Pricing Options on Assets with Predictable White Noise Returns [1]

Angel León (University of Alicante)

Enrique Sentana (CEMFI)

October 1996

Revised: May 1997

[1] We would like to thank Sudipto Bhattacharya, John Campbell, Gabriele Fiorentini, Lars Hansen, Albert Marcet, Rafael Repullo, Ken Singleton, Luis Viceira and David Webb for very useful discussions. We are also grateful to seminar audiences at CEMFI and LSE for valuable comments. Of course, the usual caveat applies.

Correspondence should be addressed to E. Sentana, CEMFI, Casado del Alisal 5, 28014 Madrid, Spain (tel.: +34 1 429 0551, fax: +34 1 429 1056, e-mail: sentana@cemfi.es).

Abstract

We study the effect of the predictability of an asset's return on the prices of options on that asset, for models in which returns are serially uncorrelated, yet predictable on the basis of a larger information set. We show that return predictability may matter in a discrete time world, especially for longer maturity options. However, discrepancies between the frequency of trading and observation become relevant in estimating the model parameters. When trading is continuous, Black-Scholes is valid, and the sample variance of holding returns over finite periods is an appropriate estimator of the variance of instantaneous returns.

1 Introduction

In a recent paper, Lo and Wang (1995) convincingly argue that the predictability of an asset's returns may affect the prices of options written on that asset, even though predictability is induced by the drift, which does not enter the option pricing formula. The rationale is that, unlike in the geometric Brownian motion process with constant drift underlying the standard Black-Scholes formula, the sample variance of discretely-sampled returns may not be an appropriate estimator of the instantaneous variance if returns are predictable. Lo and Wang (1995) show that this is indeed the case for univariate and multivariate continuous time AR(1) processes which imply serially correlated returns.

We analyze the same issue for models in which asset returns are serially uncorrelated (i.e. white noise), and therefore unpredictable from their past history alone, but they are predictable on the basis of a larger information set. In other words, the market for the primitive asset is weak- but not semistrong-form efficient. The justification for such models lies at the core of the mean-reversion literature (see e.g. Shiller (1984), Summers (1986), Poterba and Summers (1988), or Fama and French (1988)), and simply reflects the fact that negligible autocorrelations for observed returns are compatible not only with constant expected returns, but also with a smoothly time-varying expected return process whose first-order autocorrelation is high (see also Campbell (1991)). Furthermore, such models are not only a theoretical possibility. As pointed out by Campbell, Lo and MacKinlay (1997, p. 267), "this possibility seems to be empirically relevant for the US stock market".

The paper is organized as follows. In section 2, we introduce a general discrete-time version of the price process which generates white noise returns. This process nests a conditional version of the binomial model, which we use to assess whether return predictability is potentially important for option valuation within a preference-free framework. We also analyze the consequences of discrepancies between the frequency of trading and the frequency of observation of prices. Then, in section 3 we derive the continuous-time diffusion which aggregates exactly to the discrete-time model, and analyze the effects of return predictability in the limiting case of continuous trading. In order to gain some intuition, we also consider a discrete state approximation to the continuous time model. Our conclusions can be found in section 4.

2 Discrete Time Analysis

2.1 A Discrete Time Model

Let $p(t)$ denote the (log) price at instant t of a risky asset which pays no dividends, and let $x(t)$ be a predictor variable which Granger-causes prices. Let's initially consider a discrete-time world in which the highest frequency is 1. Campbell (1991) and Fiorentini and Sentana (1996) show that if the joint data generation process for $\Delta_1 p\,(t) = p(t) - p(t-1)$ and $x(t)$ is given by the following reduced-rank bivariate VAR(1):

$$\begin{pmatrix} \Delta_1 p\,(t) \\ x\,(t) - \mu \end{pmatrix} = \begin{pmatrix} \mu \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} \Delta_1 p\,(t-1) \\ x\,(t-1) - \mu \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,1}\,(t) \\ \varepsilon_{2,1}\,(t) \end{pmatrix} \quad (1)$$

where $|\alpha| < 1$ and $\boldsymbol{\varepsilon}_1(t) = (\varepsilon_{1,1}\,(t)\,, \varepsilon_{2,1}\,(t))'$ is a martingale difference sequence with $E\left[\boldsymbol{\varepsilon}_1(t)|I(t-1)\right] = \mathbf{0}$ and $V\left[\boldsymbol{\varepsilon}_1(t)|I(t-1)\right] = \boldsymbol{\Sigma}_1 = \begin{pmatrix} \sigma^2_{1,1} & \sigma_{12,1} \\ \sigma_{12,1} & \sigma^2_{2,1} \end{pmatrix}$, then the (continuously-compounded) return process $\Delta_1 p(t)$ is white noise with constant variance $\omega^2 = \sigma^2_{1,1} + \frac{\sigma^2_{2,1}}{1-\alpha^2} \geq \sigma^2_{1,1}$, provided that

$$\sigma_{12,1} = \frac{-\alpha}{1-\alpha^2}\sigma^2_{2,1} \quad (2)$$

and

$$\sigma_{1,1}^2 \geq \left(\frac{\alpha}{1-\alpha^2}\right)^2 \sigma_{2,1}^2$$

Similarly, since $\Delta_k p(t) = \sum_{j=0}^{k-1} \Delta_1 p(t-j)$, k-period holding returns, with k integer, will also be white noise, so that the variance ratio $Var(\Delta_k p(t))$ over $kVar(\Delta_1 p(t))$ will be 1 for all k.

The implications of condition (2) are perhaps easier to understand if we consider the impulse response functions of price changes with respect to the different shocks (see Fiorentini and Sentana (1996)). For the relevant case of $\alpha > 0$, the negative correlation between innovations implies that the initial positive effect on $\Delta_1 p(t)$ of a shock to $\varepsilon_{1,1}(t)$ is slowly compensated by the negative but decaying impact on $x(t)$. More interestingly, a shock to $\varepsilon_{2,1}(t)$ has a very large negative immediate impact on $\Delta_1 p(t)$, which is then slowly reversed by the positive and decaying effect on $x(t)$. The response patterns are such that a white noise marginal process is obtained for $\Delta_1 p(t)$. Campbell (1991) provides an economic intuition for such a negative correlation in the context of a dynamic Gordon growth model.

However, lack of autocorrelation at all horizons should not be taken as evidence in favour of constant expected returns. In this model, one-period holding returns are predictable on the basis of $x(t)$, which can actually be interpreted as expected returns.[1] In fact, depending on the parameter values, the R^2 of the theoretical regression of $\Delta_1 p(t)$ on $x(t-1)$ may be substantial (see Fiorentini and Sentana (1996)). Furthermore, the degree of predictability is horizon-dependent, in the sense that the ratio of the variance of the k-period ahead forecast error, $\sigma_{1,k}^2$, to the variance of the k-period return, $k\omega^2$, is a

[1] In this respect, Fiorentini and Sentana (1996) show that any reduced rank VAR(1) for $\Delta_1 p(t)$ and other variable $\delta(t)$ which Granger causes it (with a dense companion matrix) can be written as (1) with $x(t-1) = E[\Delta_1 p(t)|I(t-1)]$. The reduced rank restriction simply guarantees that expected returns follow an AR process of order not higher than 1.

nonlinear function of k (see Campbell (1993) and the discussion below). Figure 1 presents a plot of these two variances (with the normalization $\omega^2 = 1$) for parameter values broadly representative of post-war monthly US stock market returns when the corresponding lagged dividend yield is used as predictor variable (see Fiorentini and Sentana (1996), or chapter 7 of Campbell, Lo and MacKinlay (1997) for details). In particular, we choose $\alpha = 0.98$, $\rho_{12,1} = cor(\varepsilon_{1,1}(t), \varepsilon_{2,1}(t)) = -0.63$ and $R_1^2 = 1.6\%$. As can be seen, $\sigma_{1,k}^2$ is very close to k in the short-run, but then it becomes significantly smaller in the medium-run, although eventually it increases linearly again, as the long-run forecast of $x(t)$ is simply its unconditional mean, μ.

Nevertheless, given that predictability disappears under a risk neutralized measure, what is important for implementing option pricing models is to use the correct values of the relevant parameters (see Lo and Wang (1995)). In particular, since $Var(\Delta_1 p(t)) = Var(x(t)) + Var(\varepsilon_{1,1}(t)) \geq \sigma_{1,1}^2$, with equality if and only if $\sigma_{2,1}^2 = 0$, option prices computed under the assumption that $p(t)$ is a geometric random walk with constant drift and variance ω^2 may well be wrong. As pointed out by Lo and Wang (1995), the effect on prices may be particularly important for longer maturity options, even with small levels of predictability, since an option's vega is an increasing function of time to maturity.

2.2 A Discrete State Version of the Discrete Time Model

In order to assess within a preference-free framework whether return predictability is potentially important for option valuation in discrete time, we shall use a conditional version of the binomial tree approach of Cox, Ross and Rubinstein (1979), in which expected returns follow a discrete-time two-state Markov chain.

Let $z(t)$ be a binary variable which indicates whether expected returns at period t are low or high. We assume that $\Delta_1 p(t)$ and $x(t)$ evolve according to (1), with the following conditional distributions for $\boldsymbol{\varepsilon}_1(t)$:

a) when $z(t-1)=0$

	$z(t)=0$	$z(t)=1$	
$\varepsilon_{1,1}(t) \diagdown \varepsilon_{2,1}(t)$	$-\sigma_{2,1}\sqrt{\frac{1-q_{0,1}}{q_{0,1}}}$	$\sigma_{2,1}\sqrt{\frac{q_{0,1}}{1-q_{0,1}}}$	
$\sigma_{1,1}\sqrt{\frac{1-\pi_{0,1}}{\pi_{0,1}}}$	(1) $\pi_{0,1}$	0	$\pi_{0,1}$
$-\sigma_{1,1}\sqrt{\frac{\pi_{0,1}}{1-\pi_{0,1}}}$	(2) $q_{0,1}-\pi_{0,1}$	(3) $1-q_{0,1}$	$1-\pi_{0,1}$
	$q_{0,1}$	$1-q_{0,1}$	

b) when $z(t-1)=1$

	$z(t)=1$	$z(t)=0$	
$\varepsilon_{1,1}(t) \diagdown \varepsilon_{2,1}(t)$	$\sigma_{2,1}\sqrt{\frac{1-q_{1,1}}{q_{1,1}}}$	$-\sigma_{2,1}\sqrt{\frac{q_{1,1}}{1-q_{1,1}}}$	
$-\sigma_{1,1}\sqrt{\frac{1-\pi_{1,1}}{\pi_{1,1}}}$	(4) $\pi_{1,1}$	0	$\pi_{1,1}$
$\sigma_{1,1}\sqrt{\frac{\pi_{1,1}}{1-\pi_{1,1}}}$	(5) $q_{1,1}-\pi_{1,1}$	(6) $1-q_{1,1}$	$1-\pi_{1,1}$
	$q_{1,1}$	$1-q_{1,1}$	

In the relevant case of $\rho_{12,1}<0$, if we choose $q_{0,1}=q_{1,1}=(\alpha+1)/2=q_1$ and $\pi_{0,1}=\pi_{1,1}=q_1\rho_{12,1}^2/[(1-q_1)+q_1\rho_{12,1}^2]$, then it is straightforward to show that $E\left[\boldsymbol{\varepsilon}_1(t)|z(t-1)=0\right]=E\left[\boldsymbol{\varepsilon}_1(t)|z(t-1)=1\right]=\mathbf{0}$ and $V\left[\boldsymbol{\varepsilon}_1(t)|z(t-1)=0\right]=V\left[\boldsymbol{\varepsilon}_1(t)|z(t-1)=1\right]=\boldsymbol{\Sigma}_1$ as required.

Graphically, the structure of the tree is:

$$
\begin{array}{ccccccccc}
 & & & & z(t-1)=0 & & & & \\
 & & & \swarrow & \downarrow & \searrow & & & \\
 & (1) & & & (2) & & & (3) & \\
\swarrow & \downarrow & \searrow & \swarrow & \downarrow & \searrow & \swarrow & \downarrow & \searrow \\
(1) & (2) & (3) & (1) & (2) & (3) & (4) & (5) & (6)
\end{array}
$$

$$
\begin{array}{ccccccccc}
 & & & & z(t-1)=1 & & & & \\
 & & & \swarrow & \downarrow & \searrow & & & \\
 & (4) & & & (5) & & & (6) & \\
\swarrow & \downarrow & \searrow & \swarrow & \downarrow & \searrow & \swarrow & \downarrow & \searrow \\
(4) & (5) & (6) & (4) & (5) & (6) & (1) & (2) & (3)
\end{array}
$$

The main advantage of such a conditional binomial model is that even though there are three possible states of nature for each value of $z(t-1)$, in two of them, namely (2)-(3) and (5)-(6), the price of the risky asset, $p(t)$, is the same. As a consequence, we can value derivative assets with payoffs 1-period ahead into the future on the basis of the risk-neutralized versions of $\pi_{0,1}$ and $\pi_{1,1}$ alone, despite the fact that markets cannot be fully completed through dynamic trading. It is easy to see that,

$$\bar{\pi}_{0,1} = \frac{e^{(\log R_f - \mu)} - e^{-M_1}}{e^{-m_1} - e^{-M_1}}$$

$$\bar{\pi}_{1,1} = \frac{e^{M_1} - e^{(\log R_f - \mu)}}{e^{M_1} - e^{m_1}}$$

where

$$M_1 = \frac{\sigma_{2,1}}{2\sqrt{q_1(1-q_1)}} + \sigma_{1,1}\left|\rho_{12,1}\right|\sqrt{\frac{q_1}{1-q_1}} > 0$$

$$m_1 = \frac{\sigma_{2,1}}{2\sqrt{q_1(1-q_1)}} - \frac{\sigma_{1,1}}{\left|\rho_{12,1}\right|}\sqrt{\frac{1-q_1}{q_1}} \lesseqqgtr 0$$

and R_f is the constant gross return on a safe asset.

Absence of arbitrage opportunities requires $0 \leq \bar{\pi}_{0,1} \leq 1$ and $0 \leq \bar{\pi}_{1,1} \leq 1$, or equivalently $m_1 \leq \mu - \log R_f \leq M_1$ and $-M_1 \leq \mu - \log R_f \leq -m_1$ respectively, which in turn requires at least that $m_1 \leq 0$. In principle, such conditions on the "risk premium" $\mu - \log R_f$ may not be satisfied without further restrictions on the stochastic nature of the return generating process (1). For instance, if we assume that the white noise restriction (2) holds, and make $\mu = \log R_f$, then $\rho^2_{12,1} \leq \alpha/(1+\alpha)$ becomes a necessary and sufficient condition.

For derivative assets such as European call options with maturity at $t+1$, $t+2$, etc., we would need to decompose $\bar{\pi}_{0,1}$ into $\bar{q}_{0,1} - \bar{\pi}_{0,1}$ and $1 - \bar{q}_{0,1}$, and similarly $\bar{\pi}_{1,1}$ into $\bar{q}_{1,1} - \bar{\pi}_{1,1}$ and $1 - \bar{q}_{1,1}$. Although this is impossible without knowing the price of some other asset, we nevertheless know that $\bar{\pi}_{0,1} \leq \bar{q}_{0,1} \leq 1$ and $\bar{\pi}_{1,1} \leq \bar{q}_{1,1} \leq 1$. Therefore, we can bound the derivative price by computing it for every possible pair of admissible values of $\bar{q}_{0,1}, \bar{q}_{1,1}$.[2] It turns out that the bounds obtained in this way are very tight for the parameters values considered in section 2.1.

An interesting situation arises when $\mu = \log R_f$, condition (2) is satisfied, and $\rho^2_{12,1} = \alpha/(1+\alpha)$. In this case, the prices of some of the implicit contingent commodities are 0, and it turns out that the prices of European call options on the risky asset are independent of $z(t-1)$ for all exercise prices and maturities. However, as soon as $\rho^2_{12,1} < \alpha/(1+\alpha)$, option prices generally depend on $z(t-1)$. Notable exceptions are at-the-money options, whose prices turn out to be independent of whether expected returns are low or high.

Figure 2 presents the value of 1 and 4-period European call options as a function of the strike price, both when expected returns are low and when they are high.[3] As a normalization, we fix the current price of the underlying risky asset to 1, and take $\mu = \log R_f = 0$. Note that irrespectively of whether expected returns are low or high, the call price is the same as the value of the primitive asset (i.e. 1) when the strike price is 0, and tends to 0 as the strike

[2] For this purpose, it is convenient to make $\bar{q}_{0,1} - \bar{\pi}_{0,1} = \bar{\lambda}_0(1-\bar{\pi}_{0,1})$, $1-\bar{q}_{0,1} = (1-\bar{\lambda}_0)(1-\bar{\pi}_{0,1})$, $\bar{q}_{1,1} - \bar{\pi}_{1,1} = \bar{\lambda}_1(1-\bar{\pi}_{1,1})$ and $1-\bar{q}_{1,1} = (1-\bar{\lambda}_1)(1-\bar{\pi}_{1,1})$ with $\bar{\lambda}_0, \bar{\lambda}_1 \in [0,1]$.

[3] Given the structure of our model, the pseudo-pricing function discussed in Hansen and Richard (1987), which does not take into account the values of the conditioning variable, yields simply the equally-weighted average of the asset prices when $z(t-1) = 0$ and $z(t-1) = 1$.

price goes to infinite. Since options prices must be convex with respect to the strike price, and the value of at-the-money options does not depend on the state variable in this model, out–of-the-money calls attain higher prices when $z(t-1) = 1$ than when $z(t-1) = 0$, while the opposite happens to in-the-money calls. A similar pattern arises for 4-period call options. However, the difference between prices is substantially higher for 4-period options than for 1-period ones, which is in line with the evidence in Lo and Wang (1995).

2.3 The Consequences of Time Aggregation

In order to capture the predictability in returns, we would have to estimate the bivariate VAR(1) process (1). Unfortunately, if returns are predictable, any discrepancy between the frequency of trading and the frequency of observation of prices becomes relevant in estimating (1).[4]

In particular, suppose that as econometricians, we only observe $p(t)$ and $x(t)$ every k trading periods. If we assume for simplicity that $x(t)$ is a stock variable, the time-aggregated joint process for $\Delta_k p(t) = p(t) - p(t-k)$ and $x(t)$ is also a VAR(1).[5] Specifically

$$\begin{pmatrix} \Delta_k p(t) \\ x(t) - \mu \end{pmatrix} = \begin{pmatrix} \mu k \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & \frac{1-\alpha^k}{1-\alpha} \\ 0 & \alpha^k \end{pmatrix} \begin{pmatrix} \Delta_k p(t-k) \\ x(t-k) - \mu \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,k}(t) \\ \varepsilon_{2,k}(t) \end{pmatrix}$$

where $\boldsymbol{\varepsilon}_k(t)$ is such that $E[\boldsymbol{\varepsilon}_k(t)|I(t-k)] = \mathbf{0}$ and $V[\boldsymbol{\varepsilon}_k(t)|I(t-k)] = \boldsymbol{\Sigma}_k$, with

$$\sigma_{1,k}^2 = k\sigma_{1,1}^2 + \frac{2\sigma_{12,1}}{1-\alpha}(k - \frac{1-\alpha^k}{1-\alpha}) + \frac{\sigma_{2,1}^2}{(1-\alpha)^2}[k + \frac{1-\alpha^{2k}}{1-\alpha^2} - \frac{2(1-\alpha^k)}{1-\alpha}]$$

---

[4] We are grateful to John Campbell for bringing this point to our attention.

[5] In order to find out the process for the temporally aggregated data, it is more convenient to re-write (1) as $\begin{pmatrix} p(t) \\ x(t)-\mu \end{pmatrix} = \begin{pmatrix} \mu \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} p(t-1) \\ x(t-1)-\mu \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,1}(t) \\ \varepsilon_{2,1}(t) \end{pmatrix}$ and then recursively substitute backwards. Since the first column of the VAR(1) companion matrix is $(1,0)'$ for all k, we can easily write back the time-aggregated process in terms of k-period returns.

$$\sigma_{12,k} = \frac{(1-\alpha^k)\sigma_{12,1}}{1-\alpha} + \frac{\sigma^2_{2,1}}{1-\alpha}(\frac{1-\alpha^k}{1-\alpha} - \frac{1-\alpha^{2k}}{1-\alpha^2})$$

$$\sigma^2_{2,k} = \sigma^2_{2,1}\frac{1-\alpha^{2k}}{1-\alpha^2}$$

It is then straightforward to show that if (2) holds,

$$\sigma^2_{1,k} = k\sigma^2_{1,1} + \left[k - \left(\frac{1-\alpha^k}{1-\alpha}\right)^2\right]\frac{\sigma^2_{2,1}}{1-\alpha^2} = k\omega^2 - \left(\frac{1-\alpha^k}{1-\alpha}\right)^2\frac{\sigma^2_{2,1}}{1-\alpha^2} \leq k\omega^2$$

When $\alpha \leq \sqrt{2}-1$, $k\sigma^2_{1,1} \leq \sigma^2_{1,k}$ for all k. However, when $\alpha > \sqrt{2}-1$, it is possible that $\sigma^2_{1,k}/k \leq \sigma^2_{1,1}$ for k less than $1/(1-\alpha)^2$. For instance, if $\alpha = 0.98$, $\sigma^2_{1,k}/k \leq \sigma^2_{1,1}$ for any $k \leq 2500$.

The obvious solution to this problem is to recognize the temporal aggregation explicitly, and estimate the VAR(1) above from data sampled at frequency k in terms of the parameters of the underlying process (1). In many cases of interest, though, the ratio of the actual trading frequency to the frequency of observation will be unknown. Therefore, it seems natural to analyze the limiting case of continuous trading.

3 Continuous Time Analysis

3.1 A Continuous Time Diffusion

Let $\mathbf{y}(t) = \begin{pmatrix} p(t) \\ x(t) \end{pmatrix}$, $\boldsymbol{\upsilon} = \begin{pmatrix} \mu(1+\frac{\ln\alpha}{1-\alpha}) \\ -\mu\ln\alpha \end{pmatrix}$ and $\mathbf{A} = \begin{pmatrix} 0 & \frac{\ln\alpha}{1-\alpha} \\ 0 & \ln\alpha \end{pmatrix}$ with $0 < \alpha < 1$, and consider the following continuous-time diffusion

$$d\mathbf{y}(t) = [\boldsymbol{\upsilon} + \mathbf{A}\mathbf{y}(t)]dt + d\mathbf{W}(t) \tag{3}$$

where $\mathbf{W}(t)$ is a bivariate Wiener process with $E\left[d\mathbf{W}(t)\,d\mathbf{W}(t)'\right] = \boldsymbol{\Sigma}_0 dt$ and $\mathbf{y}(0) = \mathbf{y}_0$. It is well known (see e.g. Arnold (1974)) that the solution to the above system of first-order stochastic linear differential equations is

$$\mathbf{y}(t) = e^{\mathbf{A}t}\mathbf{y}_0 + \int_0^t e^{\mathbf{A}(t-r)}\boldsymbol{\upsilon}dr + \int_0^t e^{\mathbf{A}(t-r)}d\mathbf{W}(r),\ t \geq 0 \tag{4}$$

It is also well known that the exact discretization of a multivariate Ornstein-Uhlenbeck process such as (4) satisfies the following system of first-order stochastic linear difference equations:

$$\mathbf{y}\left(t\right)=\mathbf{g}_{h}+\mathbf{F}_{h}\mathbf{y}\left(t-h\right)+\boldsymbol{\varepsilon}_{h}\left(t\right),\ t=h,2h,... \tag{5}$$

where $\mathbf{g}_h=\int_0^h e^{\mathbf{A}r}\boldsymbol{\upsilon}dr$, $\mathbf{F}_h=e^{\mathbf{A}h}$, $\boldsymbol{\varepsilon}_h\left(t\right)\sim iid\,N(0,\boldsymbol{\Sigma}_h)$ and $\boldsymbol{\Sigma}_h=\int_0^h e^{\mathbf{A}r}\boldsymbol{\Sigma}_0 e^{\mathbf{A}'r}dr$ for any positive real number h (see also Bergstrom (1984)).

In our case, given the structure of the matrix $\mathbf{A}$, we obtain

$$\mathbf{g}_h=\begin{pmatrix}\mu(h-\frac{1-\alpha^h}{1-\alpha})\\ \mu(1-\alpha^h)\end{pmatrix}$$

$$\mathbf{F}_h=\begin{pmatrix}1 & \frac{1-\alpha^h}{1-\alpha}\\ 0 & \alpha^h\end{pmatrix}$$

$$\sigma^2_{1,h}=h\sigma^2_{1,0}+\frac{2}{1-\alpha}\left(h-\frac{\alpha^h-1}{\ln\alpha}\right)\sigma_{12,0}+\frac{1}{(1-\alpha)^2}\left(h+\frac{\alpha^{2h}-4\alpha^h+3}{2\ln\alpha}\right)\sigma^2_{2,0}$$

$$\sigma_{12,h}=\frac{\alpha^h-1}{\ln\alpha}\left(\sigma_{12,0}-\frac{\alpha^h-1}{2\left(1-\alpha\right)}\sigma^2_{2,0}\right)$$

$$\sigma^2_{2,h}=\frac{\alpha^{2h}-1}{2\ln\alpha}\sigma^2_{2,0}$$

If we take $h=1$, equate to $\boldsymbol{\Sigma}_1$ and solve for $\boldsymbol{\Sigma}_0$, we finally obtain that (3) aggregates exactly to (1) with Gaussian innovations if:

$$\sigma^2_{1,0}=\sigma^2_{1,1}+\frac{2}{(1-\alpha)}\left[1+\frac{\ln\alpha}{1-\alpha}\right]\sigma_{12,1}+\frac{1}{(1-\alpha)^2}\left[1+\frac{2\alpha\ln\alpha}{1-\alpha^2}\right]\sigma^2_{2,1}$$

$$\sigma_{12,0}=-\frac{\ln\alpha}{1-\alpha}\sigma_{12,1}+\frac{\ln\alpha}{1-\alpha^2}\sigma^2_{2,1}$$

$$\sigma^2_{2,0}=-\frac{2\ln\alpha}{1-\alpha^2}\sigma^2_{2,1}$$

Model (3) turns out to be a special case of the bivariate Ornstein-Uhlenbeck process discussed by Lo and Wang (1995), in which there is a unit root but no deterministic trends. However, they only analyze in detail the case of $\sigma_{12,0} = 0$, which necessarily implies serially correlated discretely-sampled returns. In particular, condition (2), which guarantees that holding returns over integer periods will be white noise, is equivalent in this continuous-time framework to:

$$\sigma_{12,0} = \frac{-1}{2\,(1-\alpha)}\sigma^2_{2,0.} \tag{6}$$

In fact, condition (6) implies that $\Delta_h p\,(t)$ is white noise for any positive real number h. To see why, let's express (3) as $\dot{\mathbf{y}}(t) = \boldsymbol{\upsilon} + \mathbf{A}\mathbf{y}(t) + \boldsymbol{\xi}(t)$, where $\boldsymbol{\xi}(t)$ is the "derivative" of $\mathbf{W}(t)$. If we re-write this expression in terms of the "differential" operator D as $(D\mathbf{I} - \mathbf{A})\mathbf{y}(t) = \boldsymbol{\upsilon} + \boldsymbol{\xi}(t)$, it is then easy to prove using well known results on filters (see e.g. Priestley (1981)) that the spectral density of $\dot{p}(t)$ is constant when condition (6) holds.[6]

Nevertheless, returns are still predictable. Specifically, the R^2 of the theoretical regression of $\Delta_h p\,(t)$ on $x(t-h)$ is

$$R^2_h = \frac{-\left(\alpha^h - 1\right)^2}{2h\,(1-\alpha)^2 \ln\alpha}\frac{\sigma^2_{2,0}}{\sigma^2_{1,0}}$$

Figure 3 presents a plot of R^2_h as a function of h for the same parameter values as before. As expected, note that it converges to 0 as $h \to 0$, since the variation induced by the drift is of order h, while the variation induced by the diffusion is of order $h^{1/2}$. Similarly, it also converges to 0 as $h \to \infty$ since the predictability becomes proportionally negligible in the long run. However, note that at the same time R^2_h can be substantial ($\simeq .334$) for $h \simeq 60$.

It is well known that in this continuous time world, the Black-Scholes pricing formula is valid even though the drift is a function of the state variable $x(t)$, and the right value of the variance parameter to use for option pricing should be $\sigma^2_{1,0}$ (see e.g. Lo and Wang (1995)). But the somewhat surprising result

[6] We are grateful to Lars Hansen for suggesting this simpler line of proof.

that we obtain from (6) is that $\sigma_{1,0}^2 = \omega^2$, so that $Var\left(\Delta_h p\left(t\right)\right)/h$ is indeed an appropriate estimator of the volatility of instantaneous returns. Therefore, it seems that the effects of predictability and time aggregation discussed in the previous section exactly offset each other in the (continuous time) limit.

3.2 A Discrete State Approximation to the Continuous Time Model

In order to gain some intuition on the above results, consider the following bivariate, trinomial $iid(\mathbf{0}, \mathbf{I})$ process with equally probable states proposed by He (1990):

$\varepsilon_1 \diagdown \varepsilon_2$	$-\sqrt{\frac{3}{2}}$	0	$\sqrt{\frac{3}{2}}$	
$-\sqrt{2}$	0	(b) $\frac{1}{3}$	0	$\frac{1}{3}$
$\sqrt{\frac{1}{2}}$	(a) $\frac{1}{3}$	0	(c) $\frac{1}{3}$	$\frac{2}{3}$
	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	

Let $\varepsilon_{1,h}(t) = \sigma_{1,h}\varepsilon_1,\ \varepsilon_{2,h}(t) = (\sigma_{12,h}/\sigma_{1,h})\varepsilon_1 + \sqrt{\sigma_{2,h}^2 - \sigma_{12,h}^2/\sigma_{1,h}^2}\varepsilon_2$ and generate $\Delta_h p\left(t\right)$ and $x\left(t\right) - \mu$ according to (5). If $h = 1/N$, with N integer, this process aggregates exactly to (1), albeit with non-Gaussian innovations. At the same time, it converges weakly to (3) as $h \rightarrow 0$ (see e.g. He (1990)).

Note that even though there are three possible states of nature for each value of $x(t-h)$, in two of them, namely (a) and (c), the price of the risky asset, $p(t)$, is the same (cf. section 2.2). As a result, if $\sigma_{2,0}^2 = \sigma_{12,0} = 0$ so that $x\left(t\right) = \mu, \forall t$, we obtain an asymmetric version of the unconditional binomial process of Cox, Ross and Rubinstein (1979).

Let $\bar{\pi}\left[x(t-h)\right]$ be the risk-neutralized probability of state (b) as a function of $x(t-h)$. It is easy to see that

$$\bar{\pi}\left[x(t-h)\right] = \frac{\exp\left\{h(\log R_f - \mu) + \frac{1-\alpha^h}{\alpha-1}[x(t-h)-\mu] - \sqrt{\frac{1}{2}}\sigma_{1,h}\right\} - 1}{\exp\left\{-(\sqrt{2}+\sqrt{\frac{1}{2}})\sigma_{1,h}\right\} - 1}$$

which remains between 0 and 1 provided that

$$-\sqrt{\frac{1}{2}}\sigma_{1,h} \leq h(\mu - \log R_f) + \frac{1-\alpha^h}{1-\alpha}[x(t-h) - \mu] \leq \sqrt{2}\sigma_{1,h}$$

Since $\sigma_{1,h}$ is of order $h^{1/2}$ whereas $h(\mu - \log R_f) + \frac{1-\alpha^h}{1-\alpha}[x(t-h) - \mu]$ is of order h, this lack-of-arbitrage condition is increasingly likely to be satisfied as $h \to 0$.

For positive $h, \bar{\pi}[x(t-h)]$ depends on the deviations of expected returns from its long-term mean, μ. However, it is not difficult to see that such a dependence vanishes at the rate h, and furthermore that $\lim_{h\to 0} \bar{\pi}[x(t-h)] = 1/3$, which is the actual probability of the corresponding state. Since $\sigma_{1,h} = h^{1/2}\sigma_{1,0} + o(h)$ when condition (6) holds, this confirms that the Black-Scholes pricing formula is valid in the limit with ω as the relevant parameter.

We can also use the trinomial model to see whether the continuous trading results provide a reasonable guide when the frequency of trading is small but finite. Given that the number of states after a unit time interval is 3^N when $h = 1/N$, we choose $N = 10$ for simplicity. For the purposes of the exercise, we arbitrarily split 1-$\bar{\pi}[x(t-h)]$ equally between the first and the last state.[7] Figure 4 presents the price of a 1-period European call options as a function of the strike price, when the current expected return is $\pm$one standard deviation away from its mean, μ. Again, we choose $\mu = \log R_f = 0$ and $p(t) = 0$ for normalization. As a benchmark, we also include prices computed on the basis of the Black-Scholes formula, as well as the asymmetric version of the Cox, Ross and Rubinstein (1979) approach mentioned above. Although the two prices computed under the assumption of no predictability are fairly accurate for a wide range of exercise prices, it seems that the initial value of $x(t)$ still

[7] Alternatively, we could assume that $x(t)$ is itself the (detrended) price of another financial asset. In that case, it is possible to prove that the risk-neutralized probabilities of states (a) and (c) also go to 1/3 as $h \to 0$. It turns out that the way in which we split $1 - \bar{\pi}[x(t-h)]$ has only an imperceptible effect on the results displayed in Figure 4.

exerts some influence on option prices over the depicted range, at least for $N = 10$.

4 Conclusions

We analyze the effect of the predictability of an asset's return on the prices of options on that asset, for a class of stochastic processes for prices which yield predictable, yet serially uncorrelated returns. For a conditional version of the binomial tree approach of Cox, Ross and Rubinstein (1979) in which expected returns follow a discrete-time two-state Markov chain, we show that return predictability matters, especially for longer maturity options (cf. Lo and Wang (1995)).

However, in a discrete time world with predictable returns, any discrepancy between the frequency of trading and the frequency of observation of prices becomes relevant in estimating the model parameters. For that reason, we also analyze the limiting case of continuous trading. In such a world, the Black-Scholes option pricing formula is valid despite the predictability, and moreover, the sample variance of holding returns over finite periods turns out to be an appropriate estimator of the variance of instantaneous returns. Therefore, it seems that what is important for implementing option pricing models is not merely the predictability of the asset return, but its serial correlation. In fact, this is also true for more general price processes. In particular, suppose that the drift follows a general linear covariance stationary process, so that the joint model for actual and expected returns can be written as

$$\begin{aligned} \dot{p}(t) &= \mu(t) + \xi_1(t) \\ \mu(t) - \mu &= g(D)\xi_2(t) \end{aligned}$$

For instance, $g(D) = (1 + Db_1 + \ldots + D^q b_q)/(D^p + D^{p-1}a_1 + \ldots + a_p)$ for the continuous ARMA(p,q) process discussed in Brockwell (1995). Here, returns are serially uncorrelated over any frequency λ if (and only if) $|g(i\lambda)|^2 \sigma_{22,0} + (g(i\lambda) + g(-i\lambda)) \sigma_{12,0}$ is not a function of λ. In that case $Var(\Delta_h p(t))/h$ is still an appropriate estimator of the volatility of instantaneous returns, despite the fact that returns remain predictable as long as $g(D) \neq 0$.

References

Arnold, L. (1974): *Stochastic Differential Equations: Theory and Applications*, Wiley, New York.

Bergstrom, A.R. (1984): "Continuous Time Stochastic Models and Issues of Aggregation over Time", in *Handbook of Econometrics* vol. 2, ed. by Z. Griliches and M.D. Intriligator, North Holland, Amsterdam.

Brockwell, P.J. (1995): "A Note on the Embedding of Discrete-Time ARMA Processes", *Journal of Time Series Analysis* 16, 451-460.

Campbell, J.Y. (1991): "A Variance Decomposition for Stock Returns", *Economic Journal* 101, 157-179.

Campbell, J.Y. (1993): "Why Long Horizons: A Study of Power against Persistent Alternatives", NBER Technical Working Paper 142.

Campbell, J.Y., Lo A.W. and MacKinlay, A.C. (1997): *The Econometrics of Financial Markets*, Princeton University Press, Princeton.

Cox, J.C., Ross, S.A. and Rubinstein, M. (1979): "Option Pricing: A Simplified Approach", *Journal of Financial Economics* 7, 229-263.

Fama, E.F. and K.R. French (1988): "Permanent and Temporary Components of Stock Prices", *Journal of Political Economy*, 96, 246-273.

Fiorentini, G. and Sentana, E. (1996): "Conditional Means of Time Series Processes and Time Series Processes for Conditional Means", LSE Financial Markets Group Discussion Paper 257.

Hansen, L.P. and Richard, S.F. (1987): "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Theories", *Econometrica* 55, 587-613.

He, H. (1990): "Convergence from Discrete- to Continuous-Time Contingent Claims Prices", *Review of Financial Studies* 3, 523-546.

Lo, A.W. and Wang, J. (1995): "Implementing Option Pricing Models when Asset Returns are Predictable", *Journal of Finance* 50, 87-129.

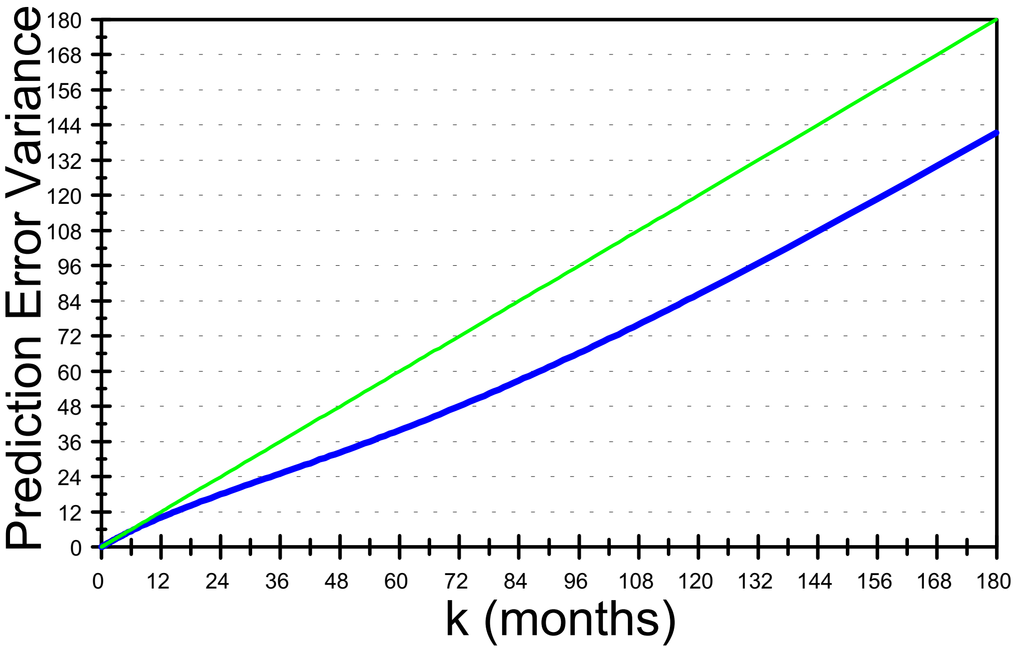
Poterba, J. and Summers L.H. (1988): "Mean Reversion in Stock Returns: Evidence and Implications". *Journal of Financial Economics*, 22, 27-60.

Priestley, M.B. (1981): *Spectral Analysis and Time Series*, Academic Press, London.

Shiller, R.J. (1984): "Stock Prices and Social Dynamics". *Brookings Papers on Economic Activity*, 2, 457-498.

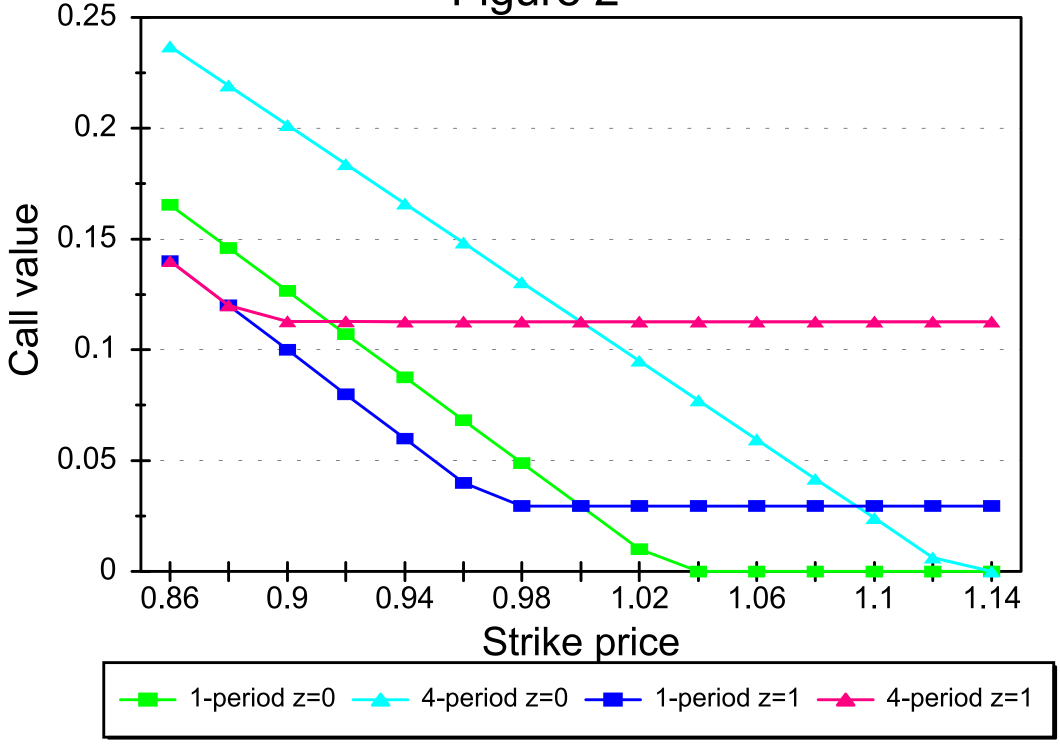
Summers, L.H. (1986): "Does the Stock Market Rationality Reflect Fundamental Values ?", *Journal of Finance,* 41, 591-601.

Figure 1



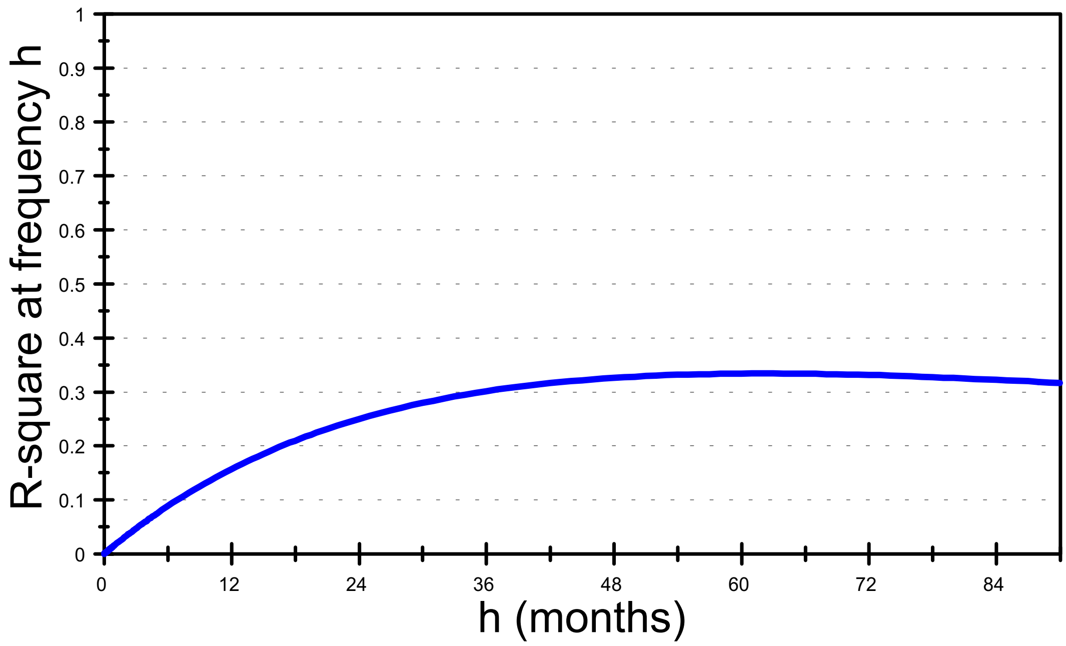
Prediction Error Variance
180
168
156
144
132
120
108
96
84
72
60
48
36
24
12
0
0 12 24 36 48 60 72 84 96 108 120 132 144 156 168 180
k (months)

Figure 2



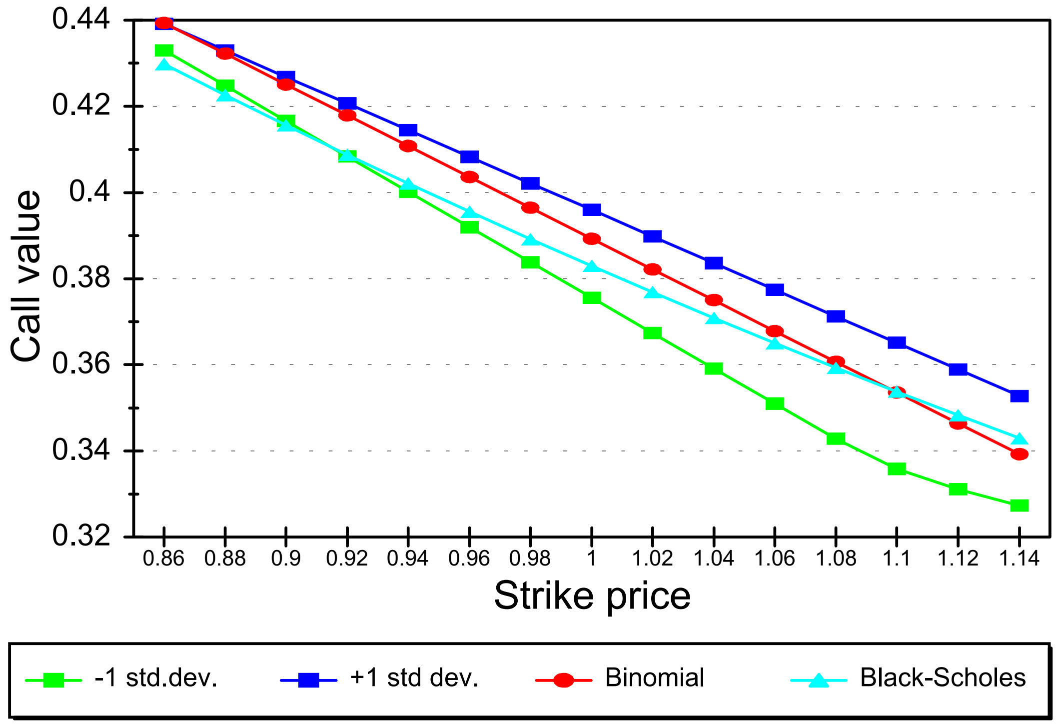
Call value
0.25
0.2
0.15
0.1
0.05
0
0.86
0.9
0.94
0.98
1.02
1.06
1.1
1.14
Strike price
1-period z=0
4-period z=0
1-period z=1
4-period z=1

Figure 3



R-square at frequency h
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
12
24
36
48
60
72
84
h (months)

Figure 4


0.44
0.42
0.4
0.38
0.36
0.34
0.32
Call value
0.86 0.88 0.9 0.92 0.94 0.96 0.98 1 1.02 1.04 1.06 1.08 1.1 1.12 1.14
Strike price
-1 std.dev.
+1 std dev.
Binomial
Black-Scholes