal{P}_t(A_1), \dots, \mathcal{P}_t(A_n)) \quad (36)$$

The market has *full predictive resolution* with respect to this partition if $\boldsymbol{\pi}_t$ can be determined from market prices.

Definition 6.2 (Predictive Span). The *predictive span* of a market is the coarsest σ -algebra $\mathcal{G} \subset \mathcal{F}$ such that the market has full predictive resolution with respect to every finite measurable partition of Ω into $\mathcal{G}$ -measurable sets.

Theorem 6.3 (Fundamental Theorem of Predictive Markets). *Let $\mathcal{K}$ be the set of traded assets, with payoff functions $\{X^{(k)}\}_{k \in \mathcal{K}}$. Then:*

- (i) *The predictive span $\mathcal{G}$ equals the σ -algebra generated by the payoff functions: $\mathcal{G} = \sigma\{X^{(k)} : k \in \mathcal{K}\}$.*
- (ii) *If $\mathcal{G} = \mathcal{F}$ (the payoff functions generate the full σ -algebra), then the market is predictively complete: it generates a full predictive measure $\mathcal{P}_t$ on $(\Omega, \mathcal{F})$.*
- (iii) *The minimum number of assets required for predictive completeness with respect to a finite partition into n states is $n - 1$ (since probabilities sum to one).*
- (iv) *For a predictively complete market, the predictive measure $\mathcal{P}_t$ is the unique measure satisfying Axioms 1–4 and consistent with observed prices.*

Proof. (i) Any event in $\mathcal{G}$ can be expressed as a Boolean combination of events $\{X^{(k)} \in B_k\}$ for Borel sets B_k . The price of an asset with payoff $\mathbf{1}\{X^{(k)} \in B_k\}$ (a digital option) encodes $\mathcal{P}_t(X^{(k)} \in B_k)$ up to the risk distortion. With sufficiently many such digital options (or equivalently, vanilla options at all strikes, via Breeden–Litzenberger), we can recover $\mathcal{P}_t$ restricted to $\mathcal{G}$.

(ii) If $\mathcal{G} = \mathcal{F}$, every event is measurable with respect to the payoff functions, so the construction in (i) determines $\mathcal{P}_t$ on all of $\mathcal{F}$.

(iii) For an n -state partition, the predictive measure is an $(n - 1)$ -dimensional simplex. Each linearly independent asset provides one constraint, so $n - 1$ independent assets suffice.

(iv) Uniqueness follows from the sufficiency axiom: different measures consistent with the same prices are identified. $\square$

6.2 Implications for Market Design

The Fundamental Theorem provides a precise criterion for evaluating market design: a well-designed market should maximize its predictive span. This has concrete implications:

- (i) **Options markets expand predictive completeness:** A market with only stock and bond cannot distinguish between symmetric and skewed predictive distributions. Adding options at multiple strikes fills in the full predictive density.
- (ii) **Prediction markets are maximally predictive:** A prediction market with binary contracts on all events of interest achieves predictive completeness by construction.
- (iii) **Financial innovation as predictive expansion:** New financial instruments (CDOs, variance swaps, volatility-of-volatility products) expand the predictive span, allowing the market to express forecasts along dimensions that were previously unresolvable.

7 Predictive Resolution and Uncertainty Relations

7.1 Quantifying Forecasting Precision

Definition 7.1 (Predictive Information). The *predictive information* of the market at time t about event A at horizon τ is:

$$I_t(A; \tau) = H[\mathcal{P}_0(A)] - H[\mathcal{P}_t(A)] \quad (37)$$

where $H[p] = -p \log p - (1-p) \log(1-p)$ is the binary entropy of the predictive probability. This measures the reduction in uncertainty achieved by the market's prediction relative to the prior.

Definition 7.2 (Predictive Resolution). The *predictive resolution* of a market at time t for horizon τ is the effective number of distinguishable states:

$$\mathcal{R}_t(\tau) = \exp \left(- \sum_{\omega \in \Omega} \mathcal{P}_t(\omega) \log \mathcal{P}_t(\omega) \right) = e^{H[\mathcal{P}_t]} \quad (38)$$

for a discrete state space, or the exponential of the differential entropy for continuous state spaces.

7.2 The Predictive Uncertainty Relation

We now derive a fundamental result: an uncertainty relation that constrains the joint precision of market predictions across time and across states.

Theorem 7.3 (Predictive Uncertainty Relation). *Let Δ_T denote the temporal resolution of the market (the shortest time horizon over which the market makes non-trivial predictions) and let Δ_S denote the state-space resolution (the finest partition of outcomes that the market can distinguish). Then:*

$$\Delta_T \geq \frac{\log_2 \Delta_S}{\mathcal{C}} \quad (39)$$

where $\mathcal{C}$ is the market's information processing capacity, measured in bits per unit time. Equivalently, $\Delta_S \leq 2^{\mathcal{C} \cdot \Delta_T}$.

Proof. The market receives information at rate $\mathcal{C}$ bits per unit time. To resolve Δ_S states, the market needs at least $\log_2 \Delta_S$ bits of information. To achieve this resolution over time horizon Δ_T , the total information budget is $\mathcal{C} \cdot \Delta_T$ bits. The constraint $\log_2 \Delta_S \leq \mathcal{C} \cdot \Delta_T$ rearranges to $\Delta_T \geq \frac{\log_2 \Delta_S}{\mathcal{C}}$. $\square$

Remark 7.4. This is the market analogue of the Heisenberg uncertainty principle in quantum mechanics, where position and momentum cannot both be simultaneously measured with arbitrary precision. In the market context, temporal precision (“when will this happen?”) trades off against state-space precision (“what will happen?”). A market that makes very precise short-term predictions can only do so for a coarse partition of outcomes, while a market that resolves fine distinctions among outcomes can only do so over longer time horizons.

7.3 Consequences of the Uncertainty Relation

Corollary 7.5 (Resolution Frontier). *For a given market capacity $\mathcal{C}$, the achievable pairs (Δ_T, Δ_S) lie on or above the curve $\Delta_T = \frac{\log_2 \Delta_S}{\mathcal{C}}$. The resolution frontier is the set of Pareto-optimal pairs $\Delta_S = 2^{\mathcal{C} \cdot \Delta_T}$.*

Corollary 7.6 (Liquidity and Resolution). *Since market capacity $\mathcal{C}$ is proportional to trading volume (more trades = more information processed per unit time), liquid markets have a lower uncertainty bound and thus better joint resolution. This provides a new explanation for why liquid markets produce better prices: they can simultaneously resolve more states over shorter time horizons.*

Corollary 7.7 (High-Frequency Trading and the Uncertainty Relation). *High-frequency traders operate in the regime of very small Δ_T . The uncertainty relation implies that their predictions must have correspondingly low state-space resolution—they can predict the direction of the next price movement, but not the magnitude. This is consistent with empirical evidence that HFT strategies achieve high Sharpe ratios through small, frequent profits rather than large, rare gains.*

8 Resolution of Classical Puzzles

8.1 The Equity Premium Puzzle

The equity premium puzzle (20) notes that the historical excess return of equities over risk-free assets ($\sim 6\%$ per year) is far too large to be explained by standard consumption-based models with plausible risk aversion parameters. In PMT, this puzzle admits a natural resolution.

Theorem 8.1 (Predictive Resolution of the Equity Premium). *In the PMT framework, the equity premium $\mathbb{E}_{\mathbb{P}}[R_m - r]$ decomposes as:*

$$\mathbb{E}_{\mathbb{P}}[R_m - r] = \underbrace{\mathbb{E}_{\mathbb{P}}[R_m] - \mathbb{E}_{\mathcal{P}}[R_m]}_{\text{predictive bias}} + \underbrace{\mathbb{E}_{\mathcal{P}}[R_m] - r - \text{risk premium}}_{\text{predictive residual}} + \underbrace{\text{risk premium}}_{\text{standard risk compensation}} \quad (40)$$

The key new term is the *predictive bias*: the systematic difference between the physical expected return and the market's predicted expected return. If the market systematically underestimates equity returns (i.e., is too pessimistic), this generates an apparent equity premium even with moderate risk aversion.

Proposition 8.2 (Pessimism Premium). *If the predictive measure systematically underestimates the probability of high-return states:*

$$\mathcal{P}(\omega) < \mathbb{P}(\omega) \quad \text{for states } \omega \text{ with } R_m(\omega) > \bar{R}_m \quad (41)$$

then the predictive bias term is positive, contributing to the equity premium. The magnitude of this contribution is:

$$\text{Pessimism premium} = \mathbb{E}_{\mathbb{P}}[R_m] - \mathbb{E}_{\mathcal{P}}[R_m] = \int R_m(\omega) d(\mathbb{P} - \mathcal{P})(\omega) \quad (42)$$

This is empirically testable: surveys of professional forecasters consistently show that median forecasts for equity returns underperform realized returns, supporting the pessimism hypothesis.

8.2 The Excess Volatility Puzzle

Shiller (24) demonstrated that stock prices are far more volatile than can be justified by subsequent changes in dividends. In PMT, excess volatility has a natural interpretation.

Theorem 8.3 (Excess Volatility as Predictive Updating). *The variance of price changes decomposes as:*

$$\text{Var}(P_{t+1} - P_t) = \underbrace{\text{Var}(\mathbb{E}_{\mathcal{P}_{t+1}}[V] - \mathbb{E}_{\mathcal{P}_t}[V])}_{\text{predictive updating}} + \underbrace{\text{Var}(\Delta R_t)}_{\text{risk preference shifts}} + \underbrace{2\text{Cov}(\Delta \mathcal{P}, \Delta R)}_{\text{interaction}} \quad (43)$$

In this decomposition, the “predictive updating” term captures genuine information-driven price changes, while the “risk preference shifts” and “interaction” terms capture changes in risk attitudes. The excess volatility puzzle arises when the total variance exceeds what the predictive updating term alone would justify—the excess is due to fluctuations in risk preferences (the risk distortion kernel R).

Proposition 8.4 (Volatility as Predictive Disagreement). *In a market with N agents holding heterogeneous predictive measures $\{\mathcal{P}_t^{(i)}\}_{i=1}^N$, the price volatility is bounded below by:*

$$\text{Var}(\Delta P_t) \geq c \cdot \sum_{i=1}^N w_i \text{KL}(\mathcal{P}_t^{(i)} \parallel \mathcal{P}_t) \quad (44)$$

where w_i is agent i ’s market weight and $c > 0$ is a constant. Greater disagreement among predictive measures implies greater volatility.

8.3 The Value of Active Management

A persistent puzzle in finance is why active management persists when index funds consistently outperform. PMT provides a resolution: active managers are part of the market’s predictive infrastructure.

Theorem 8.5 (Predictive Maintenance Cost). *The aggregate cost of active management in equilibrium equals the value of the predictive function it provides:*

$$\sum_{i \in \text{active}} \text{Fee}_i = \mathcal{C}_{\text{prediction}} \cdot \frac{\partial \mathcal{F}}{\partial \mathcal{C}} \quad (45)$$

where $\mathcal{C}_{\text{prediction}}$ is the information processing capacity provided by active managers and $\partial \mathcal{F} / \partial \mathcal{C}$ is the marginal value of prediction improvement.

Active managers are, in effect, paid to run the market’s distributed inference algorithm. The fees they charge are the cost of maintaining the market’s predictive accuracy. If all managers switched to passive indexing, the predictive free energy would increase (predictions would worsen), leading to less efficient capital allocation. The equilibrium fee level is set by the marginal value of the prediction improvement they provide.

9 Connections to Kelly Theory and de Finetti

9.1 The Kelly Criterion as Optimal Prediction Exploitation

The Kelly criterion (17) prescribes the bet size that maximizes the expected log growth rate of wealth:

$$f^* = \arg \max_f \mathbb{E}_{\mathbb{P}}[\log(1 + f \cdot R)] \quad (46)$$

where f is the fraction of wealth bet and R is the return.

In PMT, the Kelly criterion admits a natural interpretation as optimal exploitation of predictive superiority.

Theorem 9.1 (Kelly as Predictive Edge Exploitation). *The optimal Kelly bet size for an agent with predictive measure $\mathcal{P}^{(i)}$ in a market with consensus predictive measure $\mathcal{P}$ is:*

$$f_i^* = \frac{\mathbb{E}_{\mathcal{P}^{(i)}}[R] - \mathbb{E}_{\mathcal{P}}[R]}{\text{Var}_{\mathcal{P}}(R)} + O\left(\text{KL}(\mathcal{P}^{(i)} \parallel \mathcal{P})\right) \quad (47)$$

The optimal bet is proportional to the predictive edge: the difference between the agent's prediction and the market's prediction, normalized by the market's uncertainty.

Proof. Expand the Kelly objective around $f = 0$:

$$\mathbb{E}_{\mathcal{P}^{(i)}}[\log(1 + fR)] \approx f \mathbb{E}_{\mathcal{P}^{(i)}}[R] - \frac{f^2}{2} \mathbb{E}_{\mathcal{P}^{(i)}}[R^2] \quad (48)$$

$$\approx f \mathbb{E}_{\mathcal{P}^{(i)}}[R] - \frac{f^2}{2} \text{Var}_{\mathcal{P}}(R) \quad (49)$$

In equilibrium under the risk-neutral measure $\mathbb{Q}$, $\mathbb{E}_{\mathbb{Q}}[R] = r$. Under the market's predictive measure, $\mathbb{E}_{\mathcal{P}}[R] = r + \rho$ where ρ is the risk premium. The first-order condition gives:

$$f^* = \frac{\mathbb{E}_{\mathcal{P}^{(i)}}[R] - \mathbb{E}_{\mathcal{P}}[R]}{\text{Var}_{\mathcal{P}}(R)} \quad (50)$$

The correction term arises from higher-order contributions proportional to the divergence between the agent's measure and the market measure. $\square$

Corollary 9.2 (Zero Predictive Edge Implies No Trade). *If $\mathcal{P}^{(i)} = \mathcal{P}$ (the agent has no predictive advantage over the market), then $f_i^* = 0$: the agent should not trade. This is the no-trade theorem (21) restated in predictive terms.*

9.2 De Finetti's Theorem and Market Coherence

De Finetti's (9) fundamental insight was that coherent betting behavior—behavior that does not admit a Dutch book—is equivalent to the existence of a probability measure governing the bets. This provides the operational foundation for subjective probability.

PMT extends de Finetti's theorem from individual agents to markets:

Theorem 9.3 (Market Coherence Theorem). *A market is predictively coherent—in the sense that no combination of trades yields a guaranteed loss for all participants—if and only if there exists a predictive measure $\mathcal{P}$ and a positive risk distortion kernel $R > 0$ such that all asset prices satisfy the pricing functional (32).*

Proof. ($\Rightarrow$): If the market is predictively coherent, no arbitrage exists. By the fundamental theorem of asset pricing, a risk-neutral measure $\mathbb{Q}$ exists. Decomposing $\mathbb{Q} = R \cdot \mathcal{P}$ (where $\mathcal{P}$ is the measure closest to $\mathbb{P}$ in KL divergence among all measures consistent with prices) gives the result.

($\Leftarrow$): If $\mathcal{P}$ exists with $R > 0$, then $\mathbb{Q} = R \cdot \mathcal{P}$ is an equivalent martingale measure, which by the first fundamental theorem of asset pricing implies no arbitrage, which implies no Dutch book. $\square$

9.3 Scoring Rules and the Market as a Mechanism

The theory of proper scoring rules (12) provides mechanisms that incentivize truthful probability reporting. A scoring rule $S(p, \omega)$ assigns a score to a probabilistic forecast p when outcome ω is realized.

Proposition 9.4 (Market as a Proper Scoring Mechanism). *The profit-and-loss ($P\mathcal{E}L$) of trading based on a predictive measure $\mathcal{P}^{(i)}$ when the market consensus is $\mathcal{P}$ implements a proper scoring rule:*

$$\mathbb{E}_{\mathbb{P}}[P\mathcal{E}L(\mathcal{P}^{(i)})] = -\text{KL}(\mathbb{P} \parallel \mathcal{P}^{(i)}) + \text{KL}(\mathbb{P} \parallel \mathcal{P}) + \text{const.} \quad (51)$$

The expected $P\mathcal{E}L$ is maximized when $\mathcal{P}^{(i)} = \mathbb{P}$, i.e., when the agent's predictive measure matches the true distribution.

This establishes that the market mechanism itself is a proper scoring rule that incentivizes participants to reveal their best probability estimates through trading. The market doesn't just aggregate predictions—it actively incentivizes truthful prediction.

10 Predictive Dynamics: The Stochastic Evolution of $\mathcal{P}_t$

10.1 The Predictive Process

We now study the dynamics of the predictive measure itself. Rather than analyzing the time series of prices (the traditional approach), we analyze the time series of predictions—the stochastic process $\{\mathcal{P}_t\}_{t \geq 0}$ valued in the space of probability measures on $(\Omega, \mathcal{F})$.

Definition 10.1 (Predictive Process). The *predictive process* is the $\mathcal{M}(\Omega)$ -valued stochastic process $\{\mathcal{P}_t\}_{t \geq 0}$ defined by:

$$\mathcal{P}_t = \mathbb{P}(\cdot \mid \mathcal{F}_t) \quad (52)$$

under the assumption $\mathcal{P} = \mathbb{P}$ (perfect prediction). The predictive process lives on the infinite-dimensional simplex $\mathcal{M}(\Omega)$.

Theorem 10.2 (Doob's Convergence for the Predictive Process). *If the filtration $\{\mathcal{F}_t\}$ is generated by a sequence of informative signals, then the predictive process converges:*

$$\mathcal{P}_t \xrightarrow{t \rightarrow \infty} \delta_{\omega^*} \quad (53)$$

where ω^ is the realized state and δ_{ω^*} is the Dirac measure concentrated on ω^* . That is, in the long run, the market learns the truth.*

Proof. This is Lévy’s zero-one law (or Doob’s martingale convergence theorem) applied to the conditional probabilities $\mathcal{P}_t(A) = \mathbb{P}(A \mid \mathcal{F}_t)$. Since $\mathcal{P}_t(A)$ is a bounded martingale, it converges almost surely. If $\mathcal{F}_\infty = \sigma(\bigcup_t \mathcal{F}_t) = \mathcal{F}$ (the filtration eventually reveals everything), the limit is $\mathbf{1}_A(\omega^*)$. $\square$

10.2 The Stochastic Differential Equation for $\mathcal{P}_t$

In continuous time, the evolution of the predictive measure can be expressed as a stochastic partial differential equation (SPDE) on the space of measures. For the finite-dimensional case (prediction over a finite partition $\{A_1, \dots, A_n\}$), the predictive probabilities $\pi_i(t) = \mathcal{P}_t(A_i)$ satisfy:

Theorem 10.3 (Predictive SDE). *The predictive probabilities evolve according to the system of SDEs:*

$$d\pi_i(t) = \pi_i(t) \sum_{j=1}^m \left(\lambda_j^{(i)}(t) - \bar{\lambda}_j(t) \right) dW_j(t) \quad (54)$$

where $\lambda_j^{(i)}(t)$ is the sensitivity of state i ’s likelihood to the j -th information source, $\bar{\lambda}_j(t) = \sum_i \pi_i(t) \lambda_j^{(i)}(t)$ is the average sensitivity, and $\{W_j\}$ are independent Brownian motions representing information arrivals.

Proof. The result follows from applying the Kushner–Stratonovich equation to the filtering problem. The conditional probability $\pi_i(t)$ of state A_i given continuous observations $\{Y_s\}_{s \leq t}$ (the information process) satisfies the Zakai equation in unnormalized form, which when normalized to a probability measure yields (54). $\square$

Key properties of the predictive SDE:

- (i) **Drift-free:** The predictive SDE has no drift term—the evolution is a pure martingale. This is a direct consequence of Theorem 3.4.
- (ii) **Multiplicative noise:** The volatility of each probability is proportional to the probability itself, ensuring that probabilities near zero or one change slowly (strong beliefs are hard to move).
- (iii) **Conservation:** The probabilities sum to one at all times: $\sum_i d\pi_i(t) = 0$.
- (iv) **Absorption:** The states $\pi_i = 0$ and $\pi_i = 1$ are absorbing boundaries—once the market is certain, it stays certain (under the idealized model).

10.3 Information Arrival Rate and Predictive Volatility

The volatility of the predictive process is directly linked to the rate of information arrival:

Proposition 10.4 (Predictive Volatility–Information Rate Identity). *The rate of quadratic variation of the predictive probability of event A is:*

$$\frac{d\langle \mathcal{P}(\cdot(A)) \rangle_t}{dt} = \mathcal{P}_t(A)^2 (1 - \mathcal{P}_t(A))^2 \cdot \sum_{j=1}^m (\lambda_j^{(A)} - \bar{\lambda}_j)^2 \quad (55)$$

where $\langle \mathcal{P}(A) \rangle_t$ denotes the quadratic variation process. Note that the paths of $\mathcal{P}_t(A)$, driven by Brownian motion, are nowhere differentiable; the quadratic variation captures the instantaneous variability. The right-hand side is maximized when $\mathcal{P}_t(A) = 1/2$ (maximum uncertainty about A) and when the signal-to-noise ratio $(\lambda_j^{(A)} - \bar{\lambda}_j)^2$ is large (highly informative signals).

This result quantifies the intuition that predictions are most volatile when the market is most uncertain—a probability near 50% can swing dramatically in either direction, while a probability near 0% or 100% is inherently stable.

11 Quantifying Market Prediction Accuracy

11.1 Calibration and Discrimination

Following the scoring rule literature (12), we decompose predictive accuracy into two components:

Definition 11.1 (Market Calibration). The market is *calibrated* if, among all events to which it assigns probability p , approximately a fraction p actually occur:

$$\lim_{n \rightarrow \infty} \frac{1}{|\{t : \mathcal{P}_t(A_t) \in [p - \epsilon, p + \epsilon]\}|} \sum_{\substack{t: \mathcal{P}_t(A_t) \in \\ [p - \epsilon, p + \epsilon]}} \mathbf{1}_{A_t}(\omega_t) = p \quad \text{for all } p \in [0, 1] \quad (56)$$

Definition 11.2 (Market Discrimination). The market has good *discrimination* (or resolution) if it assigns high probabilities to events that occur and low probabilities to events that do not:

$$\mathbb{E}[\mathcal{P}_t(A) \mid A \text{ occurs}] \gg \mathbb{E}[\mathcal{P}_t(A) \mid A \text{ does not occur}] \quad (57)$$

Theorem 11.3 (Brier Score Decomposition for Markets). *The market's Brier score (mean squared prediction error) decomposes as:*

$$BS = \underbrace{\overline{p(1-p)}}_{\text{irreducible uncertainty}} + \underbrace{\text{CalErr}}_{\text{calibration error}} - \underbrace{\text{Disc}}_{\text{discrimination}} \quad (58)$$

The Brier score is minimized when calibration error is zero and discrimination is maximized.

11.2 The Predictive Efficiency Ratio

Definition 11.4 (Predictive Efficiency). The *predictive efficiency* of a market is:

$$\eta_P = 1 - \frac{\text{KL}(\mathbb{P} \parallel \mathcal{P})}{\text{KL}(\mathbb{P} \parallel \pi_0)} \quad (59)$$

where π_0 is the prior (uninformed) distribution. This ratio measures the fraction of the total possible information gain that the market has actually achieved. A perfect market has $\eta_P = 1$; a market that adds no information to the prior has $\eta_P = 0$.

Proposition 11.5 (Bounds on Predictive Efficiency). *Under the PMT axioms:*

$$0 \leq \eta_P \leq 1 \quad (60)$$

with the lower bound achieved when $\mathcal{P} = \pi_0$ (no information aggregation) and the upper bound when $\mathcal{P} = \mathbb{P}$ (perfect prediction). Moreover, η_P is non-decreasing in the number of informed agents N and in the liquidity L of the market:

$$\frac{\partial \eta_P}{\partial N} \geq 0, \quad \frac{\partial \eta_P}{\partial L} \geq 0 \quad (61)$$

11.3 Comparing Predictive Efficiency Across Markets

The predictive efficiency framework provides a principled way to compare markets. An empirically testable conjecture:

Conjecture 11.6 (Market Predictive Efficiency Hierarchy). *Predictive efficiency varies systematically across markets:*

$$\eta_P^{\text{large-cap equities}} > \eta_P^{\text{corporate bonds}} > \eta_P^{\text{real estate}} > \eta_P^{\text{private equity}} \quad (62)$$

with the ordering driven primarily by liquidity and the number of informed participants.

12 Predictive Market Microstructure

12.1 The Predictive Interpretation of Order Flow

In the PMT framework, every trade carries a predictive signal. When agent i buys, she is expressing the belief $\mathcal{P}^{(i)}(V > P) > \mathcal{P}(V > P)$ —her prediction assigns higher probability to favorable states than the market consensus.

Definition 12.1 (Predictive Content of a Trade). The predictive content of a trade of size q at price P is:

$$\mathcal{I}_{\text{trade}} = q \cdot \left| \frac{d\mathcal{P}_t}{dP} \right|^{-1} \approx q \cdot \sigma_{\mathcal{P}} \quad (63)$$

where $\sigma_{\mathcal{P}}$ is the standard deviation of the predictive distribution. Large trades in uncertain environments carry maximum predictive content.

Theorem 12.2 (Kyle’s Lambda as Predictive Sensitivity). *Kyle’s lambda (18)—the price impact coefficient—is the reciprocal of the market’s predictive precision:*

$$\lambda = \frac{\sigma_V}{\sigma_z} \cdot \frac{1}{\tau_{\text{total}}} \quad (64)$$

where σ_V is the prior standard deviation of fundamental value, σ_z is noise trading volatility, and τ_{total} is the total precision of all informed agents’ signals. Markets with higher aggregate predictive precision have lower price impact.

12.2 The Bid–Ask Spread as Predictive Insurance

The bid–ask spread can be understood as the price the market charges for providing instant predictive feedback:

Proposition 12.3 (Spread as Ambiguity Premium). *The bid–ask spread decomposes as:*

$$S = 2 \cdot \underbrace{\mathbb{E}[\mathcal{P}_{\text{post-trade}} - \mathcal{P}_{\text{pre-trade}}]}_{\text{expected predictive revision}} + \underbrace{\text{inventory cost}}_{\text{risk bearing}} + \underbrace{\text{processing cost}}_{\text{operational}} \quad (65)$$

The first term—the expected magnitude of the predictive revision caused by the trade—is the informational component of the spread and dominates in liquid markets.

12.3 Market Making as Predictive Mediation

Market makers occupy a unique role in the PMT framework: they are *predictive mediators* who bridge temporal gaps in the flow of predictions.

Proposition 12.4 (Market Maker as Predictive Smoother). *A market maker with inventory I_t and predictive measure $\mathcal{P}_t^{(MM)}$ optimally sets quotes:*

$$P_{bid} = \mathbb{E}_{\mathcal{P}^{(MM)}}[V] - \frac{1}{2}\gamma I_t \sigma^2 - \frac{\lambda}{2} \quad (66)$$

$$P_{ask} = \mathbb{E}_{\mathcal{P}^{(MM)}}[V] + \frac{1}{2}\gamma I_t \sigma^2 + \frac{\lambda}{2} \quad (67)$$

where γ is risk aversion, σ^2 is predictive variance, and λ is the adverse selection parameter. The quotes are centered on the market maker’s prediction, adjusted for inventory risk and the cost of trading against better-informed agents.

13 Empirical Predictions and Experimental Design

13.1 Core Empirical Predictions

1. **Calibration Test:** Option-implied probabilities, adjusted for risk preferences using the power utility inversion (17), should be well-calibrated: the fraction of events occurring at each probability level should match the predicted probability. This can be tested using large cross-sections of options data.
2. **Predictive Martingale Test:** The implied probability $\mathcal{P}_t(A)$ for any event A should be a martingale under the physical measure $\mathbb{P}$. Specifically, $\mathcal{P}_{t+1}(A) - \mathcal{P}_t(A)$ should be uncorrelated with $\mathcal{F}_t$. Violations indicate failures of Axiom 2.
3. **Free Energy Monotonicity:** The predictive free energy $\mathcal{F}[\mathcal{P}_t]$ should be non-increasing over time within a given prediction episode (e.g., from the start of an earnings season to the announcement). The rate of decrease should correlate with information flow (proxied by trading volume).

4. **Uncertainty Relation Test:** For a given market, plot Δ_T (time horizon of profitable prediction, measured by holding period of successful strategies) against Δ_S (number of states distinguished, measured by the effective number of scenarios in option-implied distributions). The data should satisfy $\Delta_T \geq \frac{\log_2 \Delta_S}{c}$ as predicted by Theorem 7.3.
5. **Predictive Efficiency Comparison:** Compare the Brier scores of market-implied probabilities across different asset classes. The ordering should follow the hierarchy predicted by Conjecture 11.6, with liquid equity markets outperforming illiquid alternatives.

13.2 Experimental Design: A Controlled Test of PMT

We propose the following experimental framework for testing PMT:

- Step 1:** Identify a universe of well-defined, verifiable events (e.g., quarterly earnings beating consensus, central bank decisions, election outcomes).
- Step 2:** Extract $\mathbb{Q}$ -probabilities from option prices using the Breeden–Litzenberger formula (7).
- Step 3:** Estimate the risk distortion kernel R using either Method 1 (prediction market calibration) or Method 2 (power utility inversion).
- Step 4:** Compute $\mathcal{P} = \mathbb{Q}/R$ and evaluate calibration, discrimination, and Brier scores.
- Step 5:** Compare PMT-derived $\mathcal{P}$ against: (a) raw option-implied $\mathbb{Q}$ -probabilities, (b) analyst consensus forecasts, (c) pure prediction market prices.

The hypothesis is that $\mathcal{P}$ (the risk-adjusted market probability) outperforms $\mathbb{Q}$ (the unadjusted option-implied probability) in calibration and discrimination, as $\mathcal{P}$ removes the systematic risk distortion that makes $\mathbb{Q}$ a biased probability estimator.

13.3 Data Requirements and Feasibility

Testing the full PMT framework requires:

- High-frequency option chain data (to extract $\mathbb{Q}$).
- Contemporaneous prediction market data (for calibration of R).
- Realized outcomes for a large sample of events.
- Trading volume and order flow data (for microstructure tests).

The richest testing ground is the period around binary events (earnings announcements, FDA decisions, elections) where both option markets and prediction markets provide probabilities.

14 Conclusion

We have introduced Predictive Market Theory, an axiomatic framework that reconceptualizes financial markets as prediction engines whose fundamental output is a probability measure over future states of the world. The theory rests on five axioms—Predictive Completeness, Bayesian Coherence, Informational Irreversibility, Predictive Sufficiency of Price, and Thermodynamic Dissipation—from which the standard results of asset pricing theory emerge as corollaries.

The key theoretical innovations are:

- (i) **The Predictive Measure $\mathcal{P}$:** A third probability measure, distinct from both the physical $\mathbb{P}$ and risk-neutral $\mathbb{Q}$ measures, encoding the market’s collective forecast purged of risk distortions.
- (ii) **The Variational Prediction Principle:** The market minimizes predictive free energy, unifying Bayesian inference, information theory, and thermodynamics within a single framework.
- (iii) **Price as derived from prediction:** Rather than deriving prediction from price, PMT derives price from prediction through the pricing functional $P = e^{-r\tau} \mathbb{E}_{\mathcal{P}}[X \cdot R]$.
- (iv) **The Predictive Uncertainty Relation:** A fundamental limit on the joint precision of temporal and state-space prediction, $\Delta_T \geq \frac{\log_2 \Delta_S}{c}$.
- (v) **Resolution of classical puzzles:** The equity premium, excess volatility, and persistence of active management receive natural explanations as consequences of the predictive structure.
- (vi) **Market coherence as an extension of de Finetti:** The market’s predictive function satisfies the same coherence conditions as individual rational probability, elevated to the collective level.

Predictive Market Theory shifts the conceptual center of gravity in financial economics from *pricing* to *prediction*. In this framework, the primary function of markets is not to “find the right price” but to “find the right probability.” Prices are instrumental—they are the medium through which the predictive function operates—but the prediction is the substance. This inversion suggests that the true measure of market quality is not price efficiency in the Fama sense but *predictive efficiency*: how well does the market’s probability measure approximate the true distribution of future outcomes?

We believe this framework opens a productive new chapter in the theory of financial markets, one that more naturally connects to adjacent fields—machine learning, information theory, statistical mechanics, and decision theory—and that provides sharper tools for both theoretical analysis and empirical investigation.

References

- [1] Arrow, K. J., & Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica*, 22(3), 265–290.
- [2] Arrow, K. J. (1964). The role of securities in the optimal allocation of risk-bearing. *Review of Economic Studies*, 31(2), 91–96.
- [3] Arrow, K. J., et al. (2008). The promise of prediction markets. *Science*, 320(5878), 877–878.
- [4] Bachelier, L. (1900). Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure*, 17, 21–86.
- [5] Blei, D. M., Kucukelbir, A., & McAuliffe, J. D. (2017). Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518), 859–877.
- [6] Breeden, D. T., & Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51(4), 621–651.
- [7] Cover, T. M., & Thomas, J. A. (2006). *Elements of Information Theory* (2nd ed.). Wiley.
- [8] Debreu, G. (1959). *Theory of Value*. Yale University Press.
- [9] de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 7(1), 1–68.
- [10] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2), 383–417.
- [11] Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.
- [12] Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359–378.
- [13] Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, 70(3), 393–408.
- [14] Hanson, R. (2003). Combinatorial information market design. *Information Systems Frontiers*, 5(1), 107–119.
- [15] Harrison, J. M., & Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11(3), 215–260.
- [16] Hayek, F. A. (1945). The use of knowledge in society. *American Economic Review*, 35(4), 519–530.
- [17] Kelly, J. L. (1956). A new interpretation of information rate. *Bell System Technical Journal*, 35(4), 917–926.

- [18] Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6), 1315–1335.
- [19] Long, M. (2026). Markets as distributed Bayesian inference engines: Price formation, information aggregation, and the thermodynamic structure of financial systems. *GrokRxiv:2603.04.markets-bayesian-inference*.
- [20] Mehra, R., & Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of Monetary Economics*, 15(2), 145–161.
- [21] Milgrom, P., & Stokey, N. (1982). Information, trade and common knowledge. *Journal of Economic Theory*, 26(1), 17–27.
- [22] Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1), 58–73.
- [23] Savage, L. J. (1954). *The Foundations of Statistics*. Wiley.
- [24] Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71(3), 421–436.
- [25] Wolfers, J., & Zitzewitz, E. (2004). Prediction markets. *Journal of Economic Perspectives*, 18(2), 107–126.