**Derivatives Pricing Research Projects**

*Deep Hedging with Reinforcement Learning & Generative Market Simulation*

Data: WRDS + Bloomberg

**Idea 1: Deep Hedging with Reinforcement Learning**

## Executive Summary

This project will develop a reinforcement learning (RL)–based framework for hedging equity options, with the goal of outperforming traditional strategies such as delta- and delta–gamma-hedging. The idea is to simulate a realistic trading environment in which an agent manages an options position by dynamically trading the underlying stock. The environment incorporates real-world frictions like transaction costs, discrete rebalancing, and stochastic volatility.

The RL agent learns how much of the underlying to hold at each point in time in order to minimize risk (such as the variance or tail risk of hedging errors) while controlling costs. We will train the agent in simulated markets calibrated to historical implied volatility surfaces from OptionMetrics, and then test its performance out-of-sample on actual SPX and SPY historical price paths.

The expected outcome is a clear demonstration of how data-driven hedging policies adapt to different volatility regimes (e.g., hedging less in calm markets, hedging more aggressively in crises) and reduce extreme hedging losses compared to classical approaches. This has direct implications for risk management and product structuring.

## Objectives & Key Questions

* Can RL reduce hedging error variance and tail risk versus delta/delta–gamma under transaction costs?
* How do RL policies adapt to volatility regimes and discrete trading?
* What is the economic significance (cost savings, CVaR reduction) in realistic settings?

## Literature & Positioning

* Deep Hedging (Buehler et al., 2019) – RL in frictional markets; we extend with calibration to listed equity options.
* Frontiers/Recent works – RL improves tail behavior; we add historical replay on SPX/SPY with OptionMetrics Greeks.
* Gap – limited studies tie RL hedging to real market surfaces and publish detailed OOS historical tests.

## Universe & Data

* Underlyings: SPX, SPY. Instruments: ATM/OTM calls/puts, 30–60 DTE; extend to sector ETFs later.
* WRDS OptionMetrics: NBBO mids, IVs, Greeks, dividends, yield curve; CRSP/Bloomberg: prices, realized vol, costs.
* Period: 2010–2025 (OOS replay windows for crises: 2011 EU, 2015–16, 2020, 2022).

## Methodology

Environment & Dynamics

* Start with GBM; extend to Heston/SABR; optional jumps for stress realism.
* Transaction costs: proportional + fixed; inventory and trade-size constraints.
* State: time-to-maturity, spot S, IV or vol estimate, Greeks (Δ, Γ, V), current hedge, recent realized vol.

RL Algorithms & Rewards

* Actor–Critic (DDPG / PPO). Continuous action = hedge ratio or trade size.
* Rewards: (i) −Var(hedge P&L) − λ·cost; (ii) −CVaRα(hedge P&L) − λ·cost (tail-aware).
* Stabilization: target networks, prioritized replay, entropy regularization.

Benchmarks & Controls

* Delta, delta–gamma (with second instrument), static hedge, hedging at fixed intervals.
* Calibrate environments to OptionMetrics surfaces; use Bloomberg spreads/fees.

## Experiments & Evaluation

1. Sim GBM → sanity (learning curve, hedge P&L distribution).
2. Add stochastic volatility; stress test across regimes.
3. Historical replays on SPY/SPX: rolling windows (e.g., 1y steps) with frozen policy.
4. Ablations: cost levels, re-hedge frequency, reward choices, discrete vs. continuous actions.

Metrics

* Mean/variance of hedging P&L; CVaR(95%), VaR(99%); cumulative cost; turnover.
* Sharpe of residual P&L; hit rates on tail containment; drawdown statistics.

## Economic Significance Benchmarks

* ≥15–30% reduction in CVaR(95%) vs. delta under realistic cost assumptions.
* ≥5–10% reduction in average cost for similar risk; or same risk with fewer trades.
* Robust across subperiods (pre/post-2016, COVID, 2022 tightening).

## Risks & Mitigations

* Overfitting to simulator → use calibrated parameters + historical replay; cross-regime tests.
* Unstable training → try PPO, tune reward scaling, normalize states, early stopping.
* Black-box policy → visualize policy surfaces (hedge vs. S, t, σ), SHAP on state features.

## Deliverables

* Paper: methodology, calibration to OptionMetrics, results (sim + OOS), robustness grid.
* Appendix: environment specs, reward definitions, hyperparameters.
* Code repo with configs; reproducible seeds; data dictionary; Excel tear sheets for key cases.

**Idea 2: Generative Market Simulation (Neural SDEs / Diffusion)**

## Executive Summary

## This project will build generative models — such as diffusion probabilistic models, GANs, or neural stochastic differential equations (SDEs) — that can learn the statistical properties of equity markets directly from historical data. Instead of assuming returns follow a lognormal distribution (as in Black–Scholes) or a specific stochastic volatility model (like Heston), the generative model will capture patterns in the data such as fat tails, volatility clustering, and regime shifts.

## Once trained on S&P 500 and SPY data, the model will be used to generate realistic synthetic market scenarios. These scenarios will then serve two purposes: (1) pricing path-dependent options such as Asian or barrier options by Monte Carlo simulation, and (2) stress-testing equity portfolios under rare but plausible crash scenarios. We will validate the realism of generated scenarios by comparing their statistical properties (distributional shape, autocorrelation, implied volatility smiles) to historical data and standard models.

## The expected outcome is that the generative model will replicate stylized facts of equity markets more accurately than classical models and produce derivative prices and stress test outcomes that better reflect real-world tail risks. This approach could provide a more robust simulation engine for both pricing and risk management applications.

## Objectives & Key Questions

* Can a generative model reproduce empirical stylized facts (heavy tails, vol clustering) better than Heston?
* Do AI-generated scenarios yield option prices closer to market (e.g., higher crash-risk premia)?
* Can conditional generation (e.g., high-VIX regime) support scenario-based risk management?

## Literature & Positioning

* Limitations of GBM/Heston; rough volatility & long memory evidence.
* Neural SDEs & diffusion models: strong fit for time-series generation; limited finance-specific empirical validation.
* Our angle: rigorous validation suite + derivatives applications (pricing & stress).

## Universe & Data

* SPX/SPY daily returns (1990–2025); optional sector ETFs; VIX for conditioning.
* WRDS CRSP (returns); OptionMetrics (IV surfaces for validation); Bloomberg (macro states).
* Optional TAQ subset for intraday stylized facts (advanced).

## Methodology

Model Families

* Diffusion models for time series (denoising diffusion).
* WGAN-GP for stability; sequence-aware architectures (1D CNN/LSTM blocks).
* Neural SDE: learn drift/diffusion; optionally enforce arbitrage-aware constraints.

Training & Conditioning

* Windows of 252 trading days; augmentation by rolling windows.
* Conditional inputs: VIX regime (low/med/high), rate levels, credit spread proxy.
* Early stopping via validation of stylized facts and distribution metrics.

Validation Suite

* Distribution: mean/σ/skew/kurtosis; QQ plots; KS/Wasserstein distances.
* Dependence: ACF/PACF of returns and |returns|; volatility clustering indices.
* Implied: Monte Carlo option prices (ATM/OTM puts/calls); does a smile emerge?
* Regime fidelity: conditional generation reproduces target regime statistics.

## Applications & Experiments

1. Baseline: compare generated vs. Heston vs. historical stylized facts.
2. Option Pricing: price Asian/barrier options via generated paths vs. Heston; compare to market IV levels.
3. Stress Testing: sample tail-heavy scenarios; compute portfolio drawdowns vs. Gaussian/Heston scenarios.
4. Multivariate: extend to SPX + VIX or SPX + sectors to capture co-movements.

## Economic Significance Benchmarks

* Generated-path option prices show higher tail premia for deep OTM puts vs. BS/Heston, closer to market levels.
* Risk: AI scenarios produce fatter lower-tail portfolio loss distributions vs. Gaussian, improving stress awareness.
* Scenario control: conditional models reliably shift to high-vol regimes when prompted.

## Risks & Mitigations

* GAN instability → prefer diffusion / WGAN-GP; monitor mode coverage.
* Evaluation ambiguity → pre-commit to validation metrics; hold-out periods; stylized fact scorecards.
* Black-box nature → analyze learned roughness (Hurst), regime mixes; report interpretable diagnostics.

## Deliverables

* Paper: model descriptions, validation suite, pricing & stress applications.
* Appendix: training details, hyperparameters, data preprocessing pipeline.
* Code repo with trained checkpoints; scripts for scenario generation; option pricing notebook.

# Appendix: Shared Infrastructure & Reproducibility

* Versioned data snapshots (WRDS pulls with date stamps); code-driven filters (OptionMetrics zero-bid, OI thresholds).
* Config-driven experiments (YAML): environments, rewards, model hyperparameters.
* Deterministic seeds; train/val/test splits by time; cluster-friendly logging (weights & biases or MLflow).
* Economic testing standard: include fees/slippage; capacity checks (OI, notional, turnover).
* Ethical/compliance: simulated only unless approvals; document data licenses; anonymize any proprietary series.

**Pros & Cons**

**Deep Hedging with Reinforcement Learning**

| **Pros** | **Cons** |
| --- | --- |
| Novel and high-impact — extends classic hedging to modern AI methods. | Implementation complexity: requires RL design, tuning, and calibration. |
| Strong integration of **machine learning + stochastic calculus**. | Computationally demanding (many simulated episodes). |
| Calibratable to **OptionMetrics/Bloomberg data**, allows real-market replay tests. | Risk of overfitting to simulator dynamics if calibration is weak. |
| Targets **practical problems** (hedging under transaction costs, discrete trading). | Interpretability challenges: black-box policy may be hard to explain. |
| Benchmarkable vs. delta/delta-gamma strategies with clear P&L metrics. | Needs careful definition of reward (variance vs CVaR vs cost). |
| Aligned with risk management & structuring desks — strong publication relevance. | Limited existing empirical studies → less guidance, more trial-and-error. |

**Generative Market Simulation (Neural SDEs / Diffusion)**

| **Pros** | **Cons** |
| --- | --- |
| Highly original — few finance papers applying diffusion/GANs to derivatives. | Training generative models for finance is technically challenging. |
| Captures **stylized facts** (fat tails, volatility clustering) missed by GBM/Heston. | Evaluation is less straightforward — need robust statistical validation. |
| Applications to both **derivative pricing** (Asian/barrier options) and **stress testing**. | Black-box nature → limited parameter interpretability vs Heston/SABR. |
| Conditional generation possible (simulate high-VIX regimes, macro stress). | Risk of arbitrage-violating or unrealistic outputs if not constrained. |
| Uses available WRDS/Bloomberg data (returns, vols, surfaces). | GAN instability; diffusion training can require GPU resources and tuning. |
| Bridges finance & frontier ML → strong novelty for publication. | Contribution is indirect: generates scenarios, not explicit closed-form pricing. |