A Bayesian HJB Framework for Robust Monetary Policy under Structural Uncertainty

Peilin Rao

November 3, 2025

Abstract

Deep structural uncertainty, exemplified by the unstable Phillips Curve, undermines single-model monetary policy. Existing paradigms are either brittle to misspecification (Certainty-Equivalence) or overly pessimistic and static (Worst-Case Robustness). This research proposes a Bayesian Hamilton-Jacobi-Bellman (HJB) framework that optimizes policy over a belief-weighted 'cloud' of models, which formalizes the dual control problem to balance stabilization with active learning. I solve this historically intractable problem using the Deep Backward Stochastic Differential Equation (Deep BSDE) method, overcoming the curse of dimensionality. The result is a tractable, adaptive policy demonstrably more robust to structural breaks than existing static approaches.

1 Motivation, Literature, and Contribution

1.1 The Central Banker's Dilemma Amidst a Faltering Phillips Curve

For several decades, central banks have navigated deep structural uncertainty regarding the true data-generating process of the economy, as exemplified by the unstable Phillips Curve. This is not only a data problem but also a core identification problem: Qu and Tkachenko (2017, 2023) demonstrate that structurally different models can be observationally equivalent, creating situations where the central bank is uncertain about the "true" model. The identification problem has significant implications for policy analysis, as the design of tools like forward guidance depends heavily on the underlying economic structure (Roulleau-Pasdeloup, 2020; Nie and Roulleau-Pasdeloup, 2023).

This motivates the core question of this proposal: when presented with a set of plausible and observationally equivalent models, how should a central bank make optimal policy decisions?

1.2 Existing Approaches and Their Limitations: Brittleness and Pessimism

Two dominant paradigms for policymakers operating under uncertainty emerge from the literature. The first paradigm is Certainty-Equivalence (CE), which chooses the optimal policy based on a single fixed best-fitting model. While simple, CE is brittle to model misspecification. The second paradigm is the robust control framework (Hansen and Sargent, 2008), which formalizes policy as a min-max game to combat model misspecification. It is still a static framework that focuses only on the worst-case scenario and may produce overly pessimistic policies; it defends against structural uncertainty but does not actively resolve it. A recent clarification by Hansen and Sargent (2023) themselves motivates our approach. Unbundling uncertainty into likelihood misspecification and prior uncertainty over a discrete set of models, they point out the need for a framework with active learning to tackle the latter uncertainty.

1.3 An Ideal but Unattainable Alternative: The Theory of Dual Control

In theory, a solution to this learning problem is the dual control theory (Feldbaum, 1961), which formalizes the trade-off between exploitation of current knowledge and exploration to reduce uncertainty. Its mathematical construct, a HJB equation over the augmented belief-state space (Kushner, 1967; Fleming and Rishel, 1975), is well established, but the curse of dimensionality (Bellman, 1957) has rendered it computationally intractable for decades, demonstrating a gap between the optimal theoretical policy and practical applications.

1.4 Research Question and Contribution: A Computationally Tractable Solution

This research addresses the gap with the question: How can central banks design optimal monetary policy that is robust against structural uncertainty, while learning about the underlying nature of the economy? Our contribution is to provide the first computationally feasible solution to this continuous-time Bayesian dual control problem in a canonical macroeconomic setting.

- **Methodological:** We use the Deep Backward Stochastic Differential Equation (Deep BSDE) method, a recent breakthrough in machine learning (Han et al., 2018) to solve the high-dimensional HJB equation. By casting the dynamic HJB problem into a BSDE framework, the resulting mesh-free method resolves the curse of dimensionality.

- **Substantive:** By solving the optimal adaptive policy, the study quantifies the "dual control premium," the economic value of policy experimentation, and presents a new normative benchmark for monetary policy. We will illustrate through simulation that this adaptive policy is superior to both the fragile CE approach and the pessimistic Worst-Case Robustness approach, especially when the environment experiences sudden structural breaks.

we aim to provide an effective computational toolkit to upgrade from a static policy framework, demonstrating how to design monetary policy that optimally balances the need to stabilize today with the need to gain knowledge about the underlying complex and evolving economic structure.

2 The Proposed Framework: A Bayesian HJB Equation

This section formally develops the continuous-time framework for deriving an optimal and robust monetary policy under structural uncertainty. The central bank's problem is cast as a stochastic optimal control problem where the state vector is augmented to include the bank's evolving beliefs about the true model of the economy. The solution is characterized by a Hamilton-Jacobi-Bellman (HJB) equation that endogenously balances the competing objectives of economic stabilization and active learning.

2.1 The Economic Environment and Objective Function

The observable economic environment is described by the state vector $\mathbf{K}_t = [\pi_t, x_t]^T$, representing inflation and the output gap. The central bank exercises control via a single instrument, the nominal interest rate i_t.

The central bank's objective is to select a policy path $\{i_s\}_{s=t}^{\infty}$ that maximizes the expected discounted sum of future flow utility. The instantaneous flow utility function, $u(\mathbf{K}_t)$, is the negative of a standard quadratic loss function:

$$u(\mathbf{K}_t) = -L(\mathbf{K}_t) = -(\pi_t^2 + \omega_x x_t^2)$$

where $\omega_x > 0$ is the constant relative weight on output gap stabilization and $\rho > 0$ is the discount rate.

The value function, $V(\mathbf{K}_t, \mathbf{p}_t)$, represents the maximized total expected discounted utility from time t onward. Crucially, it depends not only on the current economic state $\mathbf{K}_t$ but also on the central bank's belief state $\mathbf{p}_t$, which is a probability distribution over the set of possible economic models. The value function is thus defined as:

$$V(\mathbf{K}_t, \mathbf{p}_t) = \max_{\{i_s\}} \mathbb{E}_t \left[\int_t^\infty e^{-\rho(s-t)} u(\mathbf{K}_s) ds \right]$$

The full state for the policymaker is the augmented vector $(\mathbf{K}_t, \mathbf{p}_t)$, and the solution to this problem is a policy function $i^*(\mathbf{K}_t, \mathbf{p}_t)$ that maps this augmented state to an optimal policy action.

2.2 A "Cloud" of Models and Bayesian Learning Dynamics

Structural uncertainty is captured by a discrete set, or "cloud," of M candidate models, $\Theta = \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_M\}$. Each model $\boldsymbol{\theta}_m$ specifies a distinct law of motion for the economic state vector $\mathbf{K}_t = [\pi_t, x_t]^T$. The dynamics under any given model are governed by a vector Stochastic Differential Equation (SDE):

$$d\mathbf{K}_t = \boldsymbol{\mu}_{\boldsymbol{\theta}_m}(\mathbf{K}_t, i_t)dt + \boldsymbol{\Sigma} d\mathbf{B}_t$$

where $\mathbf{B}_t$ is a 2-dimensional standard Brownian motion and the diffusion matrix $\boldsymbol{\Sigma}$ is assumed, for simplicity, to be known and constant across models.

In this research, the model uncertainty is centered on the true specification of the Phillips Curve. The law of motion for the output gap (the IS curve) is assumed to be known and common across all models, following a mean-reverting Ornstein-Uhlenbeck process:

- **Output Gap (IS Curve):**

$$dx_t = \underbrace{\lambda_x(-x_t - \gamma(i_t - \pi_t))}_{\mu_x(\mathbf{K}_t, i_t)} dt + \sigma_x d\mathbf{B}_t^x$$

The model-specific component, $\boldsymbol{\mu}_{\pi,\boldsymbol{\theta}_m}$, defines the inflation dynamics according to two canonical forms:

- **Model Type 1: Backward-Looking Phillips Curve (BL)**

$$d\pi_t = \underbrace{\lambda_\pi(\kappa_{BL,m}x_t - \pi_t)}_{\mu_{\pi,\boldsymbol{\theta}_m}(\mathbf{K}_t)} dt + \sigma_\pi d\mathbf{B}_t^\pi$$

  Here, inflation dynamics are adaptive, mean-reverting to a level determined by the output gap.

- **Model Type 2: New Keynesian Phillips Curve (NK)**

$$d\pi_t = \underbrace{(\rho_{NK,m}\pi_t + \kappa_{NK,m}x_t)}_{\mu_{\pi,\boldsymbol{\theta}_m}(\mathbf{K}_t)} dt + \sigma_\pi d\mathbf{B}_t^\pi$$

  Here, inflation exhibits persistence ($\rho_{NK,m}$) and is driven by the output gap, reflecting forward-looking price-setting behavior.

A "Flat" Phillips Curve is not a distinct structural model but rather a specific parameterization of the models described above where the slope coefficient is near zero ($\kappa_{BL,m} \approx 0$ or $\kappa_{NK,m} \approx 0$). The policymaker's learning problem thus involves uncertainty over both the structural form (BL vs. NK) and the key slope parameter (κ). The specific parameter values for "steep" and "flat" slopes will be calibrated to match key empirical moments and stylized facts from the post-war U.S. economy, ensuring the model space is empirically relevant. The "cloud" of models Θ is therefore explicitly instantiated by calibrating these two structures with different parameter values. For example, a simple yet powerful four-model cloud would be:

- $\boldsymbol{\theta}_1$: **Backward-Looking, Steep Slope** (e.g., $\kappa_{BL,1}$ = high value)
- $\boldsymbol{\theta}_2$: **Backward-Looking, Flat Slope** (e.g., $\kappa_{BL,2} \approx 0$)
- $\boldsymbol{\theta}_3$: **New Keynesian, Steep Slope** (e.g., $\kappa_{NK,3}$ = high value)
- $\boldsymbol{\theta}_4$: **New Keynesian, Flat Slope** (e.g., $\kappa_{NK,4} \approx 0$)

The central bank's belief state is the probability vector $\mathbf{p}_t = [p_{1,t}, \ldots, p_{M,t}]^T$ on the $(M-1)$-simplex. Beliefs are updated via Bayesian filtering. The belief-weighted average drift of the economy is $\bar{\boldsymbol{\mu}}(\mathbf{K}_t, i_t, \mathbf{p}_t) = \sum_{m=1}^{M} p_{m,t}\boldsymbol{\mu}_{\boldsymbol{\theta}_m}(\mathbf{K}_t, i_t)$. The innovation process, representing the unpredictable component of the data from the perspective of the policymaker, is a standard martingale defined as:

$$d\mathbf{Z}_t = \mathbf{\Sigma}^{-1}(d\mathbf{K}_t - \bar{\boldsymbol{\mu}}_t dt)$$

The evolution of the belief vector $\mathbf{p}_t$ is governed by the Kushner-Stratonovich equation, the continuous-time analogue of Bayes' rule:

$$dp_{m,t} = p_{m,t} \left(\boldsymbol{h}_{\boldsymbol{\theta}_m}(\mathbf{K}_t, i_t) - \bar{\boldsymbol{h}}_t \right)^T d\mathbf{Z}_t$$

where $\boldsymbol{h}_{\boldsymbol{\theta}_m} \equiv \mathbf{\Sigma}^{-1}\boldsymbol{\mu}_{\boldsymbol{\theta}_m}$ is the model-specific "signal" and $\bar{\boldsymbol{h}}_t = \sum_m p_{m,t}\boldsymbol{h}_{\boldsymbol{\theta}_m}$ is the belief-weighted average signal. The term $(\boldsymbol{h}_{\boldsymbol{\theta}_m} - \bar{\boldsymbol{h}}_t)$ represents the prediction error of model m relative to the average belief.

2.3 The Bayesian Hamilton-Jacobi-Bellman Equation

The value function $V(\mathbf{K}, \mathbf{p})$ must satisfy the Hamilton-Jacobi-Bellman (HJB) equation that characterizes optimality. Applying the multidimensional Itô's Lemma to $V(\mathbf{K}_t, \mathbf{p}_t)$ for the joint process yields the HJB equation:

$$\rho V(\mathbf{K}, \mathbf{p}) = \max_{i_t} \left\{ u(\mathbf{K}) + \mathcal{L}V(\mathbf{K}, \mathbf{p}, i_t) \right\}$$

where $\mathcal{L}V$ is the generator of the augmented state process $(\mathbf{K}_t, \mathbf{p}_t)$. This generator can be decomposed to reveal the economic forces at play. The full generator is given by:

$$\mathcal{L}V = \bar{\boldsymbol{\mu}}^T V_{\mathbf{K}} + \frac{1}{2}\mathrm{Tr}\left(\mathbf{\Sigma}\mathbf{\Sigma}^T V_{\mathbf{KK}}\right) + \frac{1}{2}\mathrm{Tr}\left(\mathbf{\Gamma}(\mathbf{p}) V_{\mathbf{pp}}\right) + \mathrm{Tr}\left(\mathbf{C}(\mathbf{p}) V_{\mathbf{pK}}\right)$$

where $V_{\mathbf{K}}$ and $V_{\mathbf{p}}$ are gradients, $V_{\mathbf{KK}}$ and $V_{\mathbf{pp}}$ are Hessians, and $V_{\mathbf{pK}}$ is the matrix of cross-partial derivatives. The quadratic variation matrices are derived from the SDEs for $\mathbf{K}_t$ and $\mathbf{p}_t$:

- The belief covariance matrix $\mathbf{\Gamma}(\mathbf{p}) \equiv \mathbb{E}_t[d\mathbf{p}_t d\mathbf{p}_t^T]/dt$ is an $M \times M$ matrix representing the variance of belief updates. Its (i, j)-th element is derived from the quadratic covariation of the belief SDEs:

$$\mathbf{\Gamma}(\mathbf{p})_{ij} = p_i p_j (\boldsymbol{h}_{\boldsymbol{\theta}_i} - \bar{\boldsymbol{h}})^T (\boldsymbol{h}_{\boldsymbol{\theta}_j} - \bar{\boldsymbol{h}})$$

  In matrix notation, this can be expressed compactly. Let $\tilde{\mathbf{H}}$ be the $M \times 2$ matrix whose m-th row is the relative signal vector $(\boldsymbol{h}_{\boldsymbol{\theta}_m} - \bar{\boldsymbol{h}})^T$. Then the full covariance matrix is given by:

$$\boldsymbol{\Gamma}(\mathbf{p}) = \mathrm{diag}(\mathbf{p})\, \tilde{\mathbf{H}}\tilde{\mathbf{H}}^T\, \mathrm{diag}(\mathbf{p})$$

- The state-belief covariance matrix $\mathbf{C}(\mathbf{p}) \equiv \mathbb{E}_t[d\mathbf{K}_t d\mathbf{p}_t^T]/dt$ is a $2 \times M$ matrix:

$$\mathbf{C}(\mathbf{p}) = \boldsymbol{\Sigma}\left[p_1(\boldsymbol{h}_{\boldsymbol{\theta}_1} - \bar{\boldsymbol{h}}), \ldots, p_M(\boldsymbol{h}_{\boldsymbol{\theta}_M} - \bar{\boldsymbol{h}})\right]$$

Substituting this into the HJB equation and grouping terms by their economic function gives:

$$\rho V = \max_{i_t}\left\{ u(\mathbf{K}) + \underbrace{\bar{\boldsymbol{\mu}}^T V_{\mathbf{K}} + \frac{1}{2}\mathrm{Tr}\left(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T V_{\mathbf{KK}}\right)}_{\text{Term A: Exploitation}} + \underbrace{\frac{1}{2}\mathrm{Tr}\left(\boldsymbol{\Gamma}(\mathbf{p})V_{\mathbf{pp}}\right) + \mathrm{Tr}\left(\mathbf{C}(\mathbf{p})V_{\mathbf{pK}}\right)}_{\text{Term B: Exploration}} \right\}$$

- **Exploitation Term (A):** This term represents the expected change in value from the evolution of the economic state $\mathbf{K}_t$, conditional on current beliefs $\mathbf{p}_t$. It captures the myopic incentive to set the interest rate i_t to stabilize the economy according to the belief-weighted average model, $\bar{\boldsymbol{\mu}}$.

- **Exploration Term (B):** This term captures the change in value arising purely from the evolution of the belief state $\mathbf{p}_t$. This is the mathematical formalization of the dual control premium. The control i_t influences this term because it alters the drift functions $\boldsymbol{\mu}_{\boldsymbol{\theta}_m}$, which change the signals $\boldsymbol{h}_{\boldsymbol{\theta}_m}$ and thus the volatility and covariance of belief updates. A forward-looking policymaker may choose an i_t that deviates from the myopically optimal action (which only considers Term A) to "experiment" or probe the system. Such experimentation can increase the variance of belief updates ($\boldsymbol{\Gamma}(\mathbf{p})$), accelerating learning and leading to better future policy. The solution to this HJB equation, $i^*(\mathbf{K}_t, \mathbf{p}_t)$, is a policy that optimally balances the immediate goal of stabilization with the long-term value of learning.

3 Computational Strategy and Validation Plan

This section describes the method to solve the Bayesian HJB equation and the planned experimentation to validate the resulting policy function. This strategy addresses the computational bottleneck associated with dual control theory and empirically validates the performance of the adaptive policy function against appropriate benchmarks.

3.1 The Curse of Dimensionality in the Belief State

The primary problem that has historically rendered continuous-time dual control problems computationally intractable is the "curse of dimensionality." The policymaker's entire state is given by the augmented vector $\mathbf{X}_t = [\mathbf{K}_t^T, \mathbf{p}_t^T]^T$, where the economic state $\mathbf{K}_t$ is in $\mathbb{R}^2$ and the belief state $\mathbf{p}_t \in \Delta^{M-1}$ is on the $(M-1)$-dimensional simplex. The resulting state is therefore $1+M$ dimensional.

Traditional numerical approaches, such as finite difference schemes, solve HJB equations by discretizing the continuous state space with a grid, and the number of grid points increases exponentially with the number of dimensions. For a problem characterized by even a small cloud of $M = 4$ or $M = 5$ models, the memory and computation required by a grid method would become impractical. This computational barrier has stalled advances in obtaining numerical solutions to this important class of macroeconomic problems until recently. Our research addresses this barrier by changing the paradigm from approximating discrete values of the value function on the grid points to approximating the entire value function directly with a neural network function approximator.

3.2 Neural Network Approach for Solution

This research will implement a cutting-edge numerical approach called the Deep Backward Stochastic Differential Equation (Deep BSDE) method to solve the high-dimensional Bayesian HJB equation. The proposed method recasts the HJB equation into a format that fits naturally with modern deep learning methods.

The theoretical basis for this approach derives from the nonlinear Feynman-Kac theorem, which establishes a correspondence between a fully nonlinear PDE, such as the Bayesian HJB, and a system of Forward-Backward Stochastic Differential Equations (FBSDEs). The HJB equation derived in Section 2 corresponds to the following system of FBSDEs.

- **Forward SDE:** The augmented state process $\mathbf{X}_t$ evolves forward in time according to its derived law of motion, driven by the innovation process $\mathbf{Z}_t$:

$$d\mathbf{X}_t = \boldsymbol{\mu}_{\boldsymbol{X}}(\mathbf{X}_t, i_t)dt + \boldsymbol{\sigma}_{\boldsymbol{X}}(\mathbf{X}_t, i_t)d\mathbf{Z}_t, \quad \text{with a given initial state } \mathbf{X}_0$$

- **Backward SDE:** Two processes, the value process $Y_t = V(\mathbf{X}_t)$ and a gradient-related process $Z_t \in \mathbb{R}^{1\times 2}$, evolve backward in time:

$$-dY_t = f(\mathbf{X}_t, Y_t, Z_t, i_t^*)dt - Z_t^T d\mathbf{Z}_t$$

  where the "driver" function f is given by $f(\mathbf{X}_t, Y_t, Z_t, i_t^*) = u(\mathbf{K}_t, i_t^*) - \rho Y_t$, and i_t^* is the optimal policy.

The Deep BSDE algorithm works by discretizing the FBSDE system in time and leveraging deep neural networks to approximate the unknown functions at each time step, including the optimal policy i_t^* and the gradient term Z_t. The network parameters are trained by minimizing the difference between the initial value Y_0 and the terminal condition of the BSDE.

The key benefit of the deep BSDE method for this particular problem is that it avoids calculating the second-order derivatives (Hessians) of the value function ($V_{\mathbf{KK}}$, $V_{\mathbf{pp}}$, and $V_{\mathbf{pK}}$) directly. The Hessian terms in the HJB generator are hard to compute accurately and stably through automatic differentiation in higher dimensions. By projecting the HJB problem onto the BSDE framework, the algorithm can learn the necessary components from forward simulations, without forming those unstable terms, which provides a robust, more computationally efficient solution.

The training procedure will be simulation-based and constructed within the JAX framework to harness its high-performance numerical computing toolkits. The training loop will generate mini-batches of state trajectories at each iteration by simulating the forward SDE under the present policy approximation. These trajectories generate the training data for the neural networks, which helps focus the learning on a progressive approximation of the stationary distribution of the state process, the most economically relevant regions in the state space. The neural network parameters $\boldsymbol{\Theta}$ will be updated using the Adam optimizer from JAX's Optax library. This unified JAX framework is intentional: its JIT compilation is well-suited for speeding up the computationally heavy forward simulations, while its functional design enables efficient and robust gradient calculations for optimization.

3.3 Validation via Simulation Experiments

In order to quantify the economic value of the Bayesian HJB framework, the performance of the adaptive policy function $i^*(\mathbf{K}_t, \mathbf{p}_t)$ from the Bayesian HJB framework will be compared with two standard, non-learning policy frameworks through simulation experiments. The experiments are:

- **Certainty Equivalence (CE):** This agent knows there are multiple models in the cloud, but assumes from the start that the model with the greatest initial prior probability is the true model. The policy model does not change thereafter, nor does it adapt to new data.

- **Worst-Case Robustness (H-S):** This agent is, in the spirit of Hansen and Sargent, ambiguity averse, and sets the policy by optimizing against an adversarial "nature/god" that can choose any model from the cloud Θ to minimize the agent's utility.

Experiment 1: Performance under a Fixed (Hybrid) True Model.

In this experiment, we will evaluate performance in a static economic environment, where the true data-generating process (DGP) is a fixed linear combination of the models in the agent's cloud (e.g., 30% Backward-Looking, 70% New Keynesian). This setup is realistic as it suggests the true economy is more complex than any single model can capture. The lifetime utility (negative of discounted quadratic loss) will be determined for each of the three agents after thousands of simulations from this true DGP.

- **Hypothesis:** The Bayesian HJB agent will exhibit significantly greater average lifetime utility compared to both benchmark agents. It will outperform the CE agent by correctly learning that the CE agent's initial choice of a single model is misspecified, and it will outperform the H-S agent by not being overly pessimistic about a worst case that never happens.

Experiment 2: Robustness to Unforeseen Structural Breaks. This experiment aims to assess the real-time adaptability of policy frameworks. The simulation will begin with the true DGP consistent with a single model structure (e.g., New Keynesian with a steep curve). At a random time in the middle of the simulation, the DGP will change permanently to another model structure (e.g., Backward-Looking with a flat curve), simulating a structural break in the economy.

- **Hypothesis:** The static CE agents and overly pessimistic H-S agents will suffer a significant increase in cumulative loss when the DGP has a structural break, as their static policy framework cannot adapt to the new underlying economic reality. In contrast, we hypothesize that the Bayesian HJB agent's beliefs, $\mathbf{p}_t$, will change when the new data pattern emerges. The agent will then update its policy i_t^* and achieve a lower cumulative loss. This experimental setup will highlight the economic value of endogenous learning in providing robustness against structural breaks.

4 Expected Outcomes and Research Impact

This research is expected to provide the first computationally tractable solution to the canonical continuous-time dual control problem in a modern macroeconomic context. Its main result is a new, adaptive policy function, $i^*(\mathbf{K}_t, \mathbf{p}_t)$, that optimally balances short-term economic stabilization with the long-term benefit of reducing structural uncertainty through active learning. The research contribution will be twofold. Methodologically, it provides a generalizable computational toolbox, the Deep BSDE method, for this class of previously computationally intractable dynamic economic models that incorporate learning. Substantively, the research provides a new normative benchmark in monetary policy. Our simulation experiments aim to show that the adaptive policy is more robust than the brittle single model CE approach and the extremely pessimistic H-S framework, and its dual control premium provides a formal quantification of the economic value of long-run policy experimentation.

This work serves as a step towards a long-term research agenda, which aims to make this framework fully policy-relevant through two complementary thrusts. Firstly, the theoretical extension is to formulate a more comprehensive model of uncertainty. Following the dichotomy in Hansen and Sargent (2023), the next logical step is to combine our solution for prior uncertainty with robust control approaches to also address likelihood misspecification uncertainty. The second thrust is empirical applications, and the best test of the framework is to apply it to historical macroeconomic data and perform holistic counterfactual analysis. Although this direction faces challenges in model parameter calibration and identification of exogenous shocks, it aims to evaluate whether an adaptive policy could have mitigated major historical economic policy mistakes. The combination of theoretical extensions and empirical applications aims to create a new framework for central bank decision making in a world of uncertainty.

References

Bellman, R. E. (1957). *Dynamic Programming.* Princeton University Press, Princeton, NJ.

Feldbaum, A. A. (1961). Dual control theory i. *Automation and Remote Control*, 21(9):874–880. Original Russian published in Avtomatika i Telemekhanika, 21(9), pp. 1240–1249, 1960.

Fleming, W. H. and Rishel, R. W. (1975). *Deterministic and Stochastic Optimal Control*, volume 1 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, Berlin.

Han, J., Jentzen, A., and E, W. (2018). Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34):8505–8510.

Hansen, L. P. and Sargent, T. J. (2008). *Robustness.* Princeton University Press, Princeton, NJ.

Hansen, L. P. and Sargent, T. J. (2023). Risk, ambiguity, and misspecification: Decision theory, robust control, and statistics. *Journal of Economic Literature*, 61(2):534–603.

Kushner, H. J. (1967). *Stochastic Stability and Control*, volume 33 of *Mathematics in Science and Engineering.* Academic Press, New York.

Nie, P. and Roulleau-Pasdeloup, J. (2023). The promises (and perils) of control-contingent forward guidance. *Journal of Monetary Economics*, 133:1–17.

Qu, Z. and Tkachenko, D. (2017). Global identification in dsge models allowing for indeterminacy. *Quantitative Economics*, 8(3):895–936.

Qu, Z. and Tkachenko, D. (2023). Using arbitrary precision arithmetic to sharpen identification analysis for dsge models. *Journal of Applied Econometrics*, 38(4):528–543.

Roulleau-Pasdeloup, J. (2020). Optimal monetary policy and determinacy under active/passive regimes. *Journal of Monetary Economics*, 116:100–115.

A Derivation of the Kushner-Stratonovich Equation

This appendix provides a detailed derivation of the Kushner-Stratonovich equation, which governs the dynamics of the central bank's beliefs. The derivation uses the reference measure method.

A.1 Framework and Notation

- **Model Space:** A set of M possible models, $\Theta = \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_M\}$.
- **State Dynamics:** If model m is true, the state $\mathbf{K}_t \in \mathbb{R}^d$ evolves according to the Stochastic Differential Equation (SDE):

$$d\mathbf{K}_t = \boldsymbol{\mu}_{\boldsymbol{\theta}_m}(\mathbf{K}_t, i_t)dt + \boldsymbol{\Sigma} d\mathbf{B}_t$$

  where $\mathbf{B}_t$ is a standard Brownian motion under the true probability measure. For brevity, we denote $\boldsymbol{\mu}_{\boldsymbol{\theta}_m}$ as $\boldsymbol{\mu}_m$.

- **Beliefs:** The central bank's posterior probability that model m is true, given the history of observations $\mathcal{F}_t = \sigma(\mathbf{K}_s : 0 \leq s \leq t)$, is $p_{m,t}$. The belief vector is $\mathbf{p}_t = [p_{1,t}, \ldots, p_{M,t}]^T$.
- **Innovation Process:** From the policymaker's perspective, the expected drift is the belief-weighted average $\bar{\boldsymbol{\mu}}_t = \sum_{j=1}^M p_{j,t}\boldsymbol{\mu}_j$. The innovation process $d\mathbf{Z}_t$ is the part of the observed data that is a surprise to the policymaker:

$$d\mathbf{Z}_t = \boldsymbol{\Sigma}^{-1}(d\mathbf{K}_t - \bar{\boldsymbol{\mu}}_t dt)$$

  Under the policymaker's subjective probability measure, $d\mathbf{Z}_t$ is a standard Brownian motion.

- **Signal-to-Noise Ratio:** It is convenient to define the "signal" vector $\boldsymbol{h}_m = \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_m$ and its belief-weighted average $\bar{\boldsymbol{h}}_t = \sum_{j=1}^M p_{j,t}\boldsymbol{h}_j$.

The goal is to derive the SDE for $p_{m,t}$. The result is the Kushner-Stratonovich equation:

$$dp_{m,t} = p_{m,t}\left(\boldsymbol{h}_m - \bar{\boldsymbol{h}}_t\right)^T d\mathbf{Z}_t$$

A.2 Derivation via the Reference Measure Method

This derivation proceeds in four main stages: 1. Introduce a simpler, hypothetical reference measure. 2. Characterize the unnormalized beliefs under this measure (the Zakai equation). 3. Use Itô's lemma to find the dynamics of the normalized beliefs in terms of the raw observations $d\mathbf{K}_t$. 4. Translate this result into the language of the innovation process $d\mathbf{Z}_t$.

A.2.1 Step 1: The Unnormalized Filter and the Zakai Equation

The standard method to derive the filter is to introduce a reference measure, P_0, under which the process $\mathbf{K}_t$ has zero drift:

$$d\mathbf{K}_t = \mathbf{\Sigma} d\mathbf{W}_t$$

where $\mathbf{W}_t$ is a Brownian motion under P_0.

By the Kallianpur-Striebel formula, the posterior probability can be expressed as the ratio of an unnormalized filter $\pi_{m,t}$ to its sum, $p_{m,t} = \pi_{m,t}/\sum_j \pi_{j,t}$. The unnormalized filter $\pi_{m,t}$ evolves according to a simpler, linear SDE known as the Zakai equation:

$$d\pi_{m,t} = \pi_{m,t}(\boldsymbol{h}_m)^T \mathbf{\Sigma}^{-1} d\mathbf{K}_t$$

The normalizing constant is $\Pi_t = \sum_{j=1}^{M} \pi_{j,t}$. Its dynamics are found by summing the individual Zakai equations:

$$d\Pi_t = \sum_{j=1}^{M} d\pi_{j,t} = \left(\sum_{j=1}^{M} \pi_{j,t} \boldsymbol{h}_j^T \right) \mathbf{\Sigma}^{-1} d\mathbf{K}_t = \Pi_t \left(\sum_{j=1}^{M} p_{j,t} \boldsymbol{h}_j^T \right) \mathbf{\Sigma}^{-1} d\mathbf{K}_t$$

$$d\Pi_t = \Pi_t (\bar{\boldsymbol{h}}_t)^T \mathbf{\Sigma}^{-1} d\mathbf{K}_t$$

A.2.2 Step 2: Applying Itô's Lemma for Ratios

We find the SDE for the normalized belief $p_{m,t} = \pi_{m,t}/\Pi_t$ using Itô's lemma for a function $f(\pi_m, \Pi) = \pi_m/\Pi$.

$$dp_m = \frac{1}{\Pi} d\pi_m - \frac{\pi_m}{\Pi^2} d\Pi + \frac{\pi_m}{\Pi^3} (d\Pi)^2 - \frac{1}{\Pi^2} d\langle \pi_m, \Pi \rangle_t$$

The quadratic variation terms are computed using $(d\mathbf{K}_t)(d\mathbf{K}_t)^T = \mathbf{\Sigma}\mathbf{\Sigma}^T dt$:

- $(d\Pi)^2 = \Pi_t^2(\bar{\boldsymbol{h}}_t)^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T)(\boldsymbol{\Sigma}^{-1})^T\bar{\boldsymbol{h}}_t dt = \Pi_t^2\|\bar{\boldsymbol{h}}_t\|^2 dt$
- $d\langle\pi_m, \Pi\rangle_t = \pi_{m,t}\Pi_t(\boldsymbol{h}_m)^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T)(\boldsymbol{\Sigma}^{-1})^T\bar{\boldsymbol{h}}_t dt = \pi_{m,t}\Pi_t(\boldsymbol{h}_m^T\bar{\boldsymbol{h}}_t)dt$

Substituting these into the Itô formula:

$$dp_m = \frac{1}{\Pi_t}(\pi_{m,t}\boldsymbol{h}_m^T\boldsymbol{\Sigma}^{-1}d\mathbf{K}_t) - \frac{\pi_{m,t}}{\Pi_t^2}(\Pi_t\bar{\boldsymbol{h}}_t^T\boldsymbol{\Sigma}^{-1}d\mathbf{K}_t) + \frac{\pi_{m,t}}{\Pi_t^3}(\Pi_t^2\|\bar{\boldsymbol{h}}_t\|^2 dt) - \frac{1}{\Pi_t^2}(\pi_{m,t}\Pi_t\boldsymbol{h}_m^T\bar{\boldsymbol{h}}_t dt)$$

Simplifying by substituting $p_m = \pi_m/\Pi$:

$$dp_m = p_m\boldsymbol{h}_m^T\boldsymbol{\Sigma}^{-1}d\mathbf{K}_t - p_m\bar{\boldsymbol{h}}_t^T\boldsymbol{\Sigma}^{-1}d\mathbf{K}_t + p_m\|\bar{\boldsymbol{h}}_t\|^2 dt - p_m(\boldsymbol{h}_m^T\bar{\boldsymbol{h}}_t)dt$$

Grouping terms gives the SDE for p_m in terms of the observable process $d\mathbf{K}_t$:

$$dp_{m,t} = p_{m,t}(\boldsymbol{h}_m - \bar{\boldsymbol{h}}_t)^T\boldsymbol{\Sigma}^{-1}d\mathbf{K}_t + p_{m,t}\left(\|\bar{\boldsymbol{h}}_t\|^2 - \boldsymbol{h}_m^T\bar{\boldsymbol{h}}_t\right)dt$$

A.2.3 Step 3: Rewriting in Terms of the Innovation Process

The final step is to substitute the definition of the innovation process, $d\mathbf{Z}_t$. We have:

$$\boldsymbol{\Sigma}^{-1}d\mathbf{K}_t = d\mathbf{Z}_t + \bar{\boldsymbol{h}}_t dt$$

Substituting this into the SDE for $dp_{m,t}$:

$$dp_{m,t} = p_{m,t}(\boldsymbol{h}_m - \bar{\boldsymbol{h}}_t)^T(d\mathbf{Z}_t + \bar{\boldsymbol{h}}_t dt) + p_{m,t}\left(\|\bar{\boldsymbol{h}}_t\|^2 - \boldsymbol{h}_m^T\bar{\boldsymbol{h}}_t\right)dt$$

Now, we expand and collect all the drift (dt) terms:

$$\begin{aligned}\text{Drift Term} &= p_{m,t}(\boldsymbol{h}_m - \bar{\boldsymbol{h}}_t)^T\bar{\boldsymbol{h}}_t dt + p_{m,t}\left(\|\bar{\boldsymbol{h}}_t\|^2 - \boldsymbol{h}_m^T\bar{\boldsymbol{h}}_t\right)dt \\ &= p_{m,t}\left[(\boldsymbol{h}_m^T\bar{\boldsymbol{h}}_t - \bar{\boldsymbol{h}}_t^T\bar{\boldsymbol{h}}_t) + (\|\bar{\boldsymbol{h}}_t\|^2 - \boldsymbol{h}_m^T\bar{\boldsymbol{h}}_t)\right]dt \\ &= p_{m,t}\left[\boldsymbol{h}_m^T\bar{\boldsymbol{h}}_t - \|\bar{\boldsymbol{h}}_t\|^2 + \|\bar{\boldsymbol{h}}_t\|^2 - \boldsymbol{h}_m^T\bar{\boldsymbol{h}}_t\right]dt = 0\end{aligned}$$

The drift terms cancel exactly. This leaves only the martingale component driven by the innovation process.

A.3 Final Result

The resulting SDE for the evolution of the posterior belief $p_{m,t}$ is the Kushner-Stratonovich equation:

$$dp_{m,t} = p_{m,t} \left(\boldsymbol{h}_m - \bar{\boldsymbol{h}}_t \right)^T d\mathbf{Z}_t$$

This equation has a clear economic interpretation: the belief in model m increases when the observed innovation $d\mathbf{Z}_t$ aligns with the prediction error of model m relative to the average belief $(\boldsymbol{h}_m - \bar{\boldsymbol{h}}_t)$. It is a pure martingale under the agent's subjective measure, meaning that on average, the agent does not expect their beliefs to change. Changes only occur in response to genuine surprises.