ELSEVIER

Journal of Econometrics 105 (2001) 27–58

JOURNAL OF Econometrics

www.elsevier.com/locate/econbase

Forecasting multifractal volatility

Laurent Calvet[a], Adlai Fisher[b,*]

[a]*Department of Economics, Harvard University, Cambridge, MA 02138, USA*
[b]*Faculty of Commerce, University of British Columbia, 2053 Main Mall, Vancouver, BC, Canada V6T 1Z2*

Abstract

This paper develops analytical methods to forecast the distribution of future returns for a new continuous-time process, the Poisson multifractal. The process captures the thick tails, volatility persistence, and moment scaling exhibited by many financial time series. It can be interpreted as a stochastic volatility model with multiple frequencies and a Markov latent state. We assume for simplicity that the forecaster knows the true generating process with certainty but only observes past returns. The challenge in this environment is long memory and the corresponding infinite dimension of the state space. We introduce a discretized version of the model that has a finite state space and an analytical solution to the conditioning problem. As the grid step size goes to zero, the discretized model weakly converges to the continuous-time process, implying the consistency of the density forecasts. © 2001 Published by Elsevier Science S.A.

JEL classification: C22; C53; F31

Keywords: Forecasting; Long memory; Multiple frequencies; Stochastic volatility; Weak convergence

1. Introduction

This paper develops analytical methods to forecast the distribution of future returns for a new class of continuous-time processes, Poisson multifractals. The model parsimoniously captures the thick tails and volatility persistence

* Corresponding author. Tel.: +1-604-822-8331; fax: +1-604-822-4695.
E-mail address: adlai.fisher@commerce.ubc.ca (A. Fisher).

PII: S0304-4076(01)00069-0

exhibited by many financial time series. It also provides a fully stationary version of the multifractal model introduced in Mandelbrot et al. (1997), which contains residual effects of a grid-based construction. We model volatility as the multiplicative product of an infinite sequence of random functions. Each function contains Poisson arrivals of increasing frequency, and frequencies progress geometrically in the construction. This ensures that volatility clustering exists on a wide frequency range, consistent with the intuition that economic factors such as technology shocks, business and earnings cycles, and liquidity shocks have different time scales.

While the model incorporates multiple volatility components with heterogeneous frequencies, all components are drawn from the same marginal distribution. This invariance provides parsimony and can be viewed as a pragmatic approach to multi-frequency modeling. Another consequence is that moments of the price process have scaling properties that are characteristic of multifractals. We propose that such patterns may arise in fully rational equilibrium models, either exogenously from multifractal shocks, or endogenously from market incompleteness or informational cascades. Since volatility components are not directly observable in our model, past data is informative on the current latent state. Given the current level of volatility, forecasts may therefore differ considerably depending on past history.

We first introduce the continuous-time version of the model, which has an infinite state space. The model compounds a standard Brownian motion with a random Poisson time-deformation process. The construction implies that returns are uncorrelated and prices follow a semi-martingale, which precludes arbitrage in a standard two-asset economy. Squared returns have long memory, and the highest finite moment of returns is permitted to have any value greater than two. The flexible tail behavior originates in intermittent bursts of volatility, and does not require jumps or separate modelling of the conditional distribution of returns.

To facilitate forecasting, we introduce a discretized version of the process that has a Markov structure and a finite state space. A recursive algorithm allows us to calculate the current probabilities of the volatility state conditional on past data and initial beliefs. A second algorithm computes multi-step forecasts of the return distribution at future dates. We show that the discretized process weakly converges to the continuous-time process as the grid step size goes to zero, which implies the consistency of the forecasting methodology.

The Poisson multifractal approach has several features that distinguish it from forecasting in other long-memory models. In FIGARCH, the state variables relevant for forecasting consist of all past returns, and the conditioning problem reduces to choosing how many observations to use. Multi-step forecasting is problematic because of the large state space, and is usually accomplished by simulation. For standard long-memory stochastic volatility models (LMSV), the problem is further complicated since the relevant state variables are past volatility levels, which are unobserved. The forecaster thus must integrate conditional forecasts over each possible set of past volatilities. The multifractal model simplifies forecasting because it greatly reduces the volatility state space. In FIGARCH and LMSV models, n state variables are needed to model frequencies of size as low as $1/n$. The multifractal model only requires $\log_b n$ state variables to capture the same frequency range, where b is a constant of the model.

Section 2 provides a brief review of earlier work. Section 3 introduces a new, grid-free multifractal measure in continuous time. Section 4 defines the new financial model and presents its economic intuition. Section 5 introduces a discretized version of the process, develops the forecasting algorithm, and demonstrates the consistency of the forecasting methodology.

2. A review of the MMAR

This section briefly reviews the multifractal model of asset returns (MMAR) introduced in Mandelbrot et al. (1997). The MMAR is a continuous-time model that captures the outliers, volatility persistence and moment scaling of many financial time series (Calvet and Fisher, 1999). It is constructed by compounding a Brownian motion $B(t)$ with a random function $\theta(t)$ called *trading time*:

$$\ln P(t) - \ln P(0) = B[\theta(t)].$$

The main contribution of the MMAR is to specify the function $\theta(t)$ as the cumulative distribution function (c.d.f.) of a multifractal measure μ, a concept which we now recall.

2.1. The binomial measure

Multifractal measures can be built by iterating a simple procedure called a *multiplicative cascade*. We first present one of the simplest examples, the binomial measure on $[0,1]$. Consider the uniform probability measure μ_0 on the unit interval, and two positive numbers m_0 and m_1 adding up to 1. In the first step of the cascade, we define a measure μ_1 by uniformly spreading the mass m_0 on the *left* subinterval $[0,\frac{1}{2}]$, and the mass m_1 on the *right* subinterval $[\frac{1}{2},1]$. The density of μ_1 is a step function.

In the second stage of the cascade, we split $[0,\frac{1}{2}]$ into two subintervals of equal length. The left subinterval $[0,\frac{1}{2}]$ is allocated a fraction m_0 of $\mu_1[0,\frac{1}{2}]$, while the right subinterval $[\frac{1}{4},\frac{1}{2}]$ receives a fraction m_1. Applying a similar procedure to $[\frac{1}{2}, 1]$, we obtain a measure μ_2 such that

$$\mu_2[0, \tfrac{1}{4}] = m_0 m_0, \quad \mu_2[\tfrac{1}{4}, \tfrac{1}{2}] = m_0 m_1,$$
$$\mu_2[\tfrac{1}{2}, \tfrac{3}{4}] = m_1 m_0, \quad \mu_2[\tfrac{3}{4}, 1] = m_1 m_1.$$

Iteration of this procedure generates an infinite sequence (μ_k) that weakly converges to the binomial measure μ. Like many multifractals, the binomial is a continuous but singular probability measure that has no density and no point mass. Moreover since $m_0 + m_1 = 1$, each stage of the construction preserves the mass of split dyadic intervals.

This construction is easily generalized. For instance, we can uniformly split intervals into an arbitrary number $b \geqslant 2$ of cells at each stage of the cascade. Subintervals, indexed from left to right by β $(0 \leqslant \beta \leqslant b-1)$, then receive fractions $m_0, \ldots, m_{b-1}$ of the measure. Another extension randomizes the allocation of mass between subintervals. The fraction, or *multiplier*, of each cell is then a discrete random variable M_β taking values $m_0, m_1, \ldots, m_{b-1}$ with probabilities $p_0, \ldots, p_{b-1}$. This generates a *random* multifractal measure.

2.2. Multiplicative measures

We can also consider non-negative multipliers M_β $(0 \leqslant \beta \leqslant b-1)$ with *arbitrary* distributions. Assume, for simplicity, that multipliers are identically distributed ($M_\beta \overset{\mathrm{d}}{=} M \ \forall \beta$), and that multipliers defined at different stages of the construction are independent. The limit *multiplicative measure* is called *conservative* when mass is conserved exactly at each stage: $\sum M_\beta \equiv 1$, and *canonical* when it is only preserved on average: $\mathbb{E}(\sum M_\beta) = 1$ or equivalently $\mathbb{E}M = 1/b$.

The moments of multiplicative measures have interesting scaling properties. To show this, first consider the generating cascade of a *conservative* measure μ. Stage 1 uniformly splits the unit interval into cells of length b^{-1}, and allocates random masses $M_0, \ldots, M_{b-1}$ to each cell. Similarly, the measure of a b-adic cell of length $\Delta t = b^{-k}$, starting at $t = \overline{0.\eta_1 \cdots \eta_k} = \sum \eta_i b^{-i}$, is the product of k multipliers:

$$\mu(\Delta t) = M_{\eta_1} M_{\eta_1, \eta_2} \cdots M_{\eta_1, \ldots, \eta_k}. \tag{2.1}$$

Since multipliers defined at different stages of the cascade are independent, we infer that $\mathbb{E}[\mu(\Delta t)^q] = [\mathbb{E}(M^q)]^k$ or equivalently

$$\mathbb{E}[\mu(\Delta t)^q] = (\Delta t)^{\tau(q)+1}, \tag{2.2}$$

where $\tau(q) = -\log_b \mathbb{E}(M^q) - 1$. The moment of an interval's measure is thus a power function of the length Δt. This important scaling rule characterizes multifractals.

Scaling relation (2.2) easily generalizes to a *canonical* measure μ. Let Ω denote the random mass of the unit interval. The measure of a b-adic cell is of the form $\mu(\Delta t) = M_{\eta_1} M_{\eta_1,\eta_2} \cdots M_{\eta_1,\ldots,\eta_k} \Omega_{\eta_1,\ldots,\eta_k}$, where $\Omega_{\eta_1,\ldots,\eta_k} \overset{d}{=} \Omega$, and thus

$$\mathbb{E}[\mu(\Delta t)^q] = \mathbb{E}(\Omega^q)(\Delta t)^{\tau(q)+1}.$$

Because the construction relies on a b-adic grid, this relation only holds when the length Δt is of the form b^{-k}. For this reason, the multiplicative measures constructed in this section are called *grid-bound*. Section 3 will show how to construct grid-free measures.

2.3. Multifractal processes

We now recall the construction of the MMAR (see Fig. 1). Consider the price of a financial asset $P(t)$ on a bounded interval $[0, T]$, and define the *log-price* process

$$X(t) \equiv \ln P(t) - \ln P(0).$$

We model $X(t)$ by compounding a Brownian motion B with a random multifractal trading time $\theta(t)$:

$$X(t) \equiv B[\theta(t)],$$

where $\theta(t)$ is the c.d.f. of a multiplicative measure μ. We also assume for simplicity that the processes $\{B(t)\}$ and $\{\theta(t)\}$ are independent, a hypothesis that could be weakened in order to model an asymmetric relationship of volatility with positive and negative price changes.

The process $X(t)$ is a martingale with respect to its natural filtration. The price process $P(t)$ is therefore a semi-martingale, which allows use of stochastic integration to calculate gains from trade. Thus, there is no arbitrage in a two asset economy containing a risky asset with multifractal price $P(t)$, and a riskless asset with a constant rate of return.

The MMAR has long memory in volatility, as discussed in Calvet and Fisher (1999). For any integrated process Z, the *autocovariance in levels*

$$\delta_Z(t, q) = Cov(|Z(a, \Delta t)|^q, |Z(a + t, \Delta t)|^q)$$

quantifies the dependence in the size of increments $Z(a, \Delta t) = Z(a + \Delta t) - Z(t)$. It is well defined when $\mathbb{E}|Z(a, \Delta t)|^{2q}$ is finite. For a fixed q, we say that the process has *long memory in the size of increments* if the autocovariance $\delta_Z(t, q)$ is hyperbolic in t when $t/\Delta t \to \infty$. This concept is independent of the choice of q when the process Z is multifractal. Under these definitions, both trading time $\theta(t)$ and the process $X(t)$ have long memory in the size of increments.

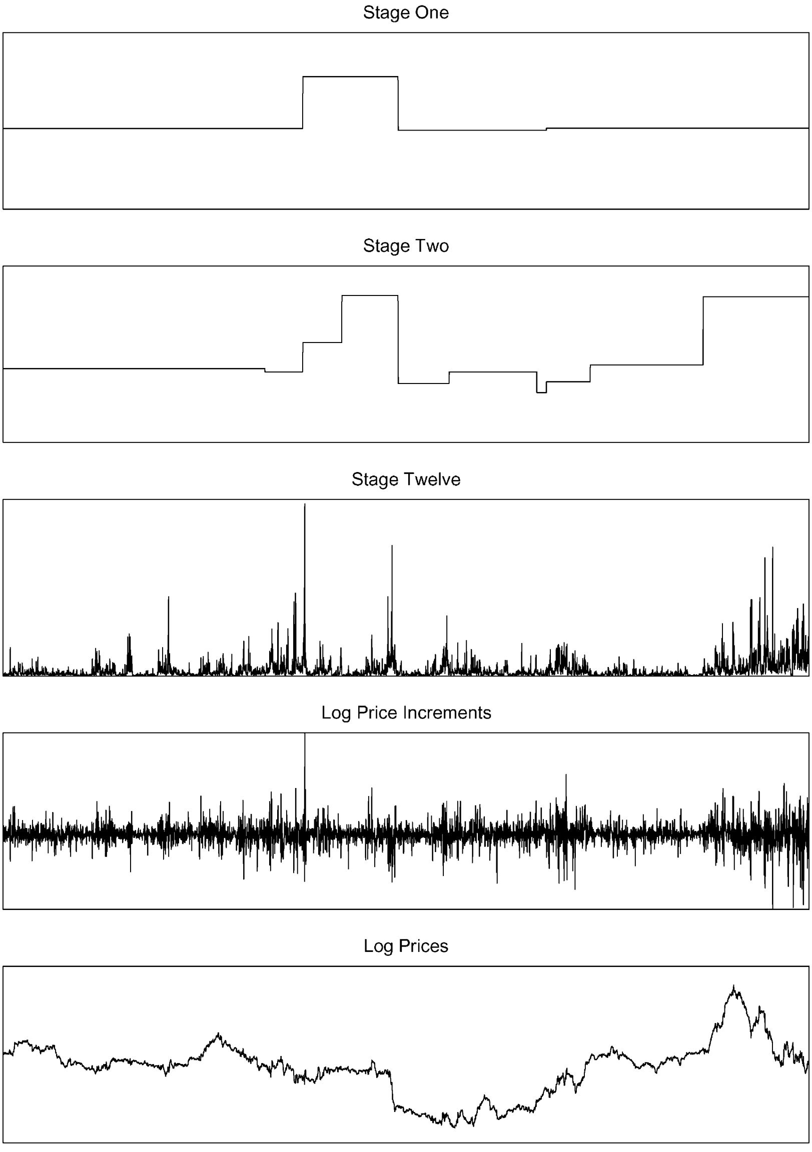


Fig. 1. Construction of a grid-free multifractal process. The first three panels show stages in the construction of a grid-free limit lognormal measure with parameter $\lambda = 1.09$. The final two panels illustrate the corresponding simulated log price increments and log prices, which are obtained by compounding the c.d.f. of the stage 12 measure with a Gaussian white noise. The return series shows long memory volatility clustering and outliers that are produced by intermittent bursts of extreme volatility. The construction fully randomizes the timing of volatility innovations, and the resulting return process is strictly stationary.

3. The Poisson multifractal

This section introduces a grid-free multifractal measure in continuous time. The main innovation consists of randomizing the instants at which the multipliers are changing. All proofs are given in the appendix.

3.1. Construction

We define a Poisson multifractal measure on interval $[0,T]$ as the limit of a multiplicative cascade. In the first stage, consider a sequence $\{T_n\}_{n=1}^{\infty}$ of independent random variables with identical exponential density $f(x;\lambda)=\lambda\exp(-\lambda x)$. The elements T_n help randomize the time intervals over which the multipliers are constant. Denoting by $N=\max\{n\colon \sum_{i=1}^{n} T_i < T\}$ the number of arrivals on $[0,T]$, we define the first stage random instants

$$t_n = \begin{cases} 0 & \text{if } n=0, \\ \sum_{i=1}^{n} T_i & \text{if } 1 \leqslant n \leqslant N, \\ T & \text{if } n = N+1. \end{cases}$$

Since the random variables $\{T_n\}_{n=1}^{\infty}$ have exponential distributions, the probability of arrival at any instant t is independent of past history. This will allow us to reinterpret the construction as a Markov model of volatility.

The intervals $\{I_n=[t_n,t_{n+1}]\colon 0 \leqslant n \leqslant N\}$ form a random partition of $[0,T]$. Let $\ell([t,t'])=t'-t$ denote the length of a given interval. The measure μ_1 is defined by drawing independent random multipliers M_n and uniformly spreading within each interval the mass

$$\mu_1(I_n)=M_n\,\ell(I_n)/T.$$

To obtain a non-degenerate limit, we impose that mass be preserved *on average* at each stage of the cascade: $\mathbb{E}\mu_1[0,T]=1$. Since $\mathbb{E}[\sum_{n=0}^{N} M_n\,\ell(I_n)|\{T_n\}_{n=1}^{\infty}]=(\mathbb{E}M)T$, we infer that

$$\mathbb{E}\mu_1[0,T]=\mathbb{E}\left[\sum_{n=0}^{N} M_n\,\ell(I_n)\right]\Big/ T=\mathbb{E}M.$$

The preservation of mass is thus equivalent to $\mathbb{E}M=1$.

Given a real number $b>1$, we then construct a sequence of measures $(\mu_k)_{k=1}^{\infty}$ in which the highest frequency of arrival $b^k\lambda$ progresses geometrically with k. The recursion procedure is defined as follows. Given the measure

μ_{k-1}, consider $j=(j_1,\dots,j_{k-1})$ and the stage $k-1$ interval

$$I_j=[t_{j_1,\dots,j_{k-1}};\ t_{j_1,\dots,j_{k-1}+1}]$$

with evenly spread mass $\mu_{k-1}(I_j)$. Let $\{T_n^j\}_{n=1}^{\infty}$ denote a sequence of independent random variables with identical exponential density $f(x;b^k\lambda)$. The sequence $\{T_n^j\}_{n=1}^{\infty}$ is assumed to be independent of all the random variables defined up to stage k. To ensure that the new arrivals belong to I_j,[1] we define the integer

$$N^j=\max\left\{n\colon \sum_{i=1}^{n} T_i^j<\ell(I_j)\right\}$$

and the arrival times

$$t_{j,n}=\begin{cases} t_j & \text{if } n=0,\\ t_j+\displaystyle\sum_{i=1}^{n}T_i^j & \text{if } 1\leqslant n\leqslant N^j,\\ t_{j_1,\dots,j_{k-1}+1} & \text{if } n=N^j+1.\end{cases}$$

On each subinterval $I_{j,n}=[t_{j,n};t_{j,n+1}]$, we then randomly draw a multiplier $M_{j,n}$ and uniformly spread the mass

$$\mu_k(I_{j,n})=(M_{j_1}\cdots M_{j_1,\dots,j_{k-1}}\ M_{j,n})\,\ell(I_{j,n})/T. \tag{3.1}$$

We observe that $\mathbb{E}[\mu_k(I_j)|\mu_{k-1}(I_j)]=\mu_{k-1}(I_j)$. The martingale convergence theorem then implies that $\{\mu_{k'}(I_j)\}_{k'=k}^{\infty}$ converges to a limit $\mu(I_j)$ when $k\to\infty$. It is useful to introduce

Condition 1. $\mathbb{E}(M^2)<b$.

We show in Appendix A:

Theorem 1. Under Condition 1, *the sequence* (μ_k) *weakly converges to a measure* μ *defined on the space* $\mathscr{C}$ *of continuous functions.*

This theorem ensures that the limit μ is well defined.[2] We call it a *Poisson multifractal measure.* Its total mass is random and satisfies by construction: $\mathbb{E}\mu[0,T]=1$.

An important advantage of Poisson multifractals lies in their implicit Markov structure, which we now briefly sketch. Given $t\in[0,T)$ and an integer $k\geqslant 1$, consider the interval $[t_{j_1,\dots,j_k};t_{j_1,\dots,j_k+1})$ containing t, and denote

[1] An alternative specification, which we do not explore in this paper, assumes the independence of arrival times defined at different stages of the cascade. The forecasting results of Section 5 immediately extend to this alternative construction.

[2] Note that Condition 1 is sufficient but not necessary to guarantee convergence.

by $M_{k,t}$ the value of the stage k multiplier $M_{j_1,\ldots,j_k}$ prevailing at date t. We can stack the value of all date t multipliers into the infinite sequence $Z_t = \{M_{k,t}\}_{k=1}^{\infty}$. Because arrival times follow Poisson processes, it easy to check that Z_t is a Markov process on the space $\mathbb{R}^{\infty}$. Moreover for every $s \geqslant t$, the distribution of the trading time increment $\theta(s) - \theta(t)$ conditional on $\mathscr{F}_t = \{(t_{j_1,\ldots,j_k}, M_{j_1,\ldots,j_k})_{t_{j_1,\ldots,j_k} \leqslant t}\}$ only depends on Z_t. We can thus interpret Z_t as the volatility state vector. The benefits of these insights will become apparent in Section 5 when we consider discretized versions of Poisson multifractals.

3.2. Properties

The mass $\mu[0,T]$ defines a random variable $\Omega(T,\lambda)$, whose distribution depends on the parameters λ and T. We observe that $\mathbb{E}\Omega(T,\lambda) = 1$ and

Proposition 1. The mass of a subinterval $[0,t]$ *satisfies*

$$\mu[0,t] \stackrel{\mathrm{d}}{=} \frac{t}{T}\Omega(t,\lambda)$$

for all $t \leqslant T$.

This proposition is very useful to analyze the scaling properties of the measure. We show in Appendix A:

Proposition 2. The random variable $\Omega(T,\lambda)$ *satisfies the invariance relation*

$$\Omega(T,\lambda) \stackrel{\mathrm{d}}{=} \Omega(1,\lambda T) \tag{3.2}$$

for all T *and* λ.

This establishes that the distribution of the random mass $\Omega(T,\lambda)$ depends only on λT. It is thus convenient to define $\Omega(\lambda) \equiv \Omega(1,\lambda)$ and henceforth use a single index to represent it. Moreover since $\mu[0,1] = \sum_{j=0}^{N} \mu[t_j, t_{j+1}]$ and $\mu[t_j,t_{j+1}] = [(t_{j+1}-t_j)/T] M_j \Omega_j [b\lambda(t_{j+1}-t_j)]$, we infer

$$\Omega(\lambda T) \stackrel{\mathrm{d}}{=} \sum_{j=0}^{N} \frac{t_{j+1}-t_j}{T} M_j \Omega_j [b\lambda(t_{j+1}-t_j)]. \tag{3.3}$$

This generalizes the invariance relation $\Omega = \sum M_\beta \Omega_\beta$ satisfied by the random mass of a grid-bound measure.

We now examine the moment properties of the Poisson multifractal.

Proposition 3. If there exists $\lambda > 0$ *such that* $\mathbb{E}[\Omega(\lambda)^q] < \infty$, *then the qth moment* $\mathbb{E}[\Omega(\lambda')^q]$ *is finite for all* $\lambda' \in (0,\infty)$.

The existence of moments is thus independent of both the time domain of the measure and the lowest frequency of the construction. In particular, the critical moment $q_{\text{crit}} = \sup\{q\colon \mathbb{E}[\Omega(\lambda)^q] < \infty\}$ does not depend on λ.

When the frequency λ is close to 0, the first stage multiplier is constant on the interval $[0,T]$ with high probability, implying $\Omega(\lambda) \approx M\Omega(b\lambda)$ and thus $\mathbb{E}[\Omega(\lambda)^q] \approx \mathbb{E}(M^q)\mathbb{E}[\Omega(b\lambda)^q]$. This suggests that $\mathbb{E}[\Omega(\lambda)^q]$ asymptotically behaves like a power function when $\lambda \to 0$. Given two functions f and g, it is convenient to use the notation $f(\lambda) \sim g(\lambda)$ when $f(\lambda)/g(\lambda) \to 1$. This leads to

Proposition 4. The qth moment of the random mass $\Omega(\lambda)$ satisfies

$$\mathbb{E}[\Omega(\lambda)^q] \sim c_q \lambda^{\tau_\theta(q)+q+1} \quad \text{as } \lambda \to 0,$$

where $\tau_\theta(q) = -\log_b \mathbb{E}(M^q) - q - 1$ and c_q is a positive constant.

We call $\tau_\theta(q)$ the *scaling function* of the measure.[3] It is concave by Hölder's inequality.

We now infer from Proposition 1 the scaling behavior of a random measure μ associated with a *fixed* parameter λ.

Corollary 1. For any $q \geqslant 0$, the qth moment of the measure satisfies

$$\mathbb{E}(\mu[0,t]^q) \sim c_{\lambda,q} t^{\tau_\theta(q)+1} \quad \text{as } t \to 0,$$

where $c_{\lambda,q}$ is a positive constant.

Estimation techniques based on the scaling of unconditional moments may thus be applied to grid-free measures. In addition, the discretized processes developed in this paper permit likelihood based estimation, which will be explored in future work.

4. The financial model

We formalize in this section the construction of the financial model. Consider the price of a financial asset $P(t)$ defined on the clock time interval $[0,T]$. We model the log-price $X(t) = \ln P(t) - \ln P(0)$ by compounding a Brownian motion with the c.d.f. of a Poisson multifractal:

[3] The multiplier, respectively, satisfies the normalization conditions $\mathbb{E}M = 1/b$ and 1 in the grid-bound and grid-free constructions. This leads to the slightly different relations between $\tau_\theta(q)$ and $\mathbb{E}(M^q)$ obtained in Section 2.2 and Proposition 4.

Assumption 1. $X(t)$ is a compound process

$$X(t) \equiv B[\theta(t)],$$

where $B(t)$ is a Brownian motion, and $\theta(t)$ is a stochastic trading time.

Assumption 2. Trading time $\theta(t)$ is the c.d.f. of a Poisson multifractal measure μ defined on $[0, T]$.

Assumption 3. The processes $\{B(t)\}$ and $\{\theta(t)\}$ are independent.

The resulting process is called the Poisson multifractal model (PMM).[4] It improves on the MMAR in two related ways. First, the PMM is a grid-free process in which volatility components switch at random instants and not at predetermined points of time. Second, the new model can be interpreted as a stochastic volatility model with a Markov latent vector, as discussed in Section 3.1. This facilitates forecasting and helps to provide economic intuition.

In addition to these improvements, the PMM shares many of the MMAR's appealing properties. Consider, for instance, the scaling function $\tau_\theta(q) = -\log_b \mathbb{E}(M^q) - q - 1$ introduced in Section 3.2, and let $\tau(q) \equiv \tau_\theta(q/2)$. By Corollary 1, returns satisfy the asymptotic scaling rule

$$\mathbb{E}[|X(t)|^q] \sim C_q t^{\tau(q)+1} \quad \text{as } t \to 0.$$

Using Monte Carlo simulations, we show in Fig. 2 that this property holds remarkably well for finite time increments. The Poisson multifractal model is therefore consistent with the moment scaling exhibited by many financial time series, including exchange rates and equities (Fisher et al., 1997; Calvet and Fisher, 1999).

The PMM has an autocorrelation structure that is very similar to the MMAR. The compound process $X(t)$ is a martingale because the increments of the Brownian motion B have unpredictable signs. We thus infer that the price $P(t) = \exp[X(t)]$ is a semi-martingale,[5] which has important consequences for arbitrage. Consider, for instance, the *two asset economy* consisting of the multifractal security with price $P(t)$, and a riskless bond with constant rate of return r. Following Harrison and Kreps (1979), we can analyze if arbitrage profits can be made by frequently rebalancing a portfolio of these two securities. Since $P(t)$ has the semi-martingale property, we infer

[4] An immediate extension would consider processes of the form $X(t) = B_H[\theta(t)]$, where B_H is a fractional Brownian motion and $\theta(t)$ is the c.d.f. of a Poisson multifractal.

[5] We refer the reader to Dothan (1990) and Durrett (1996) for excellent discussions of semi-martingales.

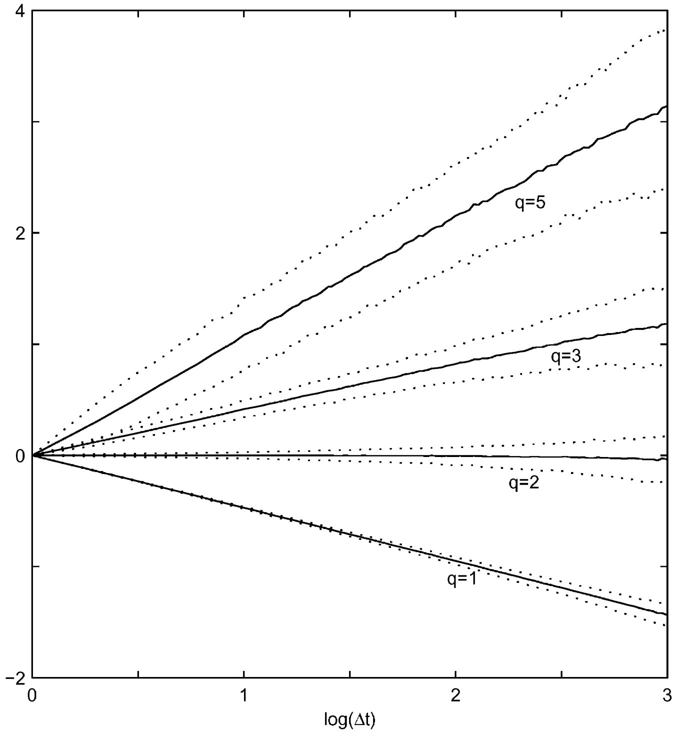
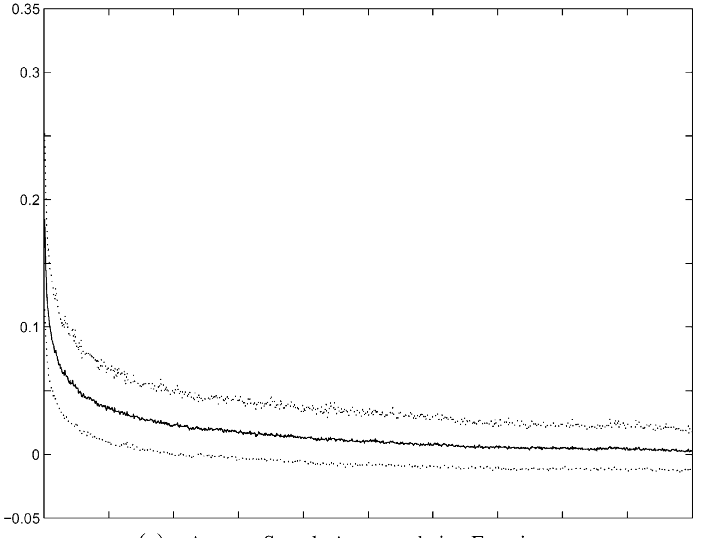


Fig. 2. Moment scaling in grid-free multifractals. This figure shows the scaling of moments $q=1,2,3,5$ for the grid-free limit lognormal process with parameter $\lambda=1.09$. The vertical axis corresponds to the logarithm of the sample moment times the number of increments with size Δt in a sample of length 20,000. The solid line shows these moments averaged over 500 independent simulations of length 20,000. The 20th and 80th percentiles are plotted in dotted lines. For convenience, each line is vertically displaced to begin at zero. All of the moments demonstrate the predicted asymptotic scaling.

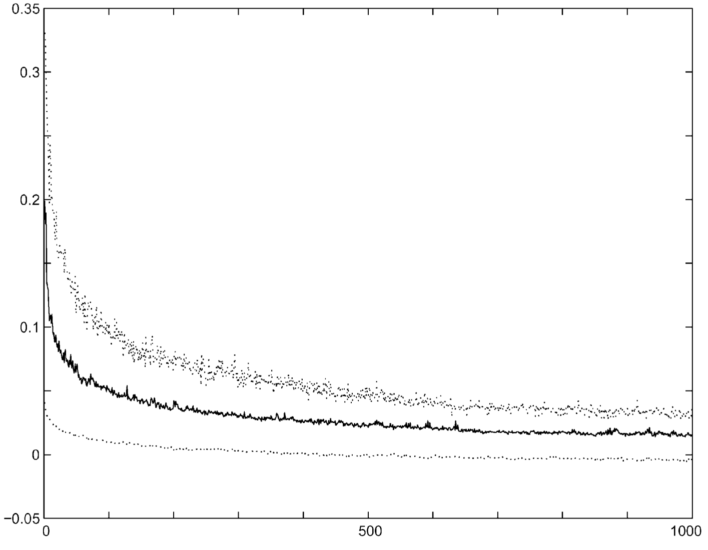
Proposition 5. There are no arbitrage opportunities in the two asset economy.

This suggests that future research may seek to embed the PMM in standard financial models. Since the price $P(t)$ is a semi-martingale, stochastic integration can be used to calculate the gains from trading multifractal assets, which in future work will greatly help us develop portfolio selection and option pricing theory.

Long memory in volatility is another important feature of the data that is captured by the PMM. In Section 2.3, we introduced a quantitative description of long memory for a continuous-time process defined on a bounded time



(a) Average Sample Autocorrelation Function



(b) Estimated Population Autocorrelation Function interval. Following this approach, we show in Fig. 3 that autocovariances decline hyperbolically in simulated PMM data. The PMM is thus consistent with the increasing evidence that financial series display long memory in squared returns (Taylor, 1986; Ding et al., 1993; Dacorogna et al., 1993) (Fig. 4).

Long memory can also be characterized by the property that the longest apparent cycle has approximately the same length as the interval of definition. Fig. 1 illustrates that such is indeed the case with PMM volatility when the lowest frequency λb is of order $1/T$. More generally, given a forecasting problem over the time interval $[0,T]$, we can set the parameter λb equal to the lowest frequency shock we want to consider. This choice is very similar, in practice, to the truncation problem encountered in forecasting fractionally integrated processes such as ARFIMA (Baillie, 1996).

4.1. Economic intuition

The PMM contains volatility components $M_{k,t}$ with heterogeneous frequencies. Thus, the lowest frequencies might correspond to business cycles and technological shocks, while other frequencies could correspond to earnings cycles or short-lived liquidity shocks. This closely captures the economic intuition that different types of volatility shocks have different degrees of persistence. In contrast, most standard models impose that all volatility innovations are statistically identical, and thus use a single decay function to fit many different types of shocks.

A volatility model with many frequencies could quickly become unwieldy without additional structure. The multifractal approach builds on the organizing principle that volatility shocks have no favored time scale. If we take the

Fig. 3. Long memory in the squared returns of a grid-free multifractal model. Each panel is based on 300 independent simulations of a grid-free limit lognormal process with parameter $\lambda = 1.09$. Each simulation has a length of 10,000 increments. The average over all 300 simulations is shown in bold in each panel, and the 10th and 90th percentiles are shown as dotted lines. In the top panel, the autocorrelation functions were calculated after demeaning each simulated series by its sample mean. In the second panel, each series is demeaned by the estimated population expectation, which was obtained by averaging the sample means across the 300 independent paths. The difference in the two panels is caused by non-ergodicity of volatility in the multifractal model: The sample mean of squared returns need not converge to its unconditional expectation. As a result, while both panels show hyperbolic decay, the second panel appears to show greater persistence at long lags. This is because low frequency volatility components are partially filtered out when we demean the squared returns series by their sample means rather than the true population mean.

model literally, we might search for equilibrium models that produce scaling behavior in volatility. Alternatively, we can view the scaling restrictions as a pragmatic approach to modelling heterogeneous frequencies in volatility shocks. Depending on one's prior, a realized price path can be consistent with different beliefs about the model's parameters. This property stems from the normalization of the Brownian motion B. Denoting by W a standard Brownian,[6] we can rewrite the model as

$$\ln P(t) - \ln P(0) = \sigma W[\theta(t)].$$

Given the price process over a finite interval $[0,t]$, we cannot distinguish σ from the realization of very low frequency multipliers. For instance, imagine that one knows that the lowest frequency volatility component M_1 is constant over $[0,t]$. An individual whose prior assigns a unit mass on a high value of σ can support her belief by inferring that the realization of M_1 was low. Conversely, an individual assigning a prior unit mass on a low σ might infer that M_1 was high. Although both beliefs explain the data, the first investor forecasts a long run increase in volatility, while the second investor anticipates a decrease.[7] Future extensions may thus seek to incorporate parameter uncertainty. For simplicity, the remainder of this paper assumes that all agents know the true process with certainty. We can then renormalize the data so that B is a standard Brownian motion.

5. Forecasting the distribution of returns

This section examines the forecasting problem for a Poisson multifractal process $X(t) = B[\theta(t)]$ defined on the time interval $[0,T]$. More specifically, given a date t and information set $\mathscr{I}_t$, we want to predict the conditional distribution of X at any future instant. We assume in Section 5.2 that trading time θ is directly observable, and then consider in Section 5.4 the case where only the compound process X is observed.

Given $s > t$, let $f_{X_s}(x|\mathscr{I}_t)$ and $f_{\theta_{t,s}}(\theta|\mathscr{I}_t)$ denote, respectively, the conditional densities of the log-price $X(s)$ and the trading time increment $\theta_{t,s} \equiv \theta(s) - \theta(t)$. By Bayes' rule,

$$f_{X_s}(x|\mathscr{I}_t) = \int_0^{+\infty} n[x; X(t), \theta] f_{\theta_{t,s}}(\theta|\mathscr{I}_t)\, \mathrm{d}\theta, \tag{5.1}$$

where $n(x;\mu,\sigma^2)$ is the density of a Gaussian random variable with mean μ and variance σ^2. The forecasting problem thus reduces to calculating the conditional density $f_{\theta_{t,s}}(\theta|\mathscr{I}_t)$ of trading time.

[6] The increments of a standard Brownian satisfy $W(t) - W(s) \sim \mathcal{N}(0, t-s)$ for all $t > s$.

[7] The PMM might thus be related to recent equilibrium models in which investors rationally maintain divergent opinions (Kurz, 1994).

The estimation of $f_{\theta_{t,s}}(\theta|\mathscr{I}_t)$ follows the same principle whether or not trading time is observable. We introduce a discretized version of the model, in which arrival times occur on a finite grid and the construction stops after $\bar{k}+1$ iterations. We solve the discretized conditioning problem and show that as $\bar{k}\to\infty$, the density forecast of the discretized model converges to the conditional density $f_{\theta_{t,s}}(\theta|\mathscr{I}_t)$.

5.1. A discrete-time Markov model of volatility

This section introduces a discretized version of the Poisson multifractal measure examined in Section 3. We first define a random measure μ^* by a recursive procedure with a *finite* number $\bar{k}+1$ of stages. In each step of the construction, arrival times belong to the regular grid $t=0,1,\ldots,T^*=c^{\bar{k}}$, where $c>1$ is a fixed integer.[8] The measure μ^* is defined on $[0,c^{\bar{k}}]$, and the corresponding trading time θ^* on the time interval $[0,T]$ is defined by

$$\theta^*(s)=\mu^*[0,sT^*/T], \quad 0\leqslant s\leqslant T.$$

The process θ^* is thus the c.d.f. of a homothetic transform of μ^*. It is continuous and piecewise linear. We view the trading time θ of Section 3.1 as the true process, and the discretized version θ^* as a filter that facilitates forecasting.

The construction of μ^* begins at stage zero with a unit mass spread evenly on the interval $[0,T^*]$. In the first stage, let $\{T_i^*\}_{i=1}^{\infty}$ be a sequence of iid random variables each having a *geometric* distribution with parameter $\gamma_1=1-\exp(-\lambda T/T^*)$ and mean $1/\gamma_1$.[9] This sequence helps define the lengths between arrival times. To keep the arrival times bounded below T^*, we denote by P the largest integer n such that $\sum_{i=1}^{n}T_i^*<T^*$. The random integers

$$t_n=\begin{cases}0, & n=0,\\ \displaystyle\sum_{i=1}^{n}T_i^*, & 1\leqslant n\leqslant P,\\ T^*, & n=P+1\end{cases}$$

form a random partition $I_n=[t_n,t_{n+1}]$ of $[0,T^*]$. We define the first stage measure by drawing independent random multipliers M_n and uniformly spreading mass $M_n\ell(I_n)/T^*$ on each interval I_n.

[8] The assumption that T^* be of the form $c^{\bar{k}}$ is used to simplify the presentation of consistency in Section 5.5. An immediate generalization of this construction could consider an arbitrary endpoint T^*. The results of Sections 5.1–5.4 immediately apply to this extension.

[9] The distribution of T_1^* satisfies $\mathbb{P}\{T_1^*=n\}=\gamma_1(1-\gamma_1)^{n-1}$ for all $n\geqslant 1$.

We proceed similarly in stages $k=2,\ldots,\bar{k}$ of the cascade by considering intervals $I_{j_1,\ldots,j_{k-1}}$, and the corresponding geometric random variables $\{T_n^{*j_1,\ldots,j_{k-1}}\}_{n\geqslant 1}$ with parameter $\gamma_k = 1-\exp(-\lambda b^{k-1}T/T^*)$. The random integer $P^{j_1,\ldots,j_{k-1}}$ denotes the number of arrival times falling strictly within the interval $I_{j_1,\ldots,j_{k-1}}$, and helps define the sequence $\{t_{j_1,\ldots,j_{k-1},n}\}_{0\leqslant n\leqslant P^{j_1,\ldots,j_{k-1}}+1}$. Details of stages $k=2,\ldots,\bar{k}$ can be filled in by referring to the continuous time construction.

Finally in stage $\bar{k}+1$, we draw iid random variables Ω_t for each interval $[t-1,t]$. We specify that each Ω_t be distributed like the limit mass $\Omega(\lambda b^{\bar{k}}T/T^*)$ of the Poisson multifractal measure defined in Section 3. This specification ties the high frequency components of the discrete- and continuous-time models. The mass of a cell $[t-1,t]\subseteq I_{j_1,\ldots,j_{\bar{k}}}$ is then defined as

$$\mu^*[t-1,t] = c^{-\bar{k}} M_{j_1}\cdots M_{j_1,\ldots,j_{\bar{k}}}\Omega_t. \tag{5.2}$$

We find it convenient to denote by $M_{1,t},\ldots,M_{\bar{k},t}$ the multipliers $M_{j_1},\ldots,$ $M_{j_1,\ldots,j_{\bar{k}}}$, and view $M_{k,t}$ as the value of the kth multiplier at date t. We can then stack all these multipliers into

$$Z_t = (M_{1,t},\ldots,M_{\bar{k},t}),$$

which we call the *volatility state vector* at date t.

The construction can be reinterpreted as a stochastic volatility model in which the latent state vector Z_t follows a Markov process, as is now shown. For a given t, consider the information set $\mathscr{I}_{t-1} = \{(t_{j_1,\ldots,j_k}, M_{j_1,\ldots,j_k})\}_{t_{j_1,\ldots,j_k}\leqslant t-1}$ of past stopping times and multipliers. We denote by κ_t the lowest frequency change between $t-1$ and t. More formally, let $\kappa_t = \bar{k}+1$ if t does not belong to the set of arrivals $\{t_{j_1,\ldots,j_k}\}$, and $\kappa_t = \inf\{k\colon t = t_{j_1,\ldots,j_k}\}$ otherwise. Since time increments are geometrically distributed, κ_t is independent of past stopping times and satisfies

$$\mathbb{P}\{\kappa_t = k\} = \begin{cases} \gamma_1, & k=1,\\ \gamma_k(1-\gamma_1)\cdots(1-\gamma_{k-1}), & k=2,\ldots,\bar{k},\\ (1-\gamma_1)\cdots(1-\gamma_{\bar{k}}), & k=\bar{k}+1. \end{cases}$$

The distribution of the future multipliers $Z_{t+1} = (M_{1,t+1},\ldots,M_{\bar{k},t+1})$ thus only depends on the current $Z_t = (M_{1,t},\ldots,M_{\bar{k},t})$. The volatility state vector Z_t follows a Markov process, a result that greatly facilitates the conditioning problem.

The Markov property implies an alternative method to construct and simulate the measure μ^*. Instead of recursively splitting $[0,c^{\bar{k}}]$ into finer intervals, we can build the measure as time progresses. At date $t=0$, we draw independent random variables $(M_{1,1},\ldots,M_{\bar{k},1},\Omega_1)$, and let $\mu^*[0,1] = c^{-\bar{k}}M_{1,1}\cdots M_{\bar{k},1}\Omega_1$. Given Z_t, we generate the mass $\mu^*[t,t+1]$ by drawing the index κ_{t+1}, new multipliers $M_{\kappa_{t+1},t+1},\ldots,M_{\bar{k},t+1}$, and a high frequency component Ω_{t+1}.

Multipliers corresponding to frequencies lower than κ_{t+1} remain unchanged: $M_{i,t+1}=M_{i,t}$ for all $i<\kappa_{t+1}$, and the mass $\mu^*[t,t+1]$ is set equal to $c^{-\bar{k}}M_{1,t+1}\cdots M_{\bar{k},t+1}\Omega_{t+1}$.

We assume in the rest of this section that the multiplier M takes a finite number of values b_m. The vector Z_t can therefore take $d=b_m^{\bar{k}}$ values $z^1,\ldots,z^d\in\mathbb{R}^{\bar{k}}$. The dynamics of the Markov chain Z_t are characterized by the transition matrix $A=(a_{i,j})_{1\leqslant i,j\leqslant d}$ with components $a_{ij}=\mathbb{P}(Z_{t+1}=z^j\mid Z_t=z^i)$. We note that

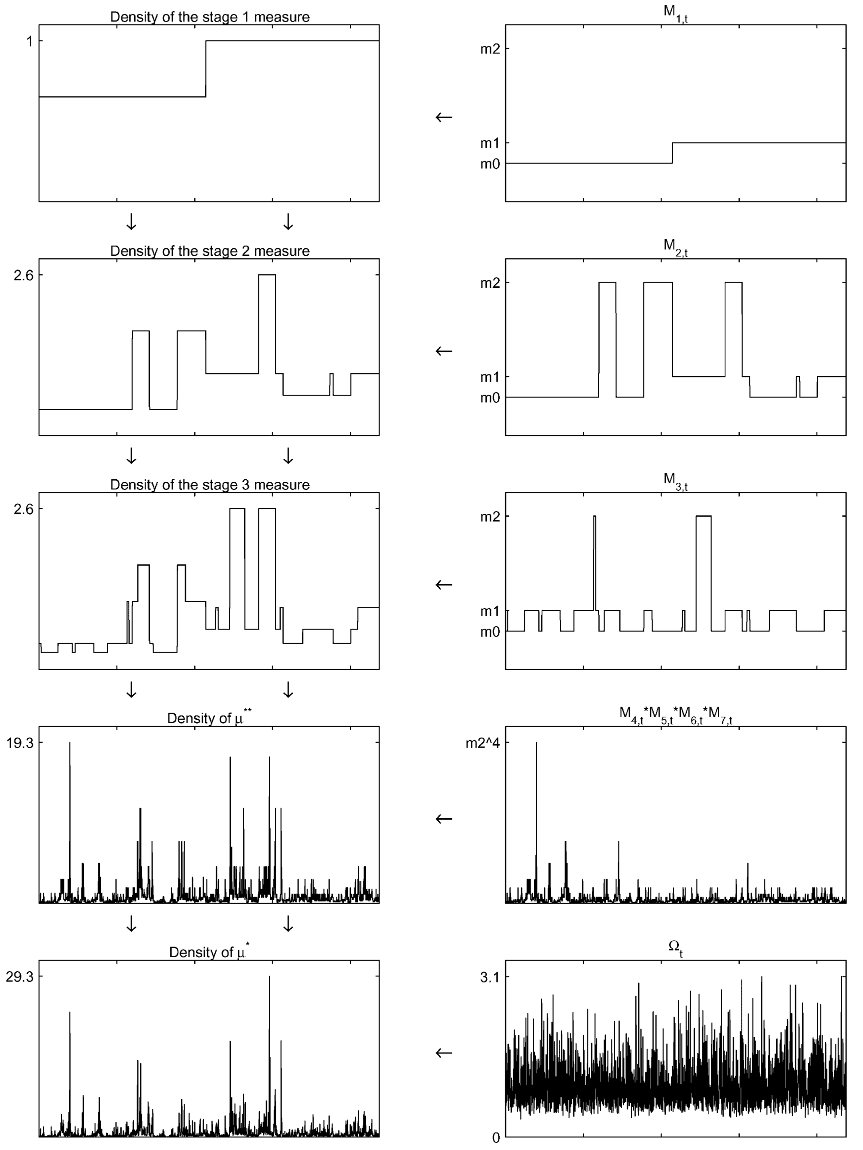
$$\begin{aligned} a_{ij} &= \sum_{k=1}^{\bar{k}+1}\mathbb{P}(\kappa_t=k)\,\mathbb{P}(Z_{t+1}=z^j\mid Z_t=z^i,\kappa_t=k) \\ &= \sum_{k=1}^{\bar{k}+1}\mathbb{P}(\kappa_t=k)1_{\{z_1^i=z_1^j,\ldots,z_{k-1}^i=z_{k-1}^j\}}\left[\prod_{l=k}^{\bar{k}}\mathbb{P}(M=z_l^j)\right], \end{aligned}$$

where z_k^i denotes the kth component of vector z^i, and $1_{\{z_1^i=z_1^j,\ldots,z_{k-1}^i=z_{k-1}^j\}}$ is the dummy variable equal to 1 if $(z_1^i,\ldots,z_{k-1}^i)=(z_1^j,\ldots,z_{k-1}^j)$, and 0 otherwise.

This alternative construction of the measure closely parallels our economic intuition of volatility. Each period, the economy can receive volatility shocks of heterogeneous frequencies. Tractability and parsimony stem from the hypothesis that all multipliers have the same marginal distribution. We also assume that when a volatility innovation arrives at a given frequency, all higher frequency components also change. Economically, this corresponds to the idea that relatively low frequency shocks to technology or demographics cause revisions in higher frequency economic variates such as the earnings cycle or aggregate liquidity. It also implies that lower frequency volatility shocks result in correspondingly larger average changes in volatility.

Despite its simplicity, the model can generate complex forecast dynamics. Thus although for each k the multiplier $\{\mathbb{E}(M_{k,t+n}|Z_t)\}_{n\geqslant 1}$ monotonically reverts to $\mathbb{E}M=1$ as the time horizon n increases, forecasts of future volatility $\mu_{t+n}=\mu^*[t+n-1,t+n]$ need not be monotonic in n. Consider, for instance, a state Z_t with a high volatility $\mu_t=M_{1,t}\cdots M_{\bar{k},t}>1$ but a low stage $\bar{k}$ →

Fig. 4 Construction of the discretized volatility measures μ^{**} and μ^* when $\bar{k}=7$ and M has a trinomial distribution. This figure is based on the trinomial multifractal construction with $[m_0,m_1,m_2]=[0.65,1.0,2.6]$ and $[p_0,p_1,p_2]=[0.5,0.39,0.11]$. The left-hand side shows increasingly refined stages in the construction of the discretized measures. The right-hand side shows the sequence of multiplicative modifications that produce the sequence of measures. The first three panels show the incorporation of the low frequency volatility components M_1, M_2, and M_3. The fourth panel shows the effect of stages four through seven. The resulting measure is the seventh stage in the construction of μ^*. It is also called μ^{**} because it is the final stage in the construction with the exception of the high frequency component Ω. The final panel shows the incorporation of the high frequency component Ω, which completes construction of the discretized measure μ^*.


Density of the stage 1 measure
1
M1,t
m2
m1
m0
Density of the stage 2 measure
2.6
M2,t
m2
m1
m0
Density of the stage 3 measure
2.6
M3,t
m2
m1
m0
Density of μ**
19.3
M4,t*M5,t*M6,t*M7,t
m2^4
Density of μ*
29.3
Ωt
3.1
0

multiplier $M_{\bar{k},t} < 1$. Because the high frequency multiplier $M_{\bar{k},t+n}$ is expected to increase, the volatility $\mathbb{E}(\mu_{t+n}|Z_t)$ may *increase* in the short run before decreasing towards 1 at longer horizons. This illustrates the kind of rich forecasting dynamics that our multi-frequency model can generate.

5.2. Inferring current multipliers

We now turn to the inference problem of a market participant who observes the trading time increments $\mu_t = \mu^*[t-1,t]$, but not the exact value of the multipliers $Z_t = (M_{1,t}, \ldots, M_{\bar{k},t})$. At a given instant t, the investor has thus information set $\mathscr{I}_t = \{\mu_s\}_{s=1}^t$. Her conditional distribution over the multipliers can be calculated recursively since Z_t is a Markov chain.

We previously imposed that Ω_t be distributed like the limit mass $\Omega(\lambda b^{\bar{k}} T/T^*)$ of the Poisson multifractal. Additionally, we choose a specification of the multiplier M such that the random mass Ω_t has a density $f_\Omega(\omega)$ with respect to Lebesgue measure. Given $a = (a_1, \ldots, a_{\bar{k}}) \in \mathbb{R}^{\bar{k}}$, we find it convenient to denote by $g(a)$ the product $\prod_{i=1}^{\bar{k}} a_i$. The relation $\mu_t = g(Z_t)\Omega_t$ implies that the density of μ_t conditional on Z_t satisfies

$$f_{\mu_t}(\mu|Z_t = z^i) = \frac{1}{g(z^i)} f_\Omega\left[\frac{\mu}{g(z^i)}\right]. \tag{5.3}$$

The investor who observes past trading time increments has conditional probabilities

$$\Pi_t^j = \mathbb{P}(Z_t = z^j|\mathscr{I}_t)$$

over the latent states $z^1, \ldots, z^d$. We note that $\Pi_{t+1}^j = \mathbb{P}(Z_{t+1} = z^j|\mathscr{I}_t, \mu_{t+1})$ satisfies

$$\Pi_{t+1}^j = f_{\mu_{t+1}}(\mu_{t+1}|Z_{t+1} = z^j)\mathbb{P}(Z_{t+1} = z^j|\mathscr{I}_t)/f_{\mu_{t+1}}(\mu_{t+1}|\mathscr{I}_t),$$

which, by (5.3), can be rewritten as

$$\Pi_{t+1}^j = \frac{f_\Omega[\mu_{t+1}/g(z^j)](\sum_{i=1}^d a_{ij}\Pi_t^i)}{g(z^j) f_\mu(\mu_{t+1}|\mathscr{I}_t)}. \tag{5.4}$$

This formula expresses the conditional probability Π_{t+1}^j as a function of the observation μ_{t+1} and the probabilities Π_t^i calculated in period t.

These results can be conveniently rewritten in vector notation. Consider the row vectors $\Pi_t = (\Pi_t^1, \ldots, \Pi_t^d) \in \mathbb{R}_+^d$, $\imath = (1, \ldots, 1) \in \mathbb{R}^d$, and

$$\omega(\mu_{t+1}) = \left[\frac{f_\Omega(\mu_{t+1}/g(z^1))}{g(z^1)}, \ldots, \frac{f_\Omega(\mu_{t+1}/g(z^d))}{g(z^d)}\right].$$

Let $x * y$ denote the Hadamard product $(x_1 y_1, \ldots, x_d y_d)$ of any $x, y \in \mathbb{R}^d$. We can rewrite (5.4) as

$$\Pi_{t+1} = \frac{\omega(\mu_{t+1}) * (\Pi_t A)}{[\omega(\mu_{t+1}) * (\Pi_t A)]\iota'}, \tag{5.5}$$

using the fact that the conditional probabilities Π^j_{t+1} add up to one.

These results imply that we can recursively calculate Π_t given an initial vector Π_0. By construction, the multipliers $(M_{1,1}, \ldots, M_{\bar{k},1})$ are independent and the initial vector Π_0 is uniquely determined in our model.[10] In empirical work however, the econometrician may want to use outside information on the starting values of the multipliers to select a different value for Π_0.

5.3. Forecasting future volatility

We now calculate the conditional density $f_{\sigma^2_{t,n}}(\sigma^2|\mathscr{I}_t)$ of future mass $\sigma^2_{t,n} \equiv \mu^*[t, t+n]$. Given the information set $\mathscr{I}_t$, the agent assigns conditional probabilities Π_t to the *current* state vector Z_t. Since Z_t is a Markov chain, Π_t contains all the information necessary to infer $f_{\sigma^2_{t,n}}(\sigma^2|\mathscr{I}_t)$. For this reason, the formal calculations developed in this section are the same whether or not trading time is directly observed.

The conditional probabilities of *future* multipliers $\hat{\Pi}_{t,n} = \mathbb{P}(Z_{t+n}|\mathscr{I}_t)$ are given by multiplication of the transition matrix:

$$\hat{\Pi}_{t,n} = \Pi_t A^n.$$

Letting $\varphi_n(\sigma^2; z^j, \Pi_t) \equiv f_{\sigma^2_{t,n}}(\sigma^2|Z_{t+n} = z^j, \mathscr{I}_t)$, we infer from Bayes' rule:

$$f_{\sigma^2_{t,n}}(\sigma^2|\mathscr{I}_t) = \sum_{j=1}^{d} \varphi_n(\sigma^2; z^j, \Pi_t)\hat{\Pi}^j_{t,n}. \tag{5.6}$$

The conditional densities $\varphi_n(\sigma^2; z^j, \Pi_t)$ in the sum can be computed by a recursive algorithm. When $n = 1$, φ_1 is independent of Π_t and satisfies

$$\varphi_1(\sigma^2; z^j) \equiv \frac{1}{g(z^j)} f_\Omega\left[\frac{\sigma^2}{g(z^j)}\right].$$

For all $n > 1$, the conditional density $\varphi_{n+1}(\sigma^2; z^j, \Pi_t)$ can be rewritten as the sum $\sum_{i=1}^d f_{\sigma^2_{t,n+1}}(\sigma^2|Z_{t+n} = z^i, Z_{t+n+1} = z^j, \mathscr{I}_t)\mathbb{P}(Z_{t+n} = z^i|Z_{t+n+1} = z^j, \mathscr{I}_t)$, or

$$\varphi_{n+1}(\sigma^2; z^j, \Pi_t) = \sum_{i=1}^{d} f_{\sigma^2_{t,n+1}}(\sigma^2|Z_{t+n} = z^i, Z_{t+n+1} = z^j, \mathscr{I}_t)a_{ij}\hat{\Pi}^i_{t,n}/\hat{\Pi}^j_{t,n+1}.$$

We now infer from Bayes' rule that $f_{\sigma^2_{t,n+1}}(\sigma^2|Z_{t+n} = z^i, Z_{t+n+1} = z^j, \mathscr{I}_t)$ equals $\int_0^{\sigma^2} f_{\sigma^2_{t,n+1}}(\sigma^2|\sigma^2_{t,n} = u, Z_{t+n+1} = z^j, \mathscr{I}_t) f_{\sigma^2_{t,n}}(u|Z_{t+n} = z^i, \mathscr{I}_t)\,du$, which simplifies

[10] The components of Π_0 satisfy $\Pi^j_0 = \Pi^{\bar{k}}_{l=1}\mathbb{P}(M = z^j_l)$ for all j.

to $\int_0^{\sigma^2} \varphi_1(\sigma^2 - u; z^j)\varphi_n(u; z^i, \Pi_t)\,\mathrm{d}u$. The sequence (φ_n) thus satisfies the recursion:

$$\varphi_{n+1}(\sigma^2; z^j, \Pi_t) = (\hat{\Pi}^j_{t,n+1})^{-1} \sum_{i=1}^{d} a_{ij} \hat{\Pi}^i_{t,n} \int_0^{\sigma^2} \varphi_1(\sigma^2 - u; z^j)\varphi_n(u; z^i, \Pi_t)\,\mathrm{d}u.$$

Eq. (5.6) then provides the forecast density $f_{\sigma^2_{t,n}}(\sigma^2|\mathscr{I}_t)$ of trading time increments.

5.4. Forecasting when trading time is unobservable

We now show how to forecast future volatility when the information set contains past returns but not trading time. Each period, the investor observes the discretized return $\rho_t = (c^{-\bar{k}} M_{1,t} \cdots M_{\bar{k},t} \Omega_t)^{1/2} \xi_t$, where $\{\xi_t\}_{t=1}^{T^*}$ is an i.i.d. sequence of centered normals independent from arrival times and multipliers. Consider the information set $\mathscr{I}^\rho_t = \{\rho_t\}_{s=1}^t$, the conditional probabilities $\Pi^j_t = \mathbb{P}(Z_t = z^j | \mathscr{I}^\rho_t)$, and the row vector

$$\omega^*(\rho_{t+1}) = \left(\frac{f_{\Omega^{1/2}\xi}[\rho_{t+1}/g(z^1)]}{g(z^1)}, \ldots, \frac{f_{\Omega^{1/2}\xi}[\rho_{t+1}/g(z^d)]}{g(z^d)} \right).$$

The inference problem has the same structure as in Section 5.2. We can thus recursively calculate Π_t using the relation

$$\Pi_{t+1} = \frac{\omega^*(\rho_{t+1}) * (\Pi_t A)}{[\omega^*(\rho_{t+1}) * (\Pi_t A)]\iota'}. \tag{5.7}$$

When only prices are observable, it may be appealing in some situations to slightly modify the discrete-time model. Using the same arrivals and multipliers as in Section 5.1, we consider the measure on $[0, T^*]$ that uniformly spreads

$$\mu^{**}[t-1, t] = c^{-\bar{k}} M_{1,t} \cdots M_{\bar{k},t}$$

on each cell. The new measure μ^{**} thus omits the high frequency component Ω_t used in (5.2). As in Section 5.1, we then define the trading time $\theta^{**}(s) = \mu^{**}[0, sT^*/T]$ by a homothetic transform of μ^{**}. When the investor observes the increments $\rho^{**}_t = (c^{-\bar{k}} M_{1,t} \cdots M_{\bar{k},t})^{1/2} \xi_t$, the updating problem has the same structure as before, and the conditional probability Π_t can again be iteratively calculated.

5.5. Consistency

In Section 3, we defined the trading time θ as the c.d.f. of a Poisson multifractal. In order to facilitate forecasting, the piecewise linear trading times θ^* and θ^{**} were then constructed in Sections 5.1 and 5.4 by a cascade of $\bar{k}+1$ stages. It is now convenient to index these discretized trading times by the number $\bar{k}$ of their multipliers. In this section, we provide conditions under which the sequences $\theta^{*\bar{k}}$ and $\theta^{**\bar{k}}$ converge to θ as $\bar{k}\to\infty$. The weak convergence of the discretized trading times to θ implies the consistency of forecasts obtained from the discretized models.

Let $\mathscr{C}$ denote the space of real-valued continuous functions defined on the interval $[0,T]$. Random functions θ, $\theta^{*\bar{k}}$ and $\theta^{**\bar{k}}$ can be viewed as probability measures on $\mathscr{C}$. We provide two alternative conditions under which the sequence of trading times $\{\theta^{*\bar{k}}\}_{\bar{k}=1}^{\infty}$ and $\{\theta^{**\bar{k}}\}_{\bar{k}=1}^{\infty}$ weakly converge to the process θ.[11]

Condition 1′. $\mathbb{E}(M^2)b<c^2$.

Condition 2. $b<c$

The second condition requires that the number of grid points grows faster than the volatility frequencies. On the other hand, Condition 1′ allows the grid size to grow at an identical or slower than the volatility frequencies. When $b=c$, this assumption reduces to $\mathbb{E}(M^2)=1+Var(M)<b$ and thus coincides with Condition 1, which was used in the existence proof of Section 3.1. We show in Appendix B:

Theorem 2. Under Conditions 1′ *or* 2, *the sequences* $\{\theta^{*\bar{k}}\}_{\bar{k}=1}^{\infty}$ *and* $\{\theta^{**\bar{k}}\}_{\bar{k}=1}^{\infty}$ *of discretized trading times weakly converge to the continuous-time process* θ.

This implies the consistency of the density forecasts presented in Sections 5.3 and 5.4.

6. Conclusion

This paper has developed analytical forecasting methods for a new stochastic process, the Poisson multifractal model (PMM). The PMM provides a fully stationary version of the compound process developed in Mandelbrot et al. (1997). It parsimoniously captures the volatility persistence, moment-scaling and thick tails that characterize many financial time series. We specify volatility as the multiplicative product of an infinite sequence of random functions, which are generated by Poisson arrivals of increasing frequencies. The model is thus consistent with the different time scales of economic variables such as

[11] Billingsley (1999), Pollard (1984), and Davidson (1994) provide excellent presentations of the weak convergence of stochatic measures and processes.

technology shocks, business and earnings cycles, and liquidity shocks. The PMM implies semi-martingale prices, and thus precludes arbitrage in a standard two-asset setting. Squared returns have long memory, and the highest finite moment of returns may have any value greater than two. This wide range of tail behaviors is fully provided by intermittent bursts of volatility, and does not require separate modelling of the tails.

Forecasting is facilitated by a discretized version of the process with a finite state space and a simple Markov structure. We show that the discretized version converges to the continuous process as the grid step size goes to zero, which implies the consistency of the forecasting methodology. Past data permits conditioning over the full set of volatility state variables, each of which has different persistence. While volatility is always reverting to the mean in the long run, forecasts can vary considerably in the short run depending on the current state. In addition, the PMM also supports differing beliefs regarding the long-run behavior of volatility.

This analysis can be extended in several directions. First, the discrete-time model allows us to use maximum likelihood techniques to estimate the parameters of the model. Second, financial analysts may consider more data than past prices to infer the current volatility state. Future research may thus construct updating rules that incorporate macroeconomic aggregates, market indices or derivative prices. Finally to model equity data, we may introduce skewness into the model by allowing a negative autocorrelation between the Brownian motion and the trading time. Overall, the multifractal approach can generate tractable models that match important features of the data, and offers promising new directions for future research in econometrics and financial theory.

Acknowledgements

We received helpful comments from D. Backus, J. Campbell, G. Chamberlain, R. Deo, F. Diebold, R. Engle, C. Hurvich, G. Kuersteiner, K. Li, B. Mandelbrot, J. Porter, J. Rosenberg, J. Stein, J. Stock, K. West, S. Zin, anonymous referees, and seminar participants at Carnegie-Mellon, the Federal Reserve Board, Harvard, MIT, the NBER, NYU, the University of British Columbia, and the University of Waterloo.

Appendix A.

A.1. Proof of Theorem 1

For every $k \geqslant 0$, let θ^k denote the c.d.f. of μ_k. The random function θ^k is continuous and piecewise linear. Let $\mathscr{C}$ denote the space of continuous functions defined on $[0,1]$. For any collection of intervals $I_1,\ldots,I_p$, the vector sequence $\{\mu_k(I_1),\mu_k(I_2),\ldots,\mu_k(I_p)\}_k$ is a positive martingale and therefore has a limit distribution. The sequence $\{\mu_k\}$ thus has at most one cluster point.

As discussed in Billingsley (1999) and Davidson (1994), a cluster point exists if the sequence $\{\theta^k\}$ is tight. For any continuous function $x\in\mathscr{C}$, it is convenient to consider the *modulus of continuity* $w(x,\delta)=\sup_{|t-s|<\delta}|x(t)-x(s)|$. Our discussion is based on the following result, proved in Billingsley (1999).

Theorem A.1. The sequence $\{\mu_k\}$ is tight if and only if these two conditions hold:

(i) *For every $\eta>0$, there exist a and K such that $\mathbb{P}\{|\theta^k(0)|\geqslant a\}\leqslant\eta\ \forall k\geqslant K$.*

(ii) *For every $\varepsilon>0$, $\lim_{\delta\to 0}\limsup_{k\to\infty}\mathbb{P}\{w(\theta^k,\delta)\geqslant\varepsilon\}=0$.*

The first condition is trivially satisfied in our setup since $\theta_k(0)\equiv 0$. We now turn to the second condition. Since the function $\limsup_k \mathbb{P}\{w(\theta^k,\delta)\geqslant\varepsilon\}$ is increasing in δ, we can restrict our attention to step sizes of the form $\delta_n=1/n$, $n=1,2,\ldots,\infty$. For a given n, consider the regularly spaced grid $t_0=0<t_1=\delta_n<\cdots<t_n=1$. Since the function θ_k is increasing, Theorem 7.4 in Billingsley (1999) implies

$$\mathbb{P}\{w(\theta^k,\delta_n)\geqslant\varepsilon\}\leqslant\sum_{i=0}^{n-1}\mathbb{P}\left\{\theta^k(t_{i+1})-\theta^k(t_i)\geqslant\frac{\varepsilon}{3}\right\}.$$

Each increment $\theta^k(t_{i+1})-\theta^k(t_i)$ is distributed like $\theta^k(\delta_n)$, implying $\mathbb{P}\{w(\theta^k,\delta_n)\geqslant\varepsilon\}\leqslant(3/\varepsilon)^q\, n\mathbb{E}[\theta_k(\delta_n)^q]$ for any $q>0$. Observing that $\theta^k(\delta_n)\stackrel{d}{=}\delta_n\mu_k[0,1]$, we infer

$$\mathbb{P}\{w(\theta^k,\delta_n)\geqslant\varepsilon\}\leqslant(3/\varepsilon)^q n^{1-q}\,\mathbb{E}\{\mu_k[0,1]^q\}.$$

The right-hand side does not converge to 0 when $q=1$, and the analysis will focus instead on the second moment ($q=2$). It is useful to consider the equality

$$\mu_k[0,1]=\sum_{j_1=0}^{N}M_{j_1}\sum_{j_2=0}^{N^{j_1}}M_{j_1,j_2}\cdots\sum_{j_k=0}^{N^{j_1,\ldots,j_{k-1}}}M_{j_1\ldots j_k}\Delta t_{j_1\ldots j_k}, \tag{A.1}$$

where $\Delta t_{j_1\ldots j_k}=(t_{j_1\ldots j_k+1}-t_{j_1\ldots j_k})$. Simple algebra implies that $m_k=\mathbb{E}\{\mu_k[0,1]^2\}$ satisfies $m_k\leqslant m_{k-1}+[\mathbb{E}(M^2)]^k\,\mathbb{E}[\sum_{j_1=0}^{N}\sum_{j_2=0}^{N^{j_1}}\cdots\sum_{j_k=0}^{N^{j_1,\ldots,j_{k-1}}}(\Delta t_{j_1\ldots j_k})^2]$. We can show:

Lemma A.1. Consider a Poisson process with frequency λ, which partitions the interval $[0,t]$ into $N+1$ intervals: $I_0=[0,t_1]$, $I_1=[t_1,t_2],\ldots,I_N=[t_N,t]$. Then $\mathbb{E}[\sum_{n=0}^{N}l(I_n)^2]\leqslant 2t/\lambda$.

This result implies $m_k \leqslant (2/\lambda)[\mathbb{E}(M^2)/b]^k + m_{k-1}$. Under Condition 1, the sequence $\{m_k\}$ is bounded and $\limsup_k \mathbb{P}\{w(\theta^k, \delta_n) \geqslant \varepsilon\} \leqslant n^{-1}(3/\varepsilon)^q \sup_k m_k \to 0$ as $n \to \infty$.

A.2. Proof of Proposition 2

We prove (3.2) by jointly constructing two measures μ and μ' defined on intervals $[0,T]$ and $[0,1]$. In the first stage, we define μ_1 on $[0,T]$ by drawing exponential variables T_i with mean $1/\lambda$, instants $t_n = \sum_{i=1}^n T_i$ and random multipliers M_n. We then consider instants $t'_n = t_n/T$ and the measure μ'_1 that uniformly spreads mass $M_n(t'_{n+1} - t'_n)$ on each interval $[t'_n, t'_{n+1}]$. We note that $\mu'_1 = \mu_1 \circ \varphi$, where $\varphi(x) \equiv x/T$. Since each $t'_n = \sum_{i=1}^n (T_i/T)$ is the sum of random exponentials with mean $1/(\lambda T)$, measure μ'_1 is the first stage in the construction of a Poisson multifractal on $[0,1]$ with parameters λT and M. This argument generalizes to every stage of the cascade, implying $\mu'[0,1] = \mu[0,T]$ or $\Omega(1, \lambda T) \stackrel{\mathrm{d}}{=} \Omega(T, \lambda)$.

A.3. Proof of Proposition 3

We first show that

$$c^q \mathbb{E}[\Omega(c\lambda)^q] \leqslant \mathbb{E}[\Omega(\lambda)^q] \leqslant n^{\max(0,1-q)} \mathbb{E}[\Omega(\lambda/n)^q] \tag{A.2}$$

for every integer $n \geqslant 1$ and real numbers $\lambda > 0$, $0 < c < 1$, $q \geqslant 0$. Let μ denote a Poisson multifractal on $[0,1]$ such that $\mu[0,1] \stackrel{\mathrm{d}}{=} \Omega(\lambda)$. By Proposition 1, the random variable $\mu[0,c] \leqslant \mu[0,1]$ is distributed like $c\Omega(c\lambda)$, implying $c^q \mathbb{E}[\Omega(c\lambda)^q] \leqslant \mathbb{E}[\Omega(\lambda)^q]$. Moreover since $\mu[0,1] = \sum_{i=0}^{n-1} \mu[i/n, (i+1)/n]$, we infer that [12] $(\mu[0,1])^q \leqslant \max(n^{q-1}, 1) \sum_{i=0}^{n-1} \{\mu[i/n, (i+1)/n]\}^q$ and thus $\mathbb{E}[\Omega(\lambda)^q] \leqslant n^{\max(0,1-q)} \mathbb{E}[\Omega(\lambda/n)^q]$. The end of the proof is then straightforward.

A.4. Proof of Proposition 4

We assume that $T = 1$, and let $g(\lambda) = \mathbb{E}[\Omega(\lambda)^q]$. Since $\mathbb{P}(N = 0) = \mathrm{e}^{-\lambda}$, simple conditioning implies

$$g(\lambda) = \mathbb{E}(M^q)\, g(b\lambda) \mathrm{e}^{-\lambda} + X\mathbb{P}(N > 0),$$

where $X = \mathbb{E}[\{\sum_{j=0}^N M_j\,(t_{j+1} - t_j)\, \Omega_j[b\lambda(t_{j+1} - t_j)]\}^q | N > 0]$. We infer from footnote 12 and relation (A.2) that $X \leqslant \mathbb{E}(M^q)\, g(b\lambda)\, \mathbb{E}[(N+1)^{\max(q,1)} | N > 0]$,

[12] Recall that $(\sum_{i=1}^n x_i)^q \leqslant \max(n^{q-1}, \sum_{i=1}^n x_i^q)$ for any $n \geqslant 1$, $(x_1, \ldots, x_n) \in \mathbb{R}_+^n$, $q \geqslant 0$.

and thus

$$1 \leqslant \frac{e^{\lambda} g(\lambda)}{\mathbb{E}(M^q)\, g(b\lambda)} \leqslant 1 + \sum_{n=1}^{+\infty} \lambda^n (n+1)^{\max(q,1)}/n!$$

This implies $g(\lambda) \sim \mathbb{E}(M^q)\, g(b\lambda)$ when $\lambda \to 0$. It can then be shown that $g(\lambda) \sim c_q \lambda^{-\log_b \mathbb{E}(M^q)}$.

Appendix B.

Our proof is based on the following result, proved in Billingsley (1999).

Theorem B.1. If

$$[\theta^{**k}(t_1), \ldots, \theta^{**k}(t_p)] \xrightarrow{\mathrm{d}} [\theta(t_1), \ldots, \theta(t_p)] \tag{B.1}$$

holds for all $t_1, \ldots, t_p$ *and if*

$$\lim_{\delta \to 0} \limsup_{k \to \infty} \mathbb{P}\{w(\theta^{**k}, \delta) \geqslant \varepsilon\} = 0 \tag{B.2}$$

for each positive ε, *then* θ^{**k} *weakly converges to* θ.

We successively establish the convergence of marginals (B.1) and the tightness condition (B.2).

B.1. Convergence of the marginals

B.1.1. Construction of coupled trading times

We assume without loss of generality that $T = 1$. Consider the sequence $\{\mu_k\}_{k=1}^{\infty}$ used in the construction of the continuous-time measure defined on $[0, 1]$. This construction relies on the exponential variables $\{T^{j_1, \ldots, j_k}\}$, the stochastic arrival times $t_{j_1, \ldots, j_k}$ and the multipliers $M_{j_1, \ldots, j_k}$. Consistent with the proof of Theorem 1, we consider the increments $\Delta t_{j_1 \ldots j_k} = (t_{j_1 \ldots j_k + 1} - t_{j_1 \ldots j_k})$, and denote by θ_k the c.d.f. of μ_k. The mass $\theta_k(1) = \mu_k[0, 1]$ is then given by (A.1).

We can also construct a coupled trading time θ^{**k} with discretized arrival times. Let $[x]$ denote for all $x \in \mathbb{R}$ the unique integer such that $[x] \leqslant x < [x] + 1$. For a given integer $c > 1$, consider the uniform grid 0, $1/c^k, \ldots, 1$. We discretize the sequence $\{t_n\}$ on the grid by letting

$$s_n = \begin{cases} 0, & n = 0, \\ \displaystyle\sum_{j=1}^{n} ([c^k T^j] + 1)/c^k, & 1 \leqslant n \leqslant P, \\ 1, & n = P + 1 \end{cases}$$

for all $n \geqslant 1$. We observe that for all $i \geqslant 1$,

$$\mathbb{P}\{s_{n+1} - s_n = i/c^k\} = \mathbb{P}\left\{\frac{i-1}{c^k} \leqslant T^j < \frac{i}{c^k}\right\} = (1-\gamma_{1,k})^{i-1}\gamma_{1,k},$$

where $\gamma_{1,k} = 1 - \exp(-\lambda/c^k)$. The random variable $c^k(s_{n+1} - s_n)$ has thus a geometric distribution with parameter $\gamma_{1,k}$. Proceeding similarly at each stage of the cascade, we construct a discretized trading time θ^{**k} coupled to θ_k with random parameters $P^{j_1,\dots,j_k}$ and $s_{j_1\dots j_k}$. Consistent with previous notation, $\theta^{**k}[(i+1)c^{-k}] - \theta^{**k}(ic^{-k}) = c^{-k}M_{j_1}\cdots M_{j_1\dots j_k}$ when $[ic^{-k},(i+1)c^{-k}] \subseteq I_{j_1\dots j_k}$. We infer that

$$\theta^{**k}(1) = \sum_{j_1=0}^{P} M_{j_1} \sum_{j_2=0}^{P^{j_1}} M_{j_1,j_2} \cdots \sum_{j_k=0}^{P^{j_1,\dots,j_{k-1}}} M_{j_1\dots j_k}\Delta s_{j_1\dots j_k},$$

where $\Delta s_{j_1\dots j_k} = s_{j_1\dots j_k+1} - s_{j_1\dots j_k}$.

The number of arrivals at each stage may differ in the continuous and discretized measures. For this reason, we extend the definition of $\Delta t_{j_1\dots j_p}$ and $N^{j_1\dots j_p}$ by letting $\Delta t_{j_1\dots j_p} = 0$, $N^{j_1\dots j_p} = 0$ when these numbers are not already defined. The definitions of $\Delta s_{j_1\dots j_p}$ and $P^{j_1\dots j_p}$ are similarly extended. It is then convenient to introduce $L^{j_1,\dots,j_p} = \min(P^{j_1,\dots,j_p}, N^{j_1,\dots,j_p})$, $H^{j_1,\dots,j_p} = \max(P^{j_1,\dots,j_p}, N^{j_1,\dots,j_p})$, and the quantity

$$\Delta_{j_1\dots j_k} = \Delta s_{j_1\dots j_k} - \Delta t_{j_1\dots j_k},$$

which quantifies the difference in length of corresponding intervals. With this notation, the trading times differ at $t=1$ by

$$\theta^{**k}(1) - \theta_k(1) = \sum_{j_1=0}^{H} M_{j_1} \sum_{j_2=0}^{H^{j_1}} M_{j_1,j_2} \cdots \sum_{j_k=0}^{H^{j_1,\dots,j_{k-1}}} M_{j_1\dots j_k}\Delta_{j_1\dots j_k}. \tag{B.3}$$

In Sections 8.1.2 and 8.1.3, we will show that the first or second moment of $|\theta^{**k}(1) - \theta_k(1)|$ converges to zero as $k \to \infty$. This will imply that $\theta^{**k}(1) = \theta_k(1) + o_{\mathrm{p}}(1) \xrightarrow{\mathrm{d}} \theta(1)$.

We first establish some useful results. For every $p \in \{1,\dots,k\}$, the average number of subintervals in the pth stage of the construction is $\mathbb{E}(\sum_{j_1=0}^{N} \cdots \sum_{j_p=0}^{N^{j_1,\dots,j_{p-1}}} 1) = \lambda b^p$, or equivalently $\mathbb{E}[\sum_{j_1=0}^{N} \cdots \sum_{j_{p-1}=0}^{N^{j_1,\dots,j_{p-2}}} N^{j_1,\dots,j_{p-1}}] = \lambda b^p$. It is also useful to bound the size of the quantities $\Delta_{j_1\dots j_k}$. Given $j = (j_1,\dots,j_p)$, consider the intervals (j,n) in the continuous and discretized constructions.

Lemma B.1. The inequalities

$$0 \leqslant \Delta_{j,n} \leqslant 1/c^k \qquad (0 \leqslant n < L^j), \tag{B.4}$$

$$|\Delta_{j,L^j}| \leqslant |\Delta_j| + c^{-k}(1 + L^j) \tag{B.5}$$

and

$$|\Delta_{j,L^j+1}| + \cdots + |\Delta_{j,H^j}| \leqslant |\Delta_j| + c^{-k}(1+L^j) \tag{B.6}$$

hold for all $p \geqslant 0$, $j = (j_1, \ldots, j_p)$.

Proof. When $n < L^j$, the relation $s_{j,n+1} - s_{j,n} = (1 + [c^k(t_{j,n+1} - t_{j,n})])/c^k$ implies (B.4). Let $j+1 = (j_1, \ldots, j_{p-1}, j_p + 1)$. We note that

$$\Delta_{j,L^j} = \Delta'_{j,L^j} - \Delta''_{j,L^j},$$

where $\Delta'_{j,L^j} = (s_{j,L^j+1} - s_j) - (t_{j,L^j+1} - t_j)$ and $\Delta''_{j,L^j} = (s_{j,L^j} - s_j) - (t_{j,L^j} - t_j)$. Since $\Delta''_{j,L^j} \in [0, c^{-k}L^j]$, we infer

$$\Delta'_{j,L^j} - c^{-k}L^j \leqslant \Delta_{j,L^j} \leqslant \Delta'_{j,L^j}. \tag{B.7a}$$

We now distinguish three cases depending on the values of the endpoints.

1. If $N^j = P^j = L^j$, we know that $s_{j,L^j+1} = s_{j+1}$, $t_{j,L^j+1} = t_{j+1}$, and therefore $\Delta'_{j,L^j} = \Delta_j$. Relation (B.7a) then implies (B.5).
2. If $N^j > P^j = L^j$, we know that $s_{j,L^j+1} = s_{j+1}$ and therefore $\Delta_j \leqslant \Delta'_{j,L^j}$. Equation (B.7a) then implies $\Delta_j - c^{-k}L^j \leqslant \Delta_{j,L^j}$. We also know that
$$s_{j+1} - s_{j,L^j} \leqslant \{[c^k(t_{j,L^j+1} - t_{j,L^j})] + 1\}/c^k \leqslant (t_{j,L^j+1} - t_{j,L^j}) + c^{-k}, \tag{B.8}$$
or equivalently $\Delta_{j,L^j} \leqslant c^{-k}$. Overall, $\Delta_j - c^{-k}L^j \leqslant \Delta_{j,L^j} \leqslant c^{-k}$, and we conclude that (B.5) holds. Moreover, inequality (B.8) implies $-t_{j,L^j+1} \leqslant -t_{j,L^j} + c^{-k} - (s_{j+1} - s_{j,L^j})$, and therefore
$$t_{j+1} - t_{j,L^j+1} \leqslant t_{j+1} - t_{j,L^j} + c^{-k} - (s_{j+1} - s_{j,L^j}) = \Delta''_{j,L^j} - \Delta_j + c^{-k},$$
which implies (B.6).
3. If $P^j > N^j = L^j$, we know that $t_{j,L^j+1} = t_{j+1}$, and therefore $\Delta'_{j,L^j} \leqslant \Delta_j$. We also know that
$$t_{j+1} - t_{j,L^j} \leqslant s_{j,L^j+1} - s_{j,L^j} \tag{B.9}$$
and therefore $\Delta_{j,L^j} > 0$. Relation (B.7a) implies $|\Delta_{j,L^j}| = \Delta_{j,L^j} \leqslant \Delta'_{j,L^j} \leqslant \Delta_j$, and inequality (B.5) holds. Moreover, we infer from (B.9) that
$$s_{j+1} - s_{j,L^j+1} \leqslant s_{j+1} - s_{j,L^j} - (t_{j+1} - t_{j,L^j}) = \Delta_j - \Delta''_{j,L^j},$$
which implies (B.6). □

B.1.2. Convergence under Condition 2

For every integers p, k, $(1 \leqslant p \leqslant k)$, consider

$$A_{k,p} = \sum_{j_1=0}^{L} \sum_{j_2=0}^{L^{j_1}} \cdots \sum_{j_p=0}^{L^{j_1,\ldots,j_{p-1}}} |\Delta_{j_1\ldots j_p}|$$

and

$$B_{k,p} = \sum_{j_1=0}^{L} \cdots \sum_{j_{p-1}=0}^{L^{j_1,\ldots,j_{p-2}}} \sum_{j_p=L^{j_1,\ldots,j_{p-1}}+1}^{H^{j_1,\ldots,j_{p-1}}} |\Delta_{j_1\ldots j_p}|.$$

We infer from relation (B.3) that the quantity $\mathbb{E}|\theta^{**k}(1) - \theta_k(1)|$ is bounded above by $\mathbb{E}(\sum_{j_1=0}^{H} \sum_{j_2=0}^{H^{j_1}} \cdots \sum_{j_k=0}^{H^{j_1,\ldots,j_{k-1}}} |\Delta_{j_1\ldots j_k}|)$, or equivalently

$$\mathbb{E}|\theta^{**k}(1) - \theta_k(1)| \leqslant \mathbb{E}(A_{k,k}) + \mathbb{E}(B_{k,1}) + \cdots + \mathbb{E}(B_{k,k}).$$

We can show

Lemma B.2. The inequalities $\mathbb{E}A_{k,p} \leqslant \lambda(1+2b)b^p/c^k$ *and* $\mathbb{E}B_{k,p} \leqslant \lambda(3b+2)b^{p-1}/c^k$ *hold for all* $p \geqslant 1$.

Proof. For any $(j_1,\ldots,j_p)$, inequality (B.4) and Lemma B.1 imply

$$\sum_{j_p=0}^{L^{j_1,\ldots,j_{p-1}}} |\Delta_{j_1\ldots j_p}| \leqslant |\Delta_{j_1\ldots j_{p-1}}| + \frac{1+2L^{j_1,\ldots,j_{p-1}}}{c^k}$$

and thus $\mathbb{E}A_{k,p} \leqslant \mathbb{E}A_{k,p-1} + c^{-k}\mathbb{E}[\sum_{j_1=0}^{L} \sum_{j_2=0}^{L^{j_1}} \cdots \sum_{j_p=0}^{L^{j_1,\ldots,j_{p-2}}} (1+2L^{j_1,\ldots,j_{p-1}})]$. We infer $\mathbb{E}A_{k,p} \leqslant \mathbb{E}A_{k,p-1} + \lambda(1+2b)b^{p-1}/c^k$, and therefore $\mathbb{E}A_{k,p} \leqslant \lambda(1+2b)(\sum_{i=0}^{p-1} b^p)/c^k \leqslant \lambda(1+2b)b^p/c^k$. A similar reasoning provides an upper bound for $\mathbb{E}B_{k,p}$. □

We infer from the lemma that $\mathbb{E}|\theta^{**k}(1) - \theta_k(1)|$ converges to 0 as $k \to \infty$, and therefore $\theta^{**k}(1) = \theta_k(1) + o_p(1) \xrightarrow{d} \theta(1)$. A straightforward adaptation of this proof implies that $\mathbb{E}|\theta^{**k}(t) - \theta_k(t)| \to 0$ for all t, and convergence condition (B.1) is therefore satisfied.

B.1.3. Convergence under Condition 1′

For every integers p and k ($1 \leqslant p \leqslant k$), it is convenient to introduce

$$D_{k,p} = \sum_{j_1=0}^{H} \sum_{j_2=0}^{H^{j_1}} \cdots \sum_{j_p=0}^{H^{j_1,\ldots,j_{p-1}}} (\Delta_{j_1,\ldots,j_p})^2$$

and

$$C_{k,p} = \sum_{j_1=0}^{H} \cdots \sum_{j_k=0}^{H^{j_1,\ldots,j_{k-1}}} \Delta_{j_1\ldots j_k} \times \left(\sum_{i_p \neq j_p}^{H^{j_1,\ldots,j_{p-1}}} \sum_{i_{p+1}=0}^{H^{j_1,\ldots,j_{p-1},i_p}} \cdots \sum_{i_k=0}^{H^{j_1,\ldots,j_{p-1},i_p,\ldots,i_{k-1}}} \Delta_{j_1\ldots j_{p-1} i_p \ldots i_k} \right).$$

Since $\sum_{j_{p+1}=0}^{H^{j_1,\ldots,j_p}} \Delta_{j_1\ldots j_p,j_{p+1}} = \Delta_{j_1\ldots j_p}$, we infer

$$\sum_{i_p \neq j_p}^{H^{j_1,\ldots,j_{p-1}}} \sum_{i_{p+1}=0}^{H^{j_1,\ldots,j_{p-1},i_p}} \cdots \sum_{i_k=0}^{H^{j_1,\ldots,j_{p-1},i_p,\ldots,i_{k-1}}} \Delta_{j_1\ldots j_{p-1}i_p\ldots i_k} = \Delta_{j_1,\ldots,j_{p-1}} - \Delta_{j_1,\ldots,j_p}.$$

It is then easy to show that $\mathbb{E}(D_{k,1}) = -\mathbb{E}(C_{k,1})$ and $\mathbb{E}(D_{k,p}) = \mathbb{E}(C_{k,p-1}) - \mathbb{E}(C_{k,p})$ for all $p \geqslant 2$.

By (B.3), the difference $X_k = \theta^{**k}(1) - \theta_k(1)$ has second moment $\mathbb{E}(X_k^2) = [\mathbb{E}(M^2)]^k \mathbb{E}C_{k,k} + \sum_{p=1}^{k} [\mathbb{E}(M^2)]^{p-1}\mathbb{E}D_{k,p}$, or equivalently

$$\mathbb{E}(X_k^2) = Var(M) \sum_{p=1}^{k} [\mathbb{E}(M^2)]^{p-1}\mathbb{E}(C_{k,p}).$$

A heuristic argument suggests that $[\mathbb{E}(M^2)]^p \mathbb{E}(C_{k,p}) \sim [\mathbb{E}(M^2)]^p b^p (c^{-k})^2$, implying that $\mathbb{E}(X_k^2)$ converges to zero if $\mathbb{E}(M^2)bc^{-2} < 1$. Using Lemma A.1, it is straightforward to show that there exists a positive number $R_{\lambda,b}$ such that $\mathbb{E}(C_{k,p}) \leqslant R_{\lambda,b} c^{-2k}(p+1)b^{p+2}$ for all k, p. We then conclude that $\mathbb{E}(X_k^2) \to 0$ under Condition 1′.

B.2. Tightness

Given $\delta = c^{-l}$, we consider the regularly spaced grid $t_0 = 0 < t_1 = \delta < \cdots < t_n = 1$, where $n = 1/\delta$. For any $k \geqslant 1$, we infer, as in Appendix 8.1, that

$$\mathbb{P}\{w(\theta^{**k}, \delta) \geqslant \varepsilon\} \leqslant \delta^{-1}\mathbb{P}\{\theta^{**k}(\delta) \geqslant \varepsilon/3\}.$$

Since $\theta^{**k}(\delta) \xrightarrow{\text{d}} \theta(\delta)$, the function $\limsup_k \mathbb{P}\{w(\theta^{**k}, \delta) \geqslant \varepsilon\}$ is therefore bounded above by $\delta^{-1}\mathbb{P}\{\theta(\delta) \geqslant \varepsilon/3\}$. Given a number $q > 0$ satisfying $\tau(q > 0$, Chebyshev's inequality implies that $\limsup_k \mathbb{P}\{w(\theta^{**k}, \delta) \geqslant \varepsilon\}$ is bounded above by $(3/\varepsilon)^q \delta^{-1}\mathbb{E}[\theta(\delta)^q]$. Letting $\delta \to 0$, we infer that $\delta^{-1}\mathbb{E}[\theta(\delta)^q] \sim c_{\lambda,q}\, \delta^{\tau(q)} \to 0$ and conclude that condition (B.2) is satisfied. We have thus established that θ^{**k} weakly converges to θ as $k \to \infty$.

B.3. Convergence of θ^{*k}

We now check that the discretized trading time θ^{*k} also converges to θ. The difference $Y_k = \theta^{*k}(1) - \theta^{**k}(1)$ satisfies

$$Y_k = c^{-k} \sum_{j_1=0}^{P} \sum_{j_2=0}^{P^{j_1}} \sum_{j_k=0}^{P^{j_1,\ldots,j_{k-1}}} \sum_{t \in I_{j_1,\ldots,j_k}} M_{j_1} M_{j_1,j_2} \cdots M_{j_1\ldots j_k}(\Omega_t - 1).$$

Since $\mathbb{E}(\Omega_s - 1)(\Omega_t - 1)] = 0$ when $s \neq t$, its second moment can be rewritten $\mathbb{E}(Y_k^2) = c^{-2k}\mathbb{E}[\sum_{j_1=0}^{P} \cdots \sum_{j_k=0}^{P^{j_1,\ldots,j_{k-1}}} \sum_{t \in I_{j_1,\ldots,j_k}} M_{j_1}^2 \cdots M_{j_1\ldots j_k}^2 (\Omega_t - 1)^2]$, or equivalently

$$\mathbb{E}(Y_k^2) = c^{-k}(\mathbb{E}M^2)^k Var(\Omega_t).$$

Recall that $\Omega_t \overset{\mathrm{d}}{=} \Omega(\lambda T b^k / c^k)$. When $b = c$ and $\mathbb{E}M^2 < b$, the distribution of Ω_t is independent of k and $\mathbb{E}(Y_k^2)$ converges to 0. When $b < c$ and $\mathbb{E}M^2 < b$, Proposition 3 implies

$$\mathbb{E}(\Omega_t^2) \sim c_q(\lambda T b^k / c^k)^{-\log_b(\mathbb{E}M^2)} \sim d_q(\mathbb{E}M^2)^{-k} c^{k \log_b(\mathbb{E}M^2)},$$

$Var(\Omega_t) \sim \mathbb{E}(\Omega_t^2) \to +\infty$, and $\mathbb{E}(Y_k^2) \sim d_q c^{k[\log_b(\mathbb{E}M^2)-1]} \to 0$.

7. For further reading

Andersen et al., 2001; Bollerslev and Mikkelsen (1996); Calvet et al., 1997; Diebold et al., 1998.

References

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001. The distribution of realized exchange rate volatility. Journal of the American Statistical Association 96, 42–55.

Baillie, R.T., 1996. Long memory processes and fractional integration in econometrics. Journal of Econometrics 73, 5–59.

Billingsley, P., 1999. Convergence of Probability Measures, 2nd Edition. Wiley, New York.

Bollerslev, T., Mikkelsen, H.O., 1996. Modeling and pricing long memory in stock market volatility. Journal of Econometrics 73, 151–184.

Calvet, L., Fisher, A., 1999. Multifractality in asset returns: Theory and evidence. Working paper.

Calvet, L., Fisher, A., Mandelbrot, B., 1997, Large deviation theory and the distribution of price changes. Cowles Foundation Discussion Paper #1165.

Dacorogna, M., Müller, U., Nagler, R., Olsen, R., Pictet, O., 1993. A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. Journal of International Money and Finance 12, 413–438.

Davidson, J., 1994. Stochastic Limit Theory. Oxford University Press, Oxford.

Diebold, F.X., Gunther, T.A., Tay, A.S., 1998. Evaluating density forecasts, with applications to financial risk management. International Economic Review 39, 863–883.

Ding, Z., Granger, C.W.J., Engle, R.F., 1993. A long memory property of stock returns and a new model. Journal of Empirical Finance 1, 83–106.

Dothan, M., 1990. Prices in Financial Markets. Oxford University Press, Oxford.

Durrett, R., 1996. Stochastic Calculus. CRC Press, New York.

Fisher, A., Calvet, L., Mandelbrot, B., 1997. Multifractality of Deutsche Mark/US dollar exchange rate. Cowles Foundation Discussion Paper #1166.

Harrison, J.M., Kreps, D., 1979. Martingales and arbitrage in multiperiod securities markets. Journal of Economic Theory 20, 381–408.

Kurz, M., 1994. On rational belief equilibria. Economic Theory 4, 859–876.

Mandelbrot, B., Fisher, A., Calvet, L., 1997. A multifractal model of asset returns. Cowles Foundation Discussion Paper #1164.

Pollard, D., 1984. Convergence of Stochastic Processes. Springer, Berlin.

Taylor, S., 1986. Modelling Financial Time Series. Wiley, New York.