Macro-Conditioned Diffusion Models for Financial Time Series Generation

Dominic Ranelli[1]

[1] Worcester Polytechnic Institute

Abstract

Financial return series are characterized by a small set of persistent stylized facts, most notably heavy tailed marginal distributions, volatility clustering, and strong dependence on macroeconomic regimes (Cont 2001). Despite decades of research, existing generative models struggle to reproduce these properties jointly (Su et al. 2025) (Stockhammer 2008). Classical econometric approaches, such as GARCH and regime-switching volatility models, impose restrictive parametric assumptions and limited functional flexibility (Engle 1982) (Bollerslev 1986) (Hamilton 1996). More recent deep generative models, including GANs, VAEs, and time-domain diffusion models, often suffer from training instability, inadequate tail behavior, or an inability to represent multi-scale temporal structure and macro-conditioned dynamics (Cobbinah et al. 2025) (Lucas et al. 2019) (Su et al. 2025).

In this paper, we propose a conditional diffusion model operating in the continuous wavelet domain for the generation of financial return series (Ramsey 2002) (Mallat 1988). By transforming returns into a time–frequency representation, the model naturally captures volatility dynamics across multiple horizons, while diffusion-based training provides a stable, likelihood-based generative framework. Conditioning variables, including macroeconomic indicators and discrete regime states, are incorporated directly into the reverse diffusion process, enabling controlled generation and counterfactual analysis.

We evaluate the proposed approach using a battery of diagnostics grounded in well-established stylized facts rather than forecasting accuracy. These include excess kurtosis, extreme absolute return quantiles, survival functions of return magnitudes, and the autocorrelation structure of squared returns (Cont 2001) (Barnes and Hughes 2002). In addition, we perform regime-conditioned counterfactual experiments in which identical latent noise realizations generate qualitatively distinct return dynamics under expansionary and crisis regimes (Hamilton 1996).

Our results demonstrate that wavelet-domain conditional diffusion models can simultaneously reproduce heavy tails, volatility clustering, and regime-dependent behavior, offering a flexible and theoretically consistent framework for financial time-series generation and scenario analysis

.

Introduction

Financial asset returns exhibit a collection of robust empirical regularities that persists across asset classes, markets, and sampling frequencies (Cont 2001). These regularities pose a fundamental challenge to both econometric and machine learning-based modeling approaches, especially in generative settings where the objective is not point prediction but true reproduction of the joint distributional and overall structure of returns. A model intended for simulation, stress testing, or scenario analysis must be evaluated not only on likelihood or forecasting accuracy, but on its ability to reproduce stylized facts in a statistically meaningful way (Rippel 2009). This paper focuses on three properties that are central to modern quantitative finance: heavy-tailed return distributions, volatility clustering, and regime-dependent dynamics. While each has been studied in isolation, capturing all three within a single, stable, and conditionally controllable generative framework remains an open challenge. In this paper, we argue that recent advances in diffusion-based modeling, when combined with a wavelet-domain representation and explicit macroeconomic conditioning, offer a promising path forward.

Stylized facts of financial returns

One of the earliest observations in financial economics is that asset return distributions exhibit heavy tails relative to the Gaussian benchmark. Mandelbrot (1963) demonstrated that speculative price changes display more extreme events than is predicted by a normal distribution, which has been confirmed across markets and horizons (Mandelbrot et al. 1963).

Recent empirical studies, including the comprehensive survey by Cont (2001), show that return distributions can typically possess high excess kurtosis, with large price movements occurring orders of magnitude more frequently than under Gaussian assumptions. This property has significant implications for risk management and motivates the use of tail-sensitive diagnostics such as extreme quintile, survival functions, and excess kurtosis in model evaluation (Cont 2001) (Jorion 1996).

A second core stylized fact is volatility clustering. This is the empirical regularity that large price changes tend to be followed by large changes, and vice versa. While raw returns often show negligible autocorrelation, nonlinear transformations (such as squared or absolute returns) display strong and persistent dependence over time. Engle (1982) formalized this phenomenon through the ARCH framework, and Bollerslev (1986) generalized it using GARCH models, showing that conditional heteroskedasticity provides an explanation for clustered volatility (Bollerslev 1986) (Engle 1982). Volatility clustering can now be understood as a defining feature of financial time series and serves as a critical benchmark for any realistic generative model.

Another related property is regime dependence. Financial markets do not operate under a single stationary environment. Instead, their statistical and geometric properties change across time in response to macroeconomic conditions, monetary policy, crises, and structural shifts. Hamilton (1989) introduced Markov regime-switching models to capture such behavior (Hamilton 1996) (Ang and Piazzesi 2003). This showed that allowing parameters to switch across latent states substantially improves explanatory power for macroeconomic and financial series. From a generative perspective, regime dependence implies that a model should not only reproduce unconditional stylized facts, but should also allow controlled variation in behavior across distinct market states (Hamilton 1989) (Jorion 1996).

Together, heavy tails, volatility clustering, and regime dependence define a set of constraints. A model that fails to reproduce any one of these properties cannot be considered a credible simulator of financial returns, regardless of its performance on simpler metrics.

Limitations of existing approaches

Traditional econometric models capture important subsets of these stylized facts, however they are limited in scope. GARCH-type models provide a principled explanation for volatility clustering through parametric conditional variance dynamics. However, they rely on strong distributional assumptions, such as Gaussian distributions. As a result, they struggle to represent complex, nonlinear dependencies across multiple horizons or across a large cross-section of assets.

Regime-switching models extend this framework by allowing parameters to change across discrete states, however introduce additional assumptions regarding number of regimes, transition dynamics, and parametric form within a regime. These assumptions can be difficult to validate empirically, especially in high-dimensional settings, and may limit flexibility as regimes evolve in a continuous manner.

Recently, deep generative models such as generative adversarial networks (GANs) and variational autoencoders (VAEs) have been applied to financial time series. While these models offer more expressiveness per sample, they also introduce challenges. GANs are unstable to train and often suffer from mode collapse, which is problematic for modeling tail behavior. VAEs are more stable, however tend to over-smooth distributions and under represent extreme events unless engineered to do so. Empirical studies have found that such models reproduce central tendencies while failing to capture heavy tails or persistent volatility dynamics.

Diffusion-based generative models have emerged as a promising alternative due to their stable, likelihood-based training and strong empirical performance in high-dimensional datasets. However, most current applications to time series operate directly in the time domain. Financial returns are highly nonstationary and exhibit structure at multiple temporal scales. As a result, modeling them exclusively as raw sequences can make it difficult for time-domain diffusion models to learn long-term volatility patterns without excessive depth or parameterization.

Motivation for wavelets and diffusion

The approach taken in this paper is motivated by two observations. First, wavelet transforms provide a natural representation for financial time series by decomposing signals into localized time-frequency components. Unlike Fourier methods, wavelets retain temporal localization, which allows the model to distinguish between short-term fluctuations and long-term volatility regimes. Wavelets have a history in financial analysis, specifically for studying multiscale volatility and dependence and are also well suited for nonstationary data.

Additionally, diffusion models offer a principled and stable framework for generative modeling. By learning to reverse a dynamic noise-injection process, diffusion models avoid adversarial training dynamics and provide a well-defined likelihood objective. This makes them the model of choice for applications where sample fidelity and distributional correctness are more important than sharpness or virtual realism.

Most importantly, diffusion models are naturally amenable to conditioning. Through incorporating macroeconomic variables or regime indicators into the denoising network, one can guide the generative process in a controlled manner. This enables not only unconditional generation of realistic returns, but also counterfactual generation under alternative macro or regime scenarios. This specifically is a capability of direct importance to quantitative finance applications such as stress testing and scenario analysis.

By combining a wavelet-domain representation with a conditional diffusion model (C-DDPM), we aim to exploit the strengths of both approaches"

- Multiscale structure from wavelets
- Flexible generation from diffusion

Contributions

The contributions of this paper are as follows:

- We introduce a wavelet-domain diffusion model for financial return generation, operating on multiscale time-freq representations rather than raw returns.
- We incorporate explicit macroeconomic and regime conditioning, enabling controlled generation throughout differing market states

Together, these contributions establish a coherent and empirically grounded framework for generative modeling of financial time series.

Related Work

This work builds on three distinct yet complementary strands of literature: classical financial econometrics, deep generative modeling for financial time series, and wavelet-based representations of economic and financial data. Each of these topics addresses important aspects of financial dynamics, yet no single framework fully resolves the joint challenge of modeling heavy-tailed distributions, volatility clustering, and regime-dependent behavior within a stable, conditionally controllable generative setting.

Financial econometrics

The econometric literature on financial time series has emphasized the inadequacy of Gaussian, homoscedastic models for asset returns. A central contribution in this area is the introduction of auto-regressive conditional heteroskedasticity (ARCH) by Engle (1982), where time-varying conditional variance is modeled as a function of past squared returns. Bollerslev (1986) generalized this framework to GARCH, which allowed volatility to depend on its own past as well as prior shocks. This was able to capture the volatility clustering observed in asset returns. These models provide an interpretable explanation for second-moment dependence and remain foundational in risk-management and option pricing.

Recent extensions introduced richer volatility dynamics through stochastic volatility (SV) models, where volatility follows a latent stochastic process rather than a deterministic recursion. SV models remove the evolution of volatility from observed returns and often provide more flexibility in capturing nonlinear volatility behavior. It should be noted that they usually require strong distributional assumptions and computationally intensive inference, specifically in multivariate or high-frequency settings (Taylor 1994).

A related line of work focuses on regime-switching models, notably the Markov-switching framework introduced by Hamilton (1989). In these models, the data-generating process transitions between a finite number of latent regimes, each with differing parameters, according to a markov chain. Regime-switching models have been applied extensively to macroeconomic time series and financial returns, and offer a natural way to model structural breaks, crisis periods, and policy-driven changes in dynamics. These models require pre-specification of the number of regimes and impose strong assumptions on regime persistence and transition dynamics, which can limit flexibility and robustness in practice (Hamilton 1989).

While ARCH/GARCH, stochastic volatility, and regime-switching models successfully capture individual stylized facts, they are inherently parametric and often struggle to scale to high-dimensional nonlinear settings. Additionally, they are not designed as general-purpose generative models capable of producing rich synthetic data under arbitrary conditioning scenarios (Cont 2001) (Engle 1982) (Bollerslev 1986).

Deep generative models in finance

Motivated by the limitations of parametric econometric models, a growing area has explored deep generative models for financial time series. Early work applied generative adversarial models (GANs) to asset returns, aiming to reproduce marginal distributions and dependence structures without explicit parametric assumptions. While GANs offer considerable expressivity, studies have reported significant training instability, sensitivity to hyperparameters, and difficulty in reproducing tail behavior consistently. Mode collapse and lack of tractable likelihood complicates rigorous evaluation, especially when the objective is faithful reproduction of rare but econometrically important extreme events (Jeon et al. 2022) (Lin et al. 2024).

Variational autoencoders (VAEs) provide a likelihood-based alternative and have been used to model latent representations of financial time series. However, VAEs are prone to posterior collapse and often generate overly smooth samples unless carefully regularized (He et al. 2019) (Potluru et al. 2023). In practice, this can lead to underestimation of tail risk and attenuation of volatility dynamics. This limits their usefulness for applications where extreme behavior is of main concern.

More recently, diffusion probabilistic models (DDPMs) have emerged as a promising class of generative models due to their stable training dynamics and strong empirical performance. Diffusion models define a forward process that gradually perturbs data with noise and learn a reverse-time denoising process, yielding a well-defined likelihood objective (Watson et al. 2021) (Yang et al. 2023). Extensions of diffusion models to time-series settings have shown competitive performance in forecasting and imputation tasks, and conditional variants allow incorporation of side information. (Lim et al. 2023) (Yang et al. 2023). Most existing applications operate directly in the time domain and do not explicitly address the multi-scale structure characteristic of financial volatility. As a result, they may require substantial model capacity to implicitly learn long-horizon dependencies, and their ability to reproduce stylized facts such as volatility clustering remains an open empirical question.

Wavelets in finance

Wavelet methods offer a complementary perspective by representing time series in the time-frequency domain, enabling localized analysis across multiple scales. Unlike Fourier-based techniques, wavelets preserve temporal localization, making them particularly well suited to nonstationary signals such as financial returns (Mallat 1999). A substantial body of work has employed wavelets to study volatility dynamics, scale dependent correlations, and market efficiency across horizons (In and Kim 2013) (Ramsey 1999).

In financial applications, wavelets have been used to decompose returns into components associated with different investment horizons (Ramsey 2002) (In and Kim 2013), which reveals that volatility and dependence structures vary systematically across scales. This multiscale view aligns naturally with empirical observations of clustered volatility and regime-dependent behavior, which are often displayed differently at short and long horizons. Wavelet-based denoising and feature extraction techniques have also been applied to improve forecasting and risk estimation.

Despite their interpretability and empirical success, wavelet methods are typically used as analytical tools rather than components of fully generative models. Existing studies rarely integrate wavelet representations with modern deep generative frameworks in a way that allows end-to-end learning and conditional simulation (Bollerslev 1986) (Ramsey 2002) (Mallat 1988) (Yang et al. 2023).

Positioning of present work

This paper sits at the intersectional of these three literature. It progresses classical econometric models by adopting a nonparametric, deep generative approach, while retaining evaluation criteria grounded in established financial stylized facts. It extends recent diffusion-based time-series models by operating in the wavelet domain, thereby making multiscale structure explicit rather than implicit. Finally, it builds on wavelet-based financial analysis by embedding wavelet representations within a conditional generative framework. This enables a regime and macro-aware simulation rather than purely descriptive analysis.

Data

Financial data

The primary object of interest in this study is a panel of daily asset returns drawn from a fixed universe of liquid financial instruments. Let $P_t^{(a)}$ denote the adjusted closing price of asset a on trading day t, where $a \in 1, ..., A$ indexes assets and $t \in 1, ..., T$ indexes time. The asset universe is held constant throughout the sample period to avoid survivorship bias (Brown, Brown, and Rannala 1995) and to ensure consistent dimensionality in the generative model, with the final universe size equal to 300 assets.

Returns are constructed using continuously compounded (log) returns, defined as:

$r_t^{(a)} = logP_t^{(a)} - logP_{t-1}^{(a)}$

Log returns are standard in empirical finance due to their time-additive property and approximate scale invariance for small price changes (Lo and MacKinlay 1997). The resulting return matrix:

$R = R_t^{(a)} \in \mathbb{R}^{T \times A}$

is the base signal from which all subsequent representations are derived.

Prior to further processing, all price series are aligned on a common trading calendar. Missing observations arising from holidays or asset-specific trading interruptions are handled consistently across assets. The sample is truncated to the maximal contiguous interval from which all series are simultaneously available. This alignment ensures that any cross-sectional or temporal dependence learned by the model reflects genuine market structure rather than leaked information (Campbell et al. 1998).

Continuous wavelet transform

Financial return series are well known to be nonstationary and to exhibit structure across multiple temporal horizons. Rather than modeling returns directly in the time domain, we transform each asset's return series into a time-frequency representation using the continuous wavelet transform (CWT) (Cont 2001). This representation makes scale-dependent volatility and transient dynamics explicit, facilitating learning by the diffusion model. Formally, for a real-valued signal $x(t)$, the continuous wavelet transform is defined as:

$$\int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{s}} \psi^*(\frac{t-\tau}{s})dt$$

where $\psi(\cdot)$ is the chosen mother wavelet, $s > 0$ is the scale parameter, τ denotes time localization and * denotes complex conjugation. The scale parameter s controls the trade-off between time and frequency resolution: small scales capture high-frequency, short horizon fluctuations, while larger scales emphasize lower frequency, long horizon components (Polikar 1996).

In this study, we use the Morlet wavelet, which consists of a complex exponential modulated by a Gaussian envelope. The Morlet wavelet is widely used in time-frequency analysis due to its favorable localization properties in both time and frequency domains. For financial applications, it provides a natural balance between capturing oscillatory behavior and isolating localized volatility bursts, making it particularly suitable for representing return dynamics across investment horizons (Ramsey 2002).

For each asset a, the CWT is computed over a fixed set of scales $s \in s_1, ..., s_S$, with the total number of scales set to 32 in the baseline configuration. Which provides, for each asset, a matrix of wavelet coefficients indexed by time and scale. These matrices are stacked across assets to form a three-dimensional tensor:

$X \in \mathbb{R}^{T \times A \times S}$

which constitutes the primary input to the generative model.

To train the diffusion model, the wavelet tensor is segmented into overlapping time windows of fixed length W (128 trading days). Each training example thus corresponds to a tensor of shape $1 \times S \times W$ representing a localized multiscale snapshot of return dynamics for a single asset. This windowing allows the model to learn local temporal structure while sharing parameters across time and assets.

Macroeconomic Feature Engineering

To enable conditional generation and regime-aware simulation, the diffusion model is conditioned on a vector of macroeconomic and market-wide variables observed at each time t, so that:

$c_t \in \mathbb{R}^d$

denotes the conditioning vector at time t, where d is the total number of conditioning features.

The conditioning vector aggregates three categories of information. First, it includes macroeconomic indicators such as interest rates, yield spreads, inflation measures, FOREX, metals, and other widely used economic aggregates. These variables capture broad economic conditions that are known to influence asset return and volatility dynamics (Ang and Piazzesi 2003).

Secondly, the macro features include market-wide features, such as aggregate returns or volatility proxies, which summarize market states beyond individual assets (Cont 2001).

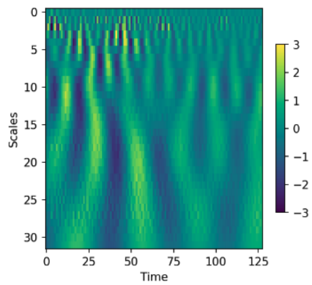

Figure 1: **Continuous wavelet scalogram of empirical financial returns.** The figure shows the magnitude of continuous wavelet coefficients for a representative asset over time, plotted as a function of scale and temporal location. Brighter regions indicate higher volatility energy at the corresponding horizon. The multiscale structure highlights the coexistence of short-term fluctuations and long-horizon volatility bursts, motivating the use of wavelet-domain representations for generative modeling.

Finally, and most importantly, the conditioning vector includes a regime indicator, derived from K-Means clustering (Darken and Moody 1990) of Principal Component Analysis (PCA) of all macro-features (Jolliffe 1993). We included a markov regime transition estimation to capture persistence of this state as well (Hamilton 1989). This variable encodes the overall market state, such as expansionary versus contractionary periods or low/high volatility environments, and is treated as an observed conditioning input rather than a latent variable inferred by the model. The regime indicator is placed as the final component of the conditioning vector to facilitate controlled counterfactual experiments in which the regime value is altered while all other inputs, specifically diffusion noise, are held fixed.

All conditioning variables are aligned to the same daily frequency as the return data. Lower frequency macroeconomic series are forward-filled to daily resolution, consistent with standard practice in empirical finance (Ghysels, Santa-Clara, and Valkanov 2004). Prior to model training, conditioning variables are standardized to zero mean and unit variance to ensure numerical stability and to prevent any feature from dominating the conditioning mechanism due to scale differences (Goodfellow 2016).

Model

The model is based on denoising diffusion probabilistic models (DDPMs), adapted to operate on multiscale wavelet representations and explicitly conditioned on macroeconomic and regime variables. The formulation below follows the general diffusion framework while introducing architectural and conditioning choices tailored to financial time-series data.

Forward diffusion process

The diffusion framework defines a fixed forward stochastic process that gradually corrupts the data with Gaussian noise over a sequence of diffusion steps (1000 here). Let $x_0 \in \mathbb{R}^{1\times S\times W}$ denote a clean wavelet-domain training example, where S is the number of wavelet scales and W is the temporal window length. The forward process constructs a sequence $x_t{}_{t=1}^{T_d}$ according to:

$q(x_t|x_{t-1}) = \mathcal{N}(\sqrt{1-\beta_t}x_{t-1}, \beta_t I)$

where $\beta_t{}_{t=1}^{T_d}$ is a predefined variance schedule with $0 < \beta_t < 1$.

This process incrementally destroys the structure in the original data by injecting Gaussian noise at each step. After a sufficiently large number of steps T_d, the distribution of x_{T_d} approaches an isotropic Gaussian, independent of the original signal. This property allows sampling to begin from pure noise and iteratively recover structure through a learned reverse process (Ho, Jain, and Abbeel 2020).

A useful closed-form expression for the forward process marginal is:

$q(x_t|x_0) = \mathcal{N}(\sqrt{\bar{a}_t}x_0, (1-\bar{a}_t)I),$

$\bar{a}_t = \prod_{i=1}^{t}(1-\beta_i)$

(Ho, Jain, and Abbeel 2020) This formulation allows efficient training by sampling x_t directly from x_0 at an arbitrary diffusion step t, without explicitly simulating all intermediate steps.

The noise schedule β_t governs how rapidly information is destroyed. In this work, we adopt a monotonic (cosine) schedule commonly used in diffusion models. A slowly increasing schedule ensures that early diffusion steps preserve substantial structure, while later steps approach near-complete noise. This has been shown to stabilize training and improve sample quality (Nichol and Dhariwal 2021b), especially for structured, high-dimensional data. For financial wavelet representations, a gradual schedule is preferred to avoid prematurely destroying scale-dependent volatility information that is critical for learning stylized facts.

Reverse process and conditional generation

The generative component of the model is the reverse diffusion process, which aims to invert the forward noising dynamics. The reverse process is parameterized by a neural network with parameters θ and defines the distribution:

$p_\theta(x_{t-1}|x_t, c) = \mathcal{N}(\mu_\theta(x_t, t, c), \sum_t)$

where $c \in \mathbb{R}^d$ is the conditioning vector containing macroeconomic variables and the regime indicator (Ho, Jain, and Abbeel 2020).

Modern diffusion models then reparameterize the model to predict the noise component ϵ added during the forward process. Under this parameterization, the network learns the function:

$\epsilon_\theta(x_t, t, c) \approx \epsilon$

where $\epsilon \sim \mathcal{N}(0, I)$. The predicted noise is then used to compute the mean of the reverse transition using a deterministic transformation involving $\bar{a}_t$ and β_t.

The conditioning vector c plays a central role in this framework. Conditioning variables are injected into the denoising network at each diffusion step, allowing the model to create its reconstruction of wavelet coefficients based on contemporaneous macroeconomic state and regime (Hamilton 1989). In contrast to latent-variable regime models, conditioning in this case is explicit and observable, as the regime indicator is provided directly as input rather than inferred. This design choice enables controlled generation and counterfactual analysis, where the effect of changing macro or regime inputs can be isolated by holding all other factors fixed.

From a financial perspective, this conditional structure mirrors the idea that return dynamics are influenced by macroeconomic and market regimes, while still allowing stochastic variability within each regime.

Training

The model is trained by minimizing a mean-squared error loss between the true noise ϵ and the network's predicted noise ϵ_θ. Specifically, it is noted as:

$\mathcal{L}(\theta) = \mathbb{E}_{x_0,t,\epsilon}[||\epsilon - \epsilon_\theta(\sqrt{\bar{a}_t}\epsilon, t, c)||_2^2]$

This loss is evaluated by sampling a diffusion step t uniformly from $1, ..., T_d$ sampling Gaussian noise ϵ, constructing the corresponding x_t and predicting ϵ using the network (Ho, Jain, and Abbeel 2020).

This training objective has multiple advantages that are relevant for financial applications. First, it defines a well-behaved, likelihood-based objective that avoids the adversarial dynamics found in GANs (Goodfellow 2016) (Cont 2001). This stability is crucial when modeling heavy-tailed data, where rare extreme events can destabilize adversarial training. Additionally, since the objective decomposes across diffusion steps, the model learns to denoise signals corrupted at varying noise levels, which improves robustness and generalization. Finally, the objective naturally accommodates conditioning without requiring additional regularization terms or auxiliary losses.

In practice, training is further stabilized through standard techniques such as gradient clipping and learning rate decay. These help prevent occasional large gradients from dominating parameter updates (Pascanu, Mikolov, and Bengio 2013).

Architecture & conditioning

The denoising network ϵ_θ is implemented as a U-Net architecture operating directly on wavelet-domain tensors. U-Nets are well suited for this task due to their hierarchical structure and skip connections, which preserves fine-grained local information while allowing global context to influence reconstruction. In the present setting, the encoder path progressively aggregates information across wavelet scales and time, while the decoder path reconstructs detailed structure, guided by skip connections from corresponding encoder layers (Ronneberger, Fischer, and Brox 2015).

The input to the network at each diffusion step is a tensor of shape $1\times S\times W$, along with an embedding of the diffusion timestep t. Early diffusion steps require delicate reconstruction of high-frequency structure, while later steps involve coarse denoising from near-Gaussian noise. Conditioning on the macroeconomic and regime vector c is incorporated through feature-wise modulation mechanisms (Zhang, Rao, and Agrawala 2023). Conditioning information is projected through learned linear layers and injected into intermediate feature maps. This allows conditioning variables to influence the denoising process at multiple levels of abstraction, rather than simply at the input layer. This improves controllability and reduces the risk that the model ignores conditioning inputs.

The regime indicator is treated identically to the other conditioning variables at the architectural level, however its placement as a dedicated scalar feature enables straightforward counterfactual analysis. By generating samples with identical diffusion noise while altering only the regime value, one can directly assess whether the model has learned regime-dependent dynamics rather than simply fitting unconditional distributions.

Generation & inverse transform

Once the conditional diffusion model has been trained, synthetic return series are generated by sampling from the learned reverse diffusion process and transforming the generated wavelet-domain representations into the time domain.

Conditional sampling from diffusion model

Generation begins by sampling an initial noise tensor:

$x_{T_d} \sim p_\theta(x_{t-1}|x_t, c), t = T_d, T_d - 1, ..., 1$

until a sample x_0 is obtained. Each reverse step removes a small amount of noise while reintroducing structure learned during training. This is guided both by the diffusion timestep t and the conditioning variables.

In this setting, the generative trajectory is influenced by the conditioning vector at each denoising step. This results in the macroeconomics variables and regime indicator modulating the evolution of the sample throughout the reverse process, rather than simply being applied as a post hoc adjustment. This enables controlled generation. By fixing the random seed used to initialize x_{T_d} and changing only the conditioning vector c, one can generate the counterfactual samples that isolate the effect of macroeconomic or regime changes on the generated dynamics.

The output of the sampling procedure is a wavelet-domain tensor:

$\hat{x}_0 \in \mathbb{R}^{1\times S\times W}$

representing a synthetic multiscale snapshot of return dynamics for a single asset over a window of length W (128).

Inverse wavelet transform

In order to evaluate the generated samples using standard financial diagnostics, the wavelet-domain outputs must be transformed back into the time domain. This is done through an inverse wavelet transform, which maps the generated wavelet coefficients back to a one-dimensional return series.

Let $\hat{W}(s, \tau)$ denote the generated wavelet coefficients for a given asset and window. The inverse transform reconstructs an estimate of the original signal $\hat{r}(t)$ by aggregating contributions across scales:

$\hat{r}(t) = \mathcal{W}^{-1}(\hat{W})(t)$

(Polikar 1996) This inverse transform is implemented using the corresponding inverse continuous wavelet transform or an equivalent reconstruction procedure consistent with the forward transform used during preprocessing.

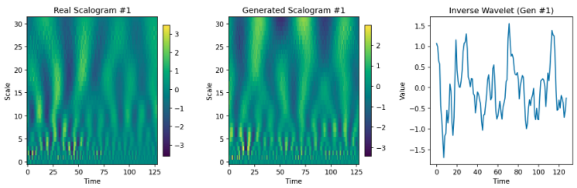

Figure 2: **Real, generated, and inverse transformed wavelet representations.** The left column shows an example continuous wavelet scalogram of empirical financial returns, the middle column displays wavelet-domain samples generated by the conditional diffusion model, and the right column shows the corresponding time-domain return series obtained via inverse wavelet transform. The generated scalogram closely reproduces the multiscale energy distribution of real data across short and long horizon scales. The inverse-transformed series exhibits realistic volatility bursts and temporal dependence. This demonstrates that the diffusion model learns a coherent wavelet-domain representation that can be safely mapped back to the return space.

The resulting reconstructed series $\hat{r}(t)$ represents a synthetic return path that can be directly compared to real returns using standard econometric tools. All diagnostics reported in the following sections (tail quantiles, excess kurtosis, etc.) are computed on these reconstructed return series.

Numerical stability considerations

Multiple numerical considerations come from reconstructing returns from generated wavelet coefficients. First, wavelet coefficients at differing scales may have substantially different magnitudes, especially during periods of high volatility. In order to prevent numerical instability during training and reconstruction, wavelet coefficients are normalized prior to diffusion modeling and rescaled appropriately during inversion (Polikar 1996).

Additionally, the diffusion process can occasionally produce extreme coefficient values (Pascanu, Mikolov, and Bengio 2013) (Nichol and Dhariwal 2021a), especially in early training. Gradient clipping and careful choice of diffusion schedule mitigate this issue while training. Also, clipping or soft normalization can be applied to generate coefficients prior to inverse transformation to prevent spurious numerical explosions.

Finally, since the inverse wavelet transform aggregates information across scales, small errors at individual scales may compound in the reconstructed signal. For this reason, evaluation focuses on distributional and dependence properties of returns, consistent with the goal of generative modeling rather than deterministic forecasting.

Evaluation Methodology

The objective of this study is not point-wise forecasting accuracy but a faithful of empirically established stylized facts of financial returns under both unconditional and conditional generation. Model evaluation is grounded in diagnostics that correspond to these stylized facts and are widely accepted in the financial econometrics literature (Cont 2001).

All diagnostics are computed on reconstructed return series obtained via the inverse wavelet transform described in section 5. Unless otherwise noted, metrics are evaluated on pooled samples across assets and time windows to obtain sufficient statistical power, consistent with standard practice in empirical finance (Campbell et al. 1998).

Heavy tails

A defining characteristic of financial returns is the presence of heavy tails, meaning that extreme outcomes occur with substantially higher probability than under a Gaussian distribution (Cont 2001). This property is central to risk management and is one of the earlier documented stylized facts of asset returns. In order to evaluate whether the generative model reproduces this behavior, we employ a set of diagnostics that capture tail thickness from multiple perspectives.

Excess kurtosis Excess kurtosis provides a scalar summary of tail heaviness relative to the normal distribution. For a return series r_t with mean μ and standard deviation σ, excess kurtosis is defined as:

$k = \mathbb{E}[(\frac{r-\mu}{\sigma})^4] - 3.$

A Gaussian distribution has zero excess kurtosis (Cont 2001), while positive values indicate heavier-than-normal tails.

Absolute return quantiles While excess kurtosis summarizes tail behavior in a single moment, it can be sensitive to outliers and does not localize behavior in specific tail regions. To complement this measure, we compute absolute return quantiles at high probability levels, specifically the 99% and 99.5% quantiles:

$q_p = Quantile(|r|, p), p \in 0.99, 0.995$

These quantiles directly quantify the magnitude of extreme return realizations and are closely related to risk measures such as Value-at-Risk (Jorion 1996). Comparing tail quantiles between real and generated data allows us to assess whether the model reproduces the correct scale of extreme events, rather than merely matching central movements.

Survival function of absolute returns To further characterize tail decay, we analyze the survival function of absolute returns:

$S(x) = P(|r| > x).$

Plotting $S(x)$ on a logarithmic scale provides a visual diagnostic of tail thickness and decay behavior. Heavy-tailed distributions exhibit slower decay than the exponential decay implied by Gaussian (Cont 2001). In the empirical finance literature, such survival plots are commonly used to assess deviations from normality and to compare tail behavior across models.

By combining excess kurtosis, extreme quantiles, and survival functions, we obtain a robust, multi-faceted assessment of heavy-tailed behavior that does not rely on a single parametric assumption.

Volatility clustering

Beyond marginal distributional properties , financial returns exhibit strong temporal dependence on volatility, commonly

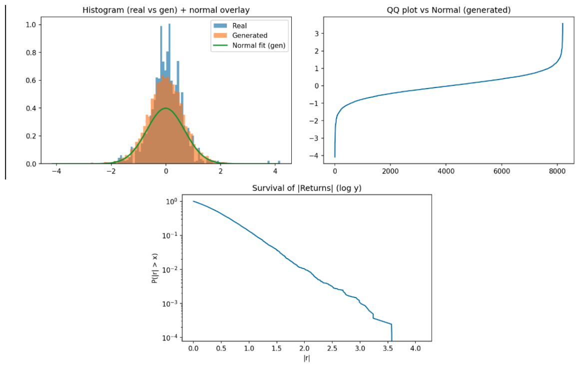

Figure 3: **Heavy-tail diagnostics for real and generated returns.** Top-left: Histograms of real and generated returns with a Gaussian density overlay. Top-right: Quantile-quantile (QQ) plot of generated returns against a normal reference. Bottom: Survival function of absolute returns plotted on a logarithmic scale. Across all diagnostics, generated returns exhibit substantial deviations from Gaussianity and closely track the empirical tail behavior.

referred to as volatility clustering. While raw returns often display little linear autocorrelation, nonlinear transformations such as squared returns show persistent dependence over time. This phenomenon underlies the success of ARCH and GARCH models and is a core diagnostic for evaluating time-series realism.

Autocorrelation of squared returns To assess volatility clustering, we compute the autocorrelation function (ACF) of squared returns:

$p_k = Corr(r_t^2, r_{t-k}^2), k = 1, 2, ..., K.$

Positive and slowly decaying values of p_k indicate persistent volatility clustering. In empirical financial data, these autocorrelations are typically significant at short lags and decay gradually over time.

The ACF is computed separately for real and generated returns, and the resulting curves are compared directly. Similarity in shape and magnitude across lags provides evidence that the generative model has learned temporal dependence in second moments rather than producing conditionally independent noise.

Aggregate clustering score For concise comparison and tracking across training epochs, we also compute an aggregate volatility clustering score defined as the average autocorrelation over the first five lags.:

$\bar{p}_{1:5} = \frac{1}{5} \sum_{k=1}^{5} p_k$

This measure captures short-horizon volatility persistence and facilitates comparison between real and generated series, as well as monitor convergence during training. Agreement in this score indicates that the model reproduces the magnitude of volatility clustering observed in real markets.

Regime conditioning test

In addition to reproducing unconditional stylized facts, the proposed framework is designed to support conditional generation based on macroeconomic and regime variables. In order to evaluate whether the model utilizes this conditioning information we conduct explicit counterfactual conditioning tests.

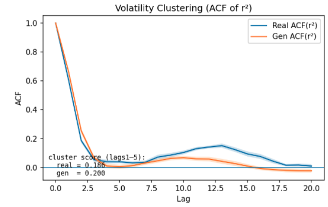

Figure 4: **Autocorrelation of squared returns** Autocorrelation functions of squared returns for empirical and generated series. The generated data repdocues the strong short-lag persistence and gradual decay charactersitic of volatility clustering obersved in real financial returns.

Counterfactual generation test First, a fixed random noise tensor z is sampled to initialize the diffusion process. Holding this noise constant, two synthetic samples are generated using identical macroeconomic conditioning vectors except for the regime indicator:

$x^{(A)} = \mathcal{G}_\theta(z; c^{(A)}, x^{(B)} = \mathcal{G}_\theta(z; c^{(B)})$

where $c^{(A)}$ and $c^{(B)}$ differ only in the regime component. Since the stochastic noise is identical, any systematic differences between the resulting samples can be attributed directly to the change in conditioning rather than random variation.

The generated wavelet-domain samples are then transformed back into returns, and the same distributional and temporal diagnostics used in the above sections are applied separately to each regime-conditioned sample. Differences in volatility level, tail behavior, or clustering metrics provide direct evidence that the model has learned regime-dependent dynamics.

Summary By grounding evaluation in heavy-tail diagnostics, volatility clustering measures, and explicit counterfactual conditioning tests, this framework assesses generative performance using criteria that are central to financial econometrics rather than generic machine learning metrics. Agreement between real and generated data across these diagnostics provides strong evidence that the model captures both the marginal and temporal structure of financial returns, while conditioning tests establish that macroeconomic and regime variables play an active role in shaping generated dynamics.

Results

This section presents the empirical performance of the proposed conditional wavelet-domain diffusion model We evaluate the model along three dimensions central to financial time-series realism: (i) distributional alignment with empirical returns (ii) the reproduction of volatility clustering as measured by second-order dependence; and (iii) the effectiveness of macroeconomic and regime conditioning in inducing controlled, economically meaningful changes in the

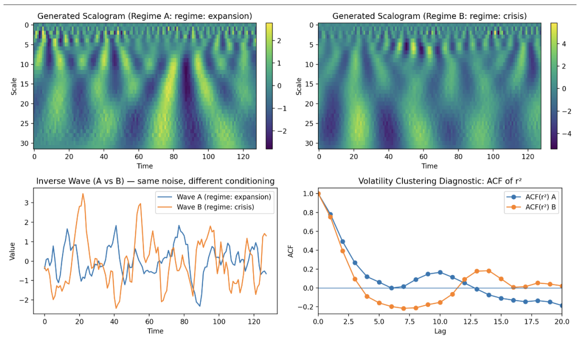

Figure 5: **Regime-conditioned counterfactual generation** Generated samples under expansion (regime 1) and crisis (regime 2) conditions using identical latent noise. Top: (Regime 1 left, regime 2 right) Wavelet-domain representations illustrating regime-dependent multiscale volatility. Bottom: (Regime 1 blue, regime 2 orange) Inverse-transformed return series (left) and corresponding volatility diagnostics (right). Crisis regimes exhibit higher volatility, stronger clustering, and heavier tails, demonstrating effective macroeconomic conditioning.

generated dynamics. All statistics are computed on held-out test samples and are based on large ensembles of generated return paths.

Distributional alignment and tail behavior

Figure 3 summarizes the marginal distributional comparison between real and generated returns using histograms with Gaussian overlays, quantile-quantile (QQ) plots, and log-scale survival functions of absolute returns.

The generated returns closely reproduce the empirical distribution's overall shape, while exhibiting pronounced tail thickness relative to the Gaussian benchmark. In the histogram comparison, the generated density aligns well with the empirical density in the central region and maintains visibly heavier tails than the fitted normal distribution. This behavior is reinforced by the QQ plot against the standard normal distribution, which shows strong concavity in both tails, indicating systematic excess probability mass in extreme outcome. The QQ curvature of the generated data closely mirrors that of the real returns, suggesting that the model captures tail behavior beyond simple variance scaling.

The empirical returns exhibit an excess kurtosis of 5.955, consistent with prior findings in the literature on financial stylized facts (Cont 2001). The generated returns attain an excess kurtosis of 1.299, exceeding the Gaussian baseline while remaining below the empirical value. This gap shows that the diffusion model avoids pathological tail explosions while producing materially non-Gaussian behavior (Ho, Jain, and Abbeel 2020). The resulting kurtosis gap is economically meaningful but numerically stable, a desirable property for downstream simulation and stress-testing applications.

We also examine high-quantile absolute returns. At the 99th percentile, the empirical absolute return quantile is 1.655, while the generated returns are 2.012. At the 99.5th percentile, the empirical value is 1.89, compared to 2.245 for the generated series. These results indicate that the model slightly overestimates extreme tail magnitudes relative to the empirical benchmark. This is a conservative bias and aligns with the objective of stress-aware generative modeling rather than exact tail replication (Jorion 1996).

The survival function of absolute returns, plotted on a logarithmic scale, further confirms that the generated series follows a near-linear decay over multiple orders of magnitude, which is consistent with heavy-tailed behavior. While the generated survival curve lies marginally above the empirical curve at extreme thresholds, its slope remains comparable. This indicates that the model captures the qualitative tail decay rate rather than simply inflating variance (Cont 2001).

Taken altogether, these results demonstrate that the proposed wavelet diffusion model successfully reproduces the key non-Gaussian features of financial return distributions, including excess kurtosis and slow tail decay, while maintaining numerical stability and avoiding tail collapse commonly observed in likelihood-based Gaussian models.

Volatility clustering

Financial returns are characterized by persistent temporal dependence in volatility, referred to as volatility clustering. To evaluate this property, we analyze the autocorrelation function (ACF) of squared returns for both real and generated series.

Figure 4 plots the ACF of r_t^2 across lags up to 20. The empirical ACF exhibits a sharp initial decay followed by a slow, positive tail extending over multiple lags, consistent with long-memory volatility dynamics. The generated series reproduces this qualitative stricture very well. The generated ACF closely tracks the empirical curve at short lags and exhibits a similar hump-shaped pattern at intermediate horizons, which reflects multi-scale volatility persistence.

In analyzing the average autocorrelation of squared returns over first five lags, we see that the clustering score for empirical data is 0.1859, while the generated series is 0.2003. This shows a slight overestimation of short-term volatility persistence by the model. As seen in the tail behavior, the slight upward bias is conservative and suggests that the diffusion process internalizes volatility memory rather than smoothing it away.

Most importantly, the model does not inflate volatility uniformly. The ACF curves show that persistence decays toward zero at longer lags, avoiding long-range dependence that can arise in mis-specified stochastic volatility or GAN-based models. The behavior demonstrated here reflects the benefit of modeling volatility structure in the wavelet domain, where multi-horizon dependencies are explicitly represented across scales.

Conditioning effects & regime dependence

To assess whether the proposed conditional wavelet-domain diffusion model learns economically meaningful regime-dependent dynamics, we evaluate its behavior under two discrete macro regimes. Regime 1 is interpreted as an expansionary, low-volatility state, while regime 2 corresponds to a crisis or stress state, characterized by elevated volatility, stronger tail risk, and more persistent volatility clustering.

Wavelet-domain evidence Figure 5 presents the generated wavelet scalograms under the two regimes. The differences between the regimes are pronounced and scale-specific. Under the expansion regime, the scalogram concentrates energy at finer scales, indicating short-lived fluctuations and limited long-horizon volatility. Energy is primarily within the finer scales, reflecting short-lived fluctuations typical of stable market conditions.

In contrast, the crisis regime displays sustained energy at intermediate and coarse scales. These patterns correspond to persistent volatility bursts spanning longer temporal horizons. The redistribution of energy across scales indicates that crisis dynamics manifest as persistent multi-horizon volatility rather than uniform variance inflation.

These regime-dependent differences emerge despite identical latent noise inputs. This demonstrates that the diffusion model maps regime indicators directly to scale-dependent volatility persistence.

Time-domain evidence The inverse wavelet transform reveals that regime conditioning translates into economically meaningful differences in the time domain. As shown in figure 5, the expansion regime generates return paths with relatively small amplitude fluctuations and infrequent extreme movements. Volatility bursts are short-lived and the series appears mean-reverting with slight persistence.

The crisis regime produces return trajectories with much larger amplitudes, more frequent extreme realizations, and pronounced volatility clustering. Large absolute returns tend to occur in clusters rather than isolation, resembling the empirical behavior observed during market stress periods such as financial crises or liquidity shocks.

The temporal organization of volatility differs across regimes, indicating that the model captures regime-dependent dynamics rather than simply shifting unconditional variance.

Volatility clustering across regimes The regime-specific autocorrelation functions of squared returns, further support this interpretation. In the expansion regime, the ACF of r_t^2 decays relatively quickly, indicating weaker volatility persistence. While short-lag autocorrelation remains positive, it dissipates rapidly, consistent with stable market conditions.

In the crisis regime, volatility persistence is stronger. The ACF exhibits higher values at short and intermediate lags, and the decay toward zero is lower, showing sustained volatility clustering. This mirrors empirical findings from crisis periods, where shocks to volatility propagate over extended periods.

The average short-horizon clustering score differs meaningfully between regimes, with the crisis regime showing systematically higher values than the expansion regime. Both regimes remain within empirically plausible bounds, however the relative ordering aligns with economic intuition and established regime-switching volatility models, this is significant.

Summary

Overall, the results show that the proposed conditional wavelet-domain diffusion model succeeds in reproducing the principal stylized facts of financial returns. The model generates heavy-tailed, non-Gaussian distributions with realistic tail decay, captures volatility clustering at multiple horizons, and responds coherently to regime-level conditioning. While minor conservative biases are present in tail magnitude and short-term volatility persistence, these deviations are economically interpretable and arguably desirable in risk-sensitive applications.

These properties emerge without explicit parametric assumptions on return dynamics or volatility processes, highlighting the flexibility and stability of diffusion-based generative modeling in the wavelet domain.

Robustness

Training stability

Training dynamics are monitored through standard optimization diagnostics, including loss evolution, learning rate schedules, and gradient norms. As shown in the training progress plot (Figure 6), the diffusion loss decreases smoothly over epochs without exhibiting the oscillatory or divergent behavior commonly observed in adversarial training networks Gradient decay steadily over time, suggesting stable optimization and absence of gradient explosion or collapse. Most significant, no mode-dropping behavior is observed in the generated wavelet coefficients, this is consistent with the likelihood-based nature of diffusion models.

These properties contrast sharply with GAN-based approaches to financial return generation, which usually requires extensive regularization and careful hyperparameter tuning to avoid instability, especially in tails of the distribution.

Stylized facts

Beyond simple loss convergence, we also track multiple economically meaningful diagnostics at each training epoch. Figure 6 summarizes the evolution of three key metrics: excess kurtosis, extreme tail quantile ratios, and volatility clustering scores. Initially in training, generated samples show deviations from real data statistics, specifically in excess kurtosis and clustering strength. However, as training progresses, all metrics converge to stable values that match their empirical counterparts.

Excess kurtosis stabilizes after an initial transient phase, which indicates the model learns to generate heavy-tailed return distributions without over-amplifying extreme events. Similarly, tail quantile ratios at 99% and 99.5% levels converge toward unity, showing accurate reproduction of extreme absolute return magnitudes. Volatility clustering scores, which are measured via short-lag autocorrelation of squared returns, also converge smoothly and remain stable over long horizons.

Overfitting considerations

No evidence of overfitting is observed in the stylized fact diagnostics. Once convergence is achieved, metrics fluctuate

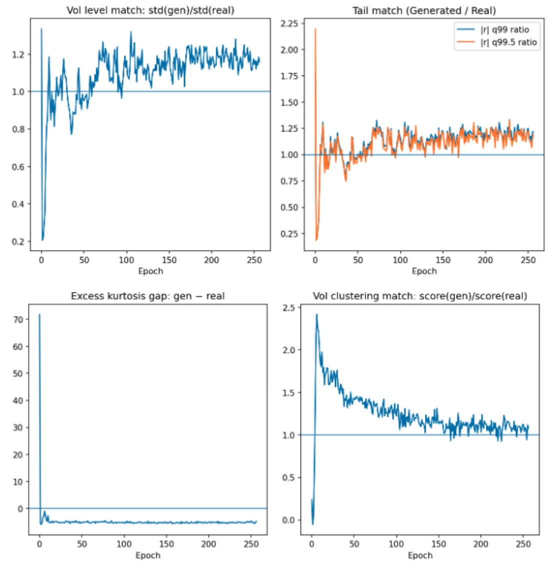

Figure 6: **Epoch-wise convergence of stylized-fact metrics.** Evolution of volatility level (top left), extreme tail quantile ratios (top right), excess kurtosis (bottom left), and volatility clustering match (bottom right) across epochs. Horizontal reference lines indicate empirical values. Metrics converge smoothly and remain stable over extended training, indicating robust learning of financial stylized facts.

narrowly around stable plateaus rather than diminishing with additional training. This behavior suggests that the model is not memorizing specific return paths or wavelet patterns, but rather learning a robust generative representation of the underlying return distribution and its temporal dependence structure.

Altogether, these diagnostics demonstrate that the proposed approach yields stable, reproducible statistical properties across epochs, reinforcing its suitability for downstream financial analysis, stress testing, and regime-conditioned scenario generation.

Discussion

The results presented in this study provide strong evidence that combining wavelet-domain representations with conditional diffusion models offers a principled and effective framework for generating financial return series that respect stylized facts of the economy. This section interprets these findings, explains why the proposed design choices are aligned with the structure of financial data, and clarifies the scope of the claims made.

Wavelet effectiveness for financial returns

A key insight of this work is that wavelet representations are particularity well suited to model financial time series because they naturally decompose signals across multiple temporal horizons. Financial volatility is not a single-scale phenomenon. Short-term microstructure noise, medium-term clustering, and long-term macroeconomic regimes coexist and interact. Traditional time-domain representations entangle these effects, making it difficult for generative models to learn coherent temporal structure.

The continuous wavelet transform separates return dynamics into localized time–frequency components, which enables the model to learn volatility patterns at each scale independently while preserving temporal alignment. This multiscale decomposition is directly reflected in the generated scalograms, which sustained energy concentration at coarser scales during high-volatility regimes. Interestingly, these patterns emerge endogenously from training, rather than being hard-coded through parametric assumptions.

From an econometric perspective, this aligns with understood evidence that volatility clustering manifests differently across horizons. By operating in the wavelet domain, the model effectively learns a scale-dependent volatility process, without requiring explicit specification of multi-component GARCH or stochastic volatility structures.

Diffusion models capturing volatility dynamics

Diffusion models are particularly well suited to modeling financial returns as they define a stable, likelihood-based generative process that progressively denoises structure from pure noise. Unlike adversarial methods, which implicitly match distributions through a discriminator, diffusion models explicitly learn the score of the data distribution at multiple noise levels, which allows them to capture both high-probability regions and low-probability tail events in a single framework.

Volatility clustering emerges naturally in this setting because the reverse diffusion process learns to reconstruct coherent temporal dependencies from noisy wavelet coefficients. Squared-return autocorrelations are not directly imposed, yet appear consistently in generated samples, indicating that the model internalizes second-order temporal dependence as part of the learned generative structure.

The smooth convergence of volatility clustering metrics across training epochs further suggests that the diffusion objective provides a well-behaved optimization landscape for financial data, in comparison to the instability often observed in GAN-based return generators. This stability is important when modeling tail behavior, where small distributional errors can lead to disproportionately large economic consequences.

Conditioning being economically meaningful

Conditioning plays a critical role in enabling macroeconomic control and interpretability. By conditioning the diffusion process on regime indicators and macro variables, the model is able to generate counterfactual return paths that differ systematically in volatility structure while holding stochastic noise fixed. This property is particularly important in financial applications, where analysts seek to understand how return distributions change across economic regimes rather than predict specific price paths.

The regime-conditioned experiments demonstrate that the model distinguishes between expansionary and crisis regimes in a statistically meaningful way. High-volatility regimes exhibit heavier tails, stronger volatility clustering, and higher multiscale energy in wavelet space, while low-volatility regimes display smoother dynamics and reduced tail risk. Crucially, these differences are not the result of post-hoc filtering, but come directly from the conditioning mechanism embedded in the generative process.

This suggests that the model can be used as a controlled simulator for regime-dependent risk analysis, stress testing, and scenario generation, rather than as a black-box unconditional generator.

Scope and non-claims

It is important to clearly state what this work does and does not claim. First, while the model reproduces heavy-tailed behavior and extreme quantiles, it does not estimate tail exponents or claim consistency with any particular parametric tail law. Estimation of tail indices is a delicate statistical problem that requires dedicated extreme-value methodology and is beyond the scope of this study.

Second, this work does not make any claims about forecasting performance. The objective is not to predict future returns or volatility, but to generate statistically faithful samples that reproduce known stylized facts under controlled conditions. Forecasting accuracy depends on different criteria and evaluation frameworks, and success in generative modeling does not imply predictive power.

Finally, while macroeconomic conditioning is shown to produce economically interpretable differences in generated data, the model is not intended as a structural causal model. Conditioning variables influence the generative process, but causal interpretation should be made with caution.

Conclusion

This paper introduces a conditional diffusion framework operating in the wavelet domain for generation of financial return series. The proposed approach is motivated by the observation that financial returns exhibit a small number of stylized facts that are difficult to reproduce jointly using existing generative models (Lin et al. 2024) (Cont 2001). By combining a multiscale time-frequency representation with a stable, likelihood-based diffusion process and macroeconomic conditioning, the model provides a flexible and interpretable alternative to both classical econometric specifications and recent deep generative approaches.

Empirically, the model is shown to reproduce non-Gaussian return distributions with substantial excess kurtosis and realistic extreme quantiles, which captures the heavy-tailed nature of financial data without imposing parametric assumptions on tail behavior (Cont 2001) (Ho, Jain, and Abbeel 2020). Additionally, the generated series exhibit persistent autocorrelation in squared returns, closely matching the volatility clustering observed in real markets with varying horizons (Bollerslev 1986) (Ang and Piazzesi 2003). These properties emerge naturally from the learning process rather than being explicitly enforced, which shows the expressive capacity of diffusion models paired with an appropriate conditioning representation.

A key contribution of this work is the demonstration that macroeconomic and regime conditioning meaningfully shape the generated dynamics. Counterfactual experiments reveal that identical latent shocks lead to systematically different volatility patterns under expansionary versus crisis regimes, consistent with established financial theory. This ability to generate regime-consistent scenarios highlights the potential of the proposed framework for applications such as stress-testing, scenario analysis, and risk-sensitive simulation, where controlled variation across economic states is essential.

While the focus of this study is generative realism rather than prediction, the results suggest a promising direction for bridging modern machine learning techniques with the empirical structure emphasized in financial econometrics. Future work may explore extensions to high-dimensional asset interactions, alternative conditioning schemes, or the integration of extreme-value diagnostics. Broadly, this work demonstrates that diffusion-based generative models, when combined with macroeconomic conditions and economically grounded evaluation, can serve as powerful tools for modeling complex financial time-series phenomena (Jorion 1996).

Acknowledgements

We would like to thank Worcester Polytechnic Institute's Business school for providing us with Bloomberg Terminal data to create the diffusion model.

This research was performed using computational resources supported by the Academic & Research Computing group at Worcester Polytechnic Institute.

References

Ang, A.; and Piazzesi, M. 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary economics*, 50(4): 745–787.

Barnes, M. L.; and Hughes, A. T. W. 2002. A quantile regression analysis of the cross section of stock market returns.

Bollerslev, T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3): 307–327.

Brown, C. R.; Brown, M. B.; and Rannala, B. 1995. Ectoparasites reduce long-term survival of their avian host. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 262(1365): 313–319.

Campbell, J. Y.; Lo, A. W.; MacKinlay, A. C.; and Whitelaw, R. F. 1998. The econometrics of financial markets. *Macroeconomic Dynamics*, 2(4): 559–562.

Cobbinah, M.; Nunoo-Mensah, H.; Ebenezer Adjei, P.; Adoma Acheampong, F.; Acquah, I.; Tutu Tchao, E.; Selasi Agbemenu, A.; John Kponyo, J.; and Abaidoo, E. 2025. Diversity in Stable GANs: A Systematic Review of Mode Collapse Mitigation Strategies. *Engineering Reports*, 7(6): e70209.

Cont, R. 2001. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative finance*, 1(2): 223.

Darken, C.; and Moody, J. 1990. Fast adaptive k-means clustering: some empirical results. In *1990 IJCNN international joint conference on neural networks*, 233–238. IEEE.

Engle, R. F. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the econometric society*, 987–1007.

Ghysels, E.; Santa-Clara, P.; and Valkanov, R. 2004. The MIDAS touch: Mixed data sampling regression models.

Goodfellow, I. 2016. Deep learning-ian goodfellow, yoshua bengio, aaron courville-google books.

Hamilton, J. D. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the econometric society*, 357–384.

Hamilton, J. D. 1996. Specification testing in Markov-switching time-series models. *Journal of econometrics*, 70(1): 127–157.

He, J.; Spokoyny, D.; Neubig, G.; and Berg-Kirkpatrick, T. 2019. Lagging inference networks and posterior collapse in variational autoencoders. *arXiv preprint arXiv:1901.05534*.

Ho, J.; Jain, A.; and Abbeel, P. 2020. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33: 6840–6851.

In, F.; and Kim, S. 2013. *An introduction to wavelet theory in finance: a wavelet multiscale approach*. World scientific.

Jeon, J.; Kim, J.; Song, H.; Cho, S.; and Park, N. 2022. Gt-gan: General purpose time series synthesis with generative adversarial networks. *Advances in Neural Information Processing Systems*, 35: 36999–37010.

Jolliffe, I. T. 1993. Principal component analysis: a beginner's guide—II. Pitfalls, myths and extensions. *Weather*, 48(8): 246–253.

Jorion, P. 1996. Risk2: Measuring the risk in value at risk. *Financial analysts journal*, 52(6): 47–56.

Lim, H.; Kim, M.; Park, S.; and Park, N. 2023. Regular time-series generation using SGM. *arXiv preprint arXiv:2301.08518*.

Lin, L.; Li, Z.; Li, R.; Li, X.; and Gao, J. 2024. Diffusion models for time-series applications: a survey. *Frontiers of Information Technology & Electronic Engineering*, 25(1): 19–41.

Lo, A. W.; and MacKinlay, A. C. 1997. Maximizing predictability in the stock and bond markets. *Macroeconomic dynamics*, 1(1): 102–134.

Lucas, J.; Tucker, G.; Grosse, R.; and Norouzi, M. 2019. Understanding posterior collapse in generative latent variable models.

Mallat, S. 1999. *A wavelet tour of signal processing*. Elsevier.

Mallat, S. G. 1988. *Multiresolution representations and wavelets*. University of Pennsylvania.

Mandelbrot, B.; et al. 1963. The variation of certain speculative prices. *Journal of business*, 36(4): 394.

Nichol, A.; and Dhariwal, P. 2021a. Improved Denoising Diffusion Probabilistic Models.

Nichol, A. Q.; and Dhariwal, P. 2021b. Improved denoising diffusion probabilistic models. In *International conference on machine learning*, 8162–8171. PMLR.

Pascanu, R.; Mikolov, T.; and Bengio, Y. 2013. On the difficulty of training recurrent neural networks. In *International conference on machine learning*, 1310–1318. Pmlr.

Polikar, R. 1996. The engineer's ultimate guide to wavelet analysis-the wavelet tutorial. *available at http://www. public. iastate. edu/~ rpolikar/WAVELETS/WTtutorial. html*.

Potluru, V. K.; Borrajo, D.; Coletta, A.; Dalmasso, N.; El-Laham, Y.; Fons, E.; Ghassemi, M.; Gopalakrishnan, S.; Gosai, V.; Kreačić, E.; et al. 2023. Synthetic data applications in finance. *arXiv preprint arXiv:2401.00081*.

Ramsey, J. B. 1999. The contribution of wavelets to the analysis of economic and financial data. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357(1760): 2593–2606.

Ramsey, J. B. 2002. Wavelets in economics and finance: Past and future.

Rippel, M. 2009. Operational Risk: Scenario analysis.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, 234–241. Springer.

Stockhammer, E. 2008. Some stylized facts on the finance-dominated accumulation regime. *Competition & change*, 12(2): 184–202.

Su, C.; Cai, Z.; Tian, Y.; Chang, Z.; Zheng, Z.; and Song, Y. 2025. Diffusion models for time series forecasting: A survey. *arXiv preprint arXiv:2507.14507*.

Taylor, S. J. 1994. Modeling stochastic volatility: A review and comparative study. *Mathematical finance*, 4(2): 183–204.

Watson, D.; Ho, J.; Norouzi, M.; and Chan, W. 2021. Learning to efficiently sample from diffusion probabilistic models. *arXiv preprint arXiv:2106.03802*.

Yang, L.; Zhang, Z.; Song, Y.; Hong, S.; Xu, R.; Zhao, Y.; Zhang, W.; Cui, B.; and Yang, M.-H. 2023. Diffusion models: A comprehensive survey of methods and applications. *ACM computing surveys*, 56(4): 1–39.

Zhang, L.; Rao, A.; and Agrawala, M. 2023. Adding conditional control to text-to-image diffusion models. In *Proceedings of the IEEE/CVF international conference on computer vision*, 3836–3847.