Macroeconomic Dynamics, **2**, 1998, 213–237. Printed in the United States of America.

ENDOGENOUS JUMPING AND ASSET PRICE DYNAMICS

G. C. LIM AND VANCE L. MARTIN
University of Melbourne

LESLIE E. TEO
University of Rochester
and
International Monetary Fund

A model of asset price dynamics is derived in which large jumps in stock prices are determined endogenously. An important property of the model is that it can lead to asset price distributions that are multimodal. The model can explain how relatively small changes in dividends can lead to relatively large changes in asset prices and it can be used to identify the time period in which bubbles begin and end. The framework is applied to modeling the U.S. stock market crash in October 1987. Some forecasting experiments also are conducted with the result that the model is able to predict the size of the eventual crash in the aggregate stock price.

Keywords: Endogenous Jumping, Asset Price Dynamics Stock Market Crash

1. INTRODUCTION

Standard models in finance assume that stock prices are equal to the expected value of discounted future dividends. The earliest tests of the present value model were conducted by LeRoy and Porter (1981) and Shiller (1981) and later by Mankiw et al. (1985, 1991), West (1988), French (1988), Campbell and Shiller (1988), and Stiglitz (1990), who found strong evidence against this model. This led to a number of extensions to the present value model: Cochrane (1992) investigated making the discount factor time varying; Pindyck (1984) allowed for a nonconstant risk premium; Shiller (1988, 1990) and West (1988) included fads and noise trading; Turner et al. (1989) extended the model for regime switching; Diba and Grossman (1988), Dezhbakhsh and Demirguc-Kunt (1990), and Flood and Garber (1980) assumed the existence of a rational bubble; Black (1988), Frankel and Froot (1988), Froot and Obstfeld (1991), and Gennotte and Leland (1990) included heterogenous agents, whereas Barsky and De Long (1993) extended the present value model by distinguishing between actual and permanent returns.

The authors thank Roger Craine, Thomas McCurdy, and economic seminar participants at Berkeley, Guelph, Queens, Rochester, Southern Methodist University, University of Sydney, University of Technology Sydney, and Yale. Address correspondence to: Vance Martin, Department of Economics, University of Melbourne, Parkville, Victoria 3052, Australia.

© 1998 Cambridge University Press 1365-1005/98 $9.50

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

The major aim of this paper is to derive a model that can be used to capture endogenously both the timing and the magnitude of large jumps in asset prices. The approach adopted has much in common with many of the existing approaches for modeling stock-market volatility. An important feature of the model is that it is possible to explain large movements in prices with small movements in the fundamentals. It is also related to the diffusion-process Poisson jump model suggested by Merton (1976) in that the starting point is based on an extension of the linear diffusion equation, which serves as the basis of most models of financial prices; see also Ahn and Thompson (1988); Akgiray and Booth (1988) among others, and Das (1993) for some recent extensions. However, in contrast with the Poisson jump model, jumping in the present model arises from the price process switching between stable equilibria during a period of multiple equilibria. Periods of multiple equilibria correspond to situations in which there is no unique view concerning the equilibrium price in the market. The occurrence of multiple equilibria is modeled by a multimodal distribution in which the modes of the distribution may correspond to views of the separate groups. Jumps occur when agents switch their views, which results in a change in the relative importance of the modes.

A major advantage of this approach is that the timing of the jump is determined endogenously and does not need to be modeled ex post as in the case with models based on dummy variables as adopted by Flood and Garber (1980). This means that it is possible to forecast jumps such as stock-market crashes. The approach also offers a simple descriptive tool for identifying both the start and the end of a period in which there is more than one view of the equilibrium market price. The model can be estimated by standard maximum likelihood methods using iterative optimization routines. An important feature of the model is that higher-order moments in the distribution are used to model the observed price distribution. This is achieved by adopting a distribution from the generalized exponential family that nests the normal distribution as a special case.

The rest of the paper proceeds as follows. The endogenous switching model is derived in Section 2. Estimation and testing issues are discussed in Section 3, whereas some of the properties of the model are demonstrated in Section 4. The framework is applied in Section 5 to modeling the U.S. stock-market crash in 1987. Concluding comments are given in Section 6.

2. ENDOGENOUS SWITCHING MODEL

In this section, an aggregate model of asset markets is derived. Two main features of the model are: Prices adjust according to the gap between demand and supply of assets; and agents are heterogeneous, at least on the supply side, with one group utilizing portfolio-insurance hedging strategies.

It is assumed that the market price adjusts to the gap between the quantity of assets demanded, Q_t^D, and supplied, Q_t^S:

$$dP_t = -\gamma\left(Q_t^S - Q_t^D\right)dt + \sigma\, dW_t, \tag{1}$$

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

where ($\gamma > 0$) is a parameter controlling the speed with which the market returns to equilibrium after a shock, as given by the Wiener process, dW_t, and σ^2 is the instantaneous variance. The mean is simply

$$\mu = \gamma\left(Q_t^S - Q_t^D\right),$$

which shows that, on average, there is no price movement when there is market equilibrium.

As with standard microeconomic models of markets, the demand and the supply of assets are assumed to be functions of the asset price. For simplicity, agents on the demand side of the market are assumed to be homogeneous and adjust linearly and inversely to changes in price signals of the asset [see, e.g., Gennotte and Leland (1990)]:

$$Q_t^D = \phi_0 - \phi_1 P_t, \tag{2}$$

where $\phi_i > 0$, $i = 1, 2$.

On the supply side, agents are assumed to be heterogeneous, being divided into two groups, depending on whether they engage in hedging strategies based on portfolio insurance, $Q_{2,t}^S$, or not, $Q_{1,t}^S$:

$$Q_t^S = Q_{1,t}^S + Q_{2,t}^S. \tag{3}$$

The key feature of the group not engaged in portfolio insurance is that it responds positively to increases in price movements,

$$Q_{1,t}^S = \delta_0 + \delta_1 P_t, \tag{4}$$

where δ_i are positive parameters. This is in contrast to the portfolio insurance group which acts in a perverse way in the market by selling assets when prices are falling so as to hedge against further losses [Rubinstein and Leland (1981)]. To model this characteristic formally, assume that hedgers use put options to protect themselves from falling prices; see Gennotte and Leland (1990, p. 1009) who considered a portfolio insurance model based on a dynamic hedging strategy that replicates a put option. Letting K be the strike price of the option, the put-option contract is exercised when the actual price falls below the strike price, $P_t < K$. If the size of the portfolio is N, then the number of assets sold when the contract is exercised is

$$Q_{2,t}^S = h_t N, \tag{5}$$

where h_t is the optimal hedge ratio. Assuming that put options are priced using Black–Scholes model and that the option contract is for one year, equation (5) becomes [see Tucker (1991, p. 459)]

$$Q_{2,t}^S = \left[1 - \Phi\left(\frac{\log(P_t/K) + r - d_t + 0.5\nu^2}{\nu}\right)\exp(-d_t)\right]N, \tag{6}$$

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

where $\Phi(\cdot)$ is the cumulative normal distribution; r is the interest rate that, for simplicity, is assumed to be constant; d_t is the dividend yield; and ν^2 represents the volatility of returns.

The assumption that the contract is a one-year contract is made just for convenience without changing the qualitative properties of the model. The use of a put option based on Black–Scholes to replicate a hedging strategy to a certain extent is arbitrary: The key feature of the model is that supply increases with a decrease in price. Further, although the model could be expanded to allow for more than one strike price, the key point is that there exists a strike price such that the observed price of the asset falls below it, thereby sparking an increase in supply.

There are two key features of equation (6) that are worth highlighting. First, there is an inverse relationship between $Q^S_{2,t}$ and price P_t. That is, the supply schedule is downward sloping—when there is a large enough fall in price, as given by $P_t < K$, put-option contracts then are exercised. Second, the relationship between price and quantity is nonlinear, which can give rise to a nonlinear backward-bending aggregate supply schedule, causing the market to exhibit multiple equilibria for certain regions of P_t. To highlight this property, substitute (2)–(4) and (6) into (1). The resulting stochastic differential equation is nonlinear in price:

$$dP_t = -\gamma\Bigg(\delta_0 - \phi_0 + (\delta_1 + \phi_1)P_t + \left\{1 - \Phi\left[\frac{\log(P_t/K) + r - d_t + 0.5\nu^2}{\nu}\right]\exp(-d_t)\right\}N\Bigg)\,dt + \sigma\,dW_t. \quad (7)$$

It is convenient to re-express (7) in an alternative diffeomorphic form. To perform this task, note that the cumulative normal distribution can be approximated as follows:

$$\begin{aligned}\Phi(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\exp(-0.5s^2)\,ds \\ &= 0.5 + \int_{0}^{x} \frac{1}{\sqrt{2\pi}}\exp(-0.5s^2)\,ds \\ &\simeq 0.5 + \frac{1}{\sqrt{2\pi}}\left(x - \frac{x^3}{6} + \cdots\right).\end{aligned}$$

From the definition of $x = [\log(P/K) + r - d_t + 0.5\nu^2]/\nu$, this expression can be approximated using a Taylor-series expansion around the strike price K:

$$x \simeq \frac{P}{K\nu} + \frac{2(r - d_t) + \nu^2 - 2}{2\nu}.$$

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

Substituting into the above for x gives the following diffeomorphic form which is a cubic in P_t:

$$dP_t = -\sum_{i=0}^{3} \alpha_i P_t^i \, dt + \sigma \, dW_t, \tag{8}$$

where the α_i's are functions of $\{\pi, \delta_i, \phi_i, \nu, K, N, r, d_t\}$. The stochastic differential equation (8) is qualitatively the same as (7) in that both processes exhibit the same number of equilibria. Of course, these approximations are based on the assumption that movements in P_t are small, which may not be appropriate at the time of a stock-market crash in which there is a large movement in the price of the asset. This is also a problem for constructing portfolio insurance based on dynamic put-option replicating strategies; see Tucker (1991, p. 461). For this case, although it may be more appropriate to use an asymptotic expansion to obtain a more accurate approximation, the qualitative properties of the equilibrium manifold still can be captured adequately by a cubic.

Associated with (8) is the distribution of prices $f(p_t)$ at each point in time. This is derived by solving the Kolmogorov forward equation [see Cox and Miller (1984)],

$$\frac{\partial f}{\partial t} = \frac{\partial(\mu f)}{\partial p} + \frac{1}{2}\frac{\partial^2(\sigma^2 f)}{\partial p^2}, \tag{9}$$

where, from (8), the instantaneous mean is $\mu = \sum_{i=0}^{3} \alpha_i P_t^i$. It is also possible to derive the long-run distribution by setting $\partial f/\partial t = 0$ in (9), which yields

$$f^*(p) = \exp\left[-\int_0^p \left(\frac{2\mu(s) + d\sigma^2(s)/ds}{\sigma^2(s)} \, ds - \eta^* \right) \right], \tag{10}$$

where η^* is the normalizing constant to ensure that $\int f^*(p)\, dp = 1$. This density represents a generalization of the exponential family of distributions, which has been studied by Cobb et al. (1983) and Lye and Martin (1993). Substituting in the expression $\mu = \sum_{i=0}^{3} \alpha_i P_t^i$ gives

$$f^*(p) = \exp\left[\varphi_1 p + \varphi_2 p^2 + \varphi_3 p^3 + \varphi_4 p^4 - \eta^*\right], \tag{11}$$

where the normalizing constant is $\eta^* = \log \int \exp[\varphi_1 p + \varphi_2 p^2 + \varphi_3 p^3 + \varphi_4 p^4]\, dp$, and the φ_j's are functions of $\{\alpha_0, \alpha_1, \alpha_2, \alpha_3, \sigma^2\}$.

A property of the model given by (8) is that, as the cubic equation in price can give rise to three roots at most, there can be three equilibria at most—two stable and one unstable. The occurrence of multiple equilibria is translated into the distribution of prices in (11) exhibiting multimodality. In particular, this distribution can have up to two modes and one antimode, where the modes correspond to the stable equilibria and the unstable mode corresponds to the unstable equilibrium.

An important feature of this distribution for modeling a crash is that it is possible to endogenously model large jumps in the asset price. In addition, the occurrence of a jump can occur with only a small change in the market fundamentals. This

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

property is highlighted in Section 4 using some Monte Carlo experiments. This distribution also can exhibit other properties, such as skewness and kurtosis, that are characteristic of financial data. The distribution is referred to as a generalized normal distribution because it contains as a special case the unimodal distribution when $\varphi_3 = \varphi_4 = 0$.

3. ESTIMATION AND TESTING ISSUES

Maximum likelihood methods can be used to estimate the parameters of the generalized distributions; see Cobb et al. (1983) and Lye and Martin (1993). The regularity conditions needed to ensure that the maximum likelihood estimator is best and asymptotically normal are immediately satisfied because the family of generalized distributions belongs to the exponential family.

Before setting out the maximum likelihood estimator, it is convenient to reparameterize the distribution in (11) by defining [see Creedy et al. (1996)]

$$u_t = (P_t - \Lambda_t)/\upsilon, \tag{12}$$

where Λ_t and υ are implicitly defined below. This yields the following distribution for u_t:

$$f(u_t) = \exp\left[\theta_1 u_t + \theta_2\left(u_t^2/2\right) - u_t^4/4 - \eta_t\right], \tag{13}$$

where, now, the normalizing constant is

$$\eta_t = \log \int \exp\left[\theta_1 u + \theta_2(u^2/2) - u^4/4\right] du, \tag{14}$$

and

$$\begin{aligned}
\varphi_1 &= \Lambda_t^3/\upsilon^4 - 2\Lambda_t\left(\theta_2/\upsilon^2\right) + \theta_1/\upsilon,\\
\varphi_2 &= -3\Lambda_t^2/2\upsilon^4 + \theta_2/\upsilon^2,\\
\varphi_3 &= \Lambda_t/\upsilon^4,\\
\varphi_4 &= -1/4\upsilon^4.
\end{aligned}$$

A natural interpretation of Λ_t, and hence u_t in (12), is obtained by taking expectations conditional on information at time t of both sides:

$$E_t[u_t] = E_t\left[\frac{p_t - \Lambda_t}{\upsilon}\right] = 0, \tag{15}$$

where the property that $E_t[u_t] = 0$ immediately follows by noting that (12) represents a transformation to standardize the distribution of u_t to have zero mean. This expression shows that Λ_t is the expected price and that u_t is the associated error

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

term and υ is the associated standard deviation. Using a standard present-values argument, the expected price is given by [see, e.g., Barsky and De Long (1993)]

$$\Lambda_t = \lambda_0 + \lambda_1 D_t, \tag{16}$$

where D_t is the dividend payment. For this class of models, the parameter λ_1 equals the inverse of the rate of interest, whereas $\lambda_0 = 0$. For the model investigated by Barsky and De Long (1993), λ_1 is the inverse of the rate of interest adjusted by a permanent dividend growth factor and $\lambda_0 = 0$.

For a sample of size $t = 1, 2, \ldots, T$, the log likelihood in terms of P_t is

$$\begin{aligned}\log L &= \sum_{t=1}^{T}\left[\log f\left(\frac{P_t - \Lambda_t}{\upsilon}\right) - \log \upsilon\right] \\ &= \sum_{t=1}^{T}\left[\theta_1\left(\frac{P_t - \Lambda_t}{\upsilon}\right) + \frac{\theta_2}{2}\left(\frac{P_t - \Lambda_t}{\upsilon}\right)^2\right. \\ &\quad \left. - \frac{1}{4}\left(\frac{P_t - \Lambda_t}{\upsilon}\right)^4 - \log \upsilon\right] - \sum_{t=1}^{T}\eta_t. \end{aligned} \tag{17}$$

The first-order derivatives with respect to θ_j are

$$\frac{\partial \log L}{\partial \theta_j} = \sum_{t=1}^{T}\left[\left(\frac{P_t - \Lambda_t}{\upsilon}\right)^j - \frac{\partial \eta_t}{\partial \theta_j}\right], \qquad j = 1, 2.$$

From the definition of η_t in (14),

$$\begin{aligned}\frac{\partial \log L}{\partial \theta_j} &= \sum_{t=1}^{T}\left\{\left(\frac{P_t - \Lambda_t}{\upsilon}\right)^j - \frac{\int u^j \exp\left[\theta_1 u + \theta_2(u^2/2) - u^4/4\right] du}{\int \exp\left[\theta_1 u + \theta_2(u^2/2) - u^4/4\right] du}\right\} \\ &= \sum_{t=1}^{T}\left\{\left(\frac{P_t - \Lambda_t}{\upsilon}\right)^j - E\left[\left(\frac{P_t - \Lambda_t}{\upsilon}\right)^j\right]\right\}, \qquad j = 1, 2.\end{aligned}$$

Evaluating at zero shows that the maximum likelihood estimator is found by equating the sample moments with the theoretical moments. In general, the log likelihood can be maximized using conventional iterative optimization algorithms. The information matrix can be computed, which then can be used to perform hypothesis tests. It is, however, necessary to use a numerical integration routine to compute the normalizing constant in (14) when maximizing the likelihood because this expression generally does not lend itself to a closed-form solution. All routines are written in GAUSS using the procedures MAXLIK to maximize (17) and INTQUAD1 to perform the numerical integration. For convenience, numerical derivatives are used to compute the gradient vector and the information matrix used in the optimizing algorithm.

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

3.1. Prediction Conventions

For models based on unimodal distributions, the prediction convention usually is chosen on the basis of the mean of the distribution. When the distribution is multimodal, the choice of the mean of the distribution is no longer appropriate. An extreme example is given by a distribution that is both bimodal and symmetrical. Here, the mean of the distribution is located at the antimode—a value corresponding to a region with the smallest probability mass. Thus, it is necessary to consider an alternative criterion. One possibility, known as the global convention, is to choose the value corresponding to the global mode. Another possibility, known as the delay convention, is to choose the mode that is closest to the realized value of the dependent variable at lag τ. This latter convention is appropriate for modeling processes exhibiting inertia. Although the choice of conventions is, to a certain extent, arbitrary, the choice can be based on empirical grounds such as goodness of fit.

3.2. Identification of Bimodality

As an aid for identifying the properties of the distribution over time, Cardan's discriminant can be used when the number of modes of the distribution is, at most, two. For the distribution given by (13), the number of modes is found by solving $df(u_t)/du = 0$. That is, the number of modes corresponds to the roots of the cubic

$$\theta_{1,t} + \theta_{2,t} u_t - u_t^3 = 0. \tag{18}$$

For this equation, Cardan's discriminant is

$$\delta_t = (\theta_{1,t}/2)^2 - (\theta_{2,t}/3)^3. \tag{19}$$

For positive (negative) values of δ_t, the distribution is unimodal (bimodal). Cardan's disciminant is used here simply as a descriptive tool for identifying the shape of the price distribution. The asymptotic distribution of δ_t may be calculated from the distribution of the θ's using the delta method.

4. SIMULATING CRASHES

To highlight the jumping property of the model, the following simulation experiment is conducted. Define the deviation between the stock price, P_t, and the market fundamentals, as given by dividends D_t, as

$$u_t = P_t - D_t. \tag{20}$$

The error distribution is specified as

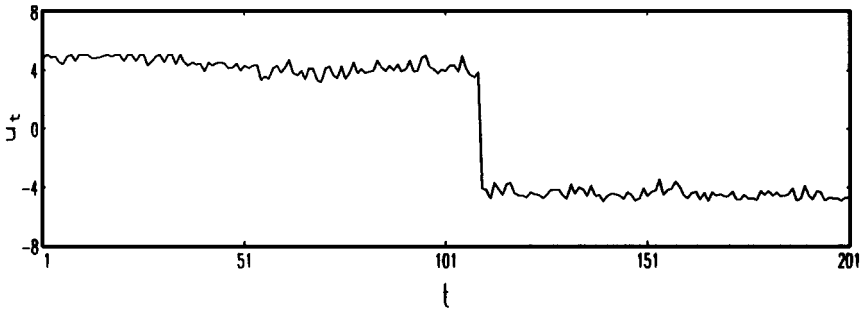
$$f(u_t) = \exp\left[10 d_t u_t + u_t^2 - 0.005 u_t^4 - \eta_t^s\right], \tag{21}$$

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

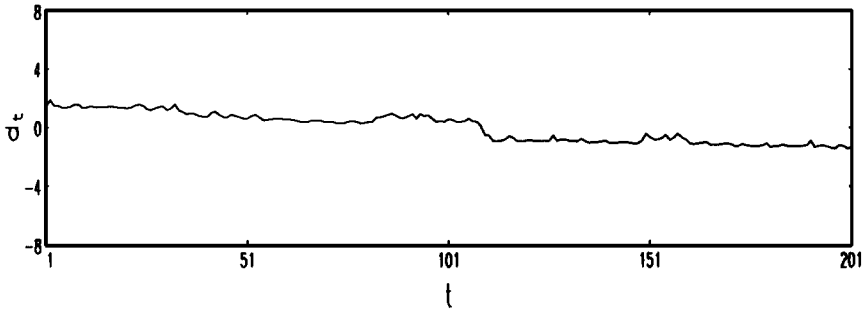
where the normalizing constant is

$$\eta_t^s = \log \int \exp\left[10 d_t s + s^2 - 0.005 s^4\right] ds, \tag{22}$$

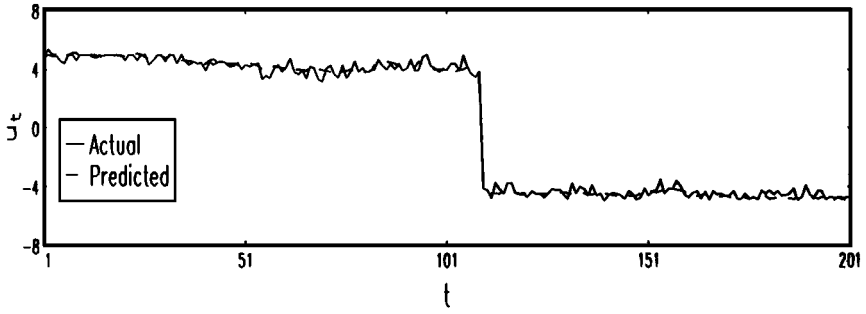
and d_t is the dividend yield. Comparing the coefficient on u_t^4 in (13) and (21) shows that $\upsilon = (4 \times 0.005)^{-0.25}$. Linear and constant relationships between dividend yields and those implied by the model in (13) are specified to highlight how jumping behavior can arise for even simple specifications. The dividend yield is plotted in Figure 1B. These data consist of $T = 201$ observations, which is a



A) Simulated price (deviation) data



B) Dividend data



C) Actual and predicted

FIGURE 1. Simulation experiment: Estimation is based on total sample $t = 1, 2, \ldots, 201$.

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

subsample of the U.S. dividend yield series used in the empirical section below and are standardized to have zero mean and unit variance. The important characteristic to note is that this series is smoothly evolving over time and does not exhibit any large movements.

To simulate u_t, the inverse cumulative density technique is used. This amounts to constructing the cumulative density from (21) for a given d_t. By drawing a uniform random number, the associated random number for u_t is, in turn, also drawn. This experiment is repeated for each value of d_t given in Figure 1B.

The simulated series for u_t is given in Figure 1A. In contrast with the dividend yield series in Figure 1B, the simulated u_t process does exhibit a relatively large fall at observation $T = 110$. From the definition of u_t in (20), this fall reflects that, for a smoothly evolving dividend stream, the market price P_t significantly deviates from the market fundamentals D_t.

The results of estimating the parameters of the model

$$f(u_t) = \exp\left[\alpha r_t u_t + \beta u_t^2 + \phi u_t^4 - \eta_t^s\right] \tag{23}$$

by maximum likelihood for the total period ($T = 201$) are given in Table 1. Within-sample predictions of the model using the global convention are given in Figure 1C. It is clear that the model captures not only the magnitude of the large fall in u_t and hence in the share price, but it also correctly captures the timing of the fall.

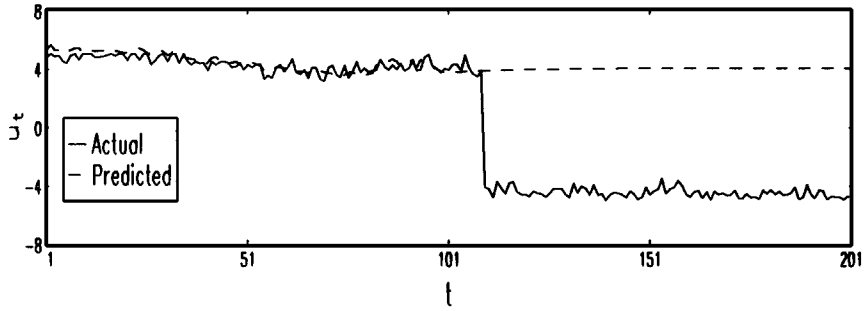
The good predictive properties of the model demonstrated in Figure 1C could be conditional on the inclusion of the fall in the price during the estimation period. What is also of interest is whether a fall in the price can be predicted from the model when no fall exists during the estimation period. To investigate this property,

TABLE 1. Simulation experiment: Maximum likelihood parameter estimates of equation (23)

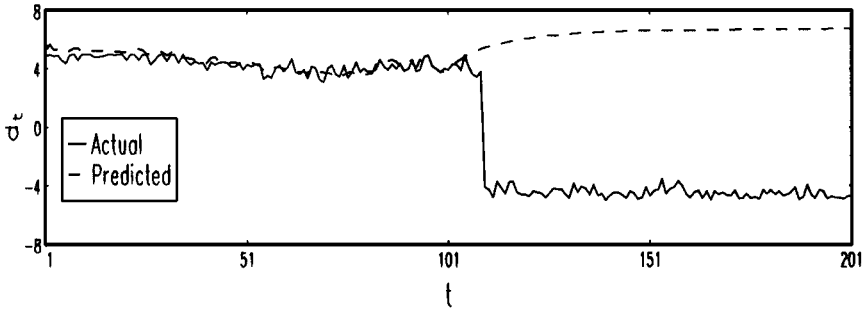
Parameter	True	Estimate[a]
	A. $T = 201$ (*Total sample*)	
α	10.000	9.920
		(8.675)
β	1.000	0.930
		(4.882)
ϕ	−0.005	−0.048
		(7.147)
	B. $T = 100$ (*Subsample*)	
α	10.000	15.034
		(7.310)
β	1.000	0.479
		(1.966)
ϕ	−0.005	−0.046
		(4.443)

[a] Absolute t-stats are in parentheses.

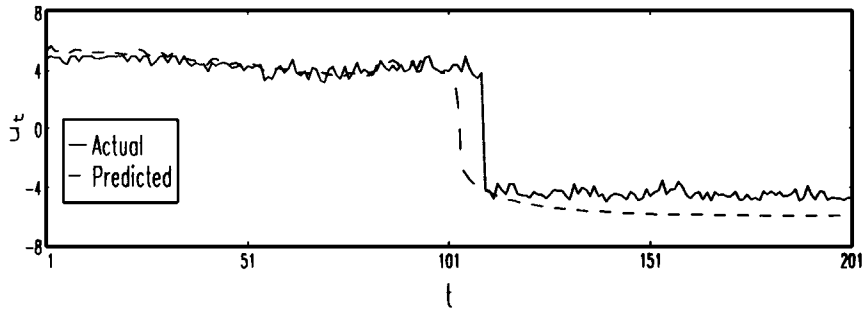
https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press



A) Actual and predicted (zero shock)



B) Actual and predicted (positive shock)



C) Actual and predicted (negative shock)

FIGURE 2. Simulation experiment: Estimation is based on the subsample $t = 1, 2, \ldots, 100$.

the model is re-estimated over the first $T = 100$ observations, which is a period of tranquility. This subperiod is terminated 10 days before the actual crash eventuates. The parameter estimates for the subperiod are given in Table 1. The predictive properties of the estimated model are given in Figure 2. In generating forecasts of u_t, forecasts of d_t are generated using an AR(3) model with a constant. The results are given in Figure 2A and show that there is no evidence of a jump. However, to see if the underlying process is close to a region of bimodality and hence there

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

is a nonzero probability of jumping, the forecasts of d_t are changed by adding (subtracting) a permanent one-standard-deviation (SD) shock. The size of the shock is equal to the standard deviation of the residuals from the estimated AR(3) model. Figure 2B shows that the positive 1-SD shock yields a smooth predicted time path of u_t. This is in contrast to the negative shock that results in a large fall in u_t, which is similar to the magnitude that eventually takes place.

This example highlights the importance of conducting sensitive analysis in identifying potential large jumps in asset prices. If, for a broad range of shocks, no jumps are predicted, the relationship between the asset price and fundamentals can be expected to be smooth. If, on the other hand, jumps are detected for certain shocks, then there is a nonzero probability of predicting a jump.

5. THE 1987 STOCK-MARKET CRASH

One of the more difficult challenges in constructing models of stock markets is to explain the simultaneous occurrence of a sharp fall in the stock price, as is characteristic of the 1987 stock-market crash, with a smoothly evolving dividend process. This phenomenon is demonstrated in Figure 3, which gives for the period January 2, 1986, to December 30, 1988, the daily U.S. stock index as well as the dividend payments and yields.[1] Although both dividend series move relatively smoothly over the period, the stock index exhibits a massive fall on October 19, 1987. On this day the value of the price index fell from 2.48 trillion to 2.03 trillion, a 22% fall, whereas for dividends, the fall was negligible.

In this section, the nonlinear theoretical framework discussed above is applied to modeling the U.S. stock-market crash of October 1987. The total number of observations is $T = 757$. The data are standardized to have zero mean and unit variance.

5.1. Model Specification

An initial empirical specification of the nonlinear model that is used to capture the large movement in the stock index identified in Figure 1A is given by

$$P_t = \lambda_1 D_t + \upsilon u_t, \tag{24}$$

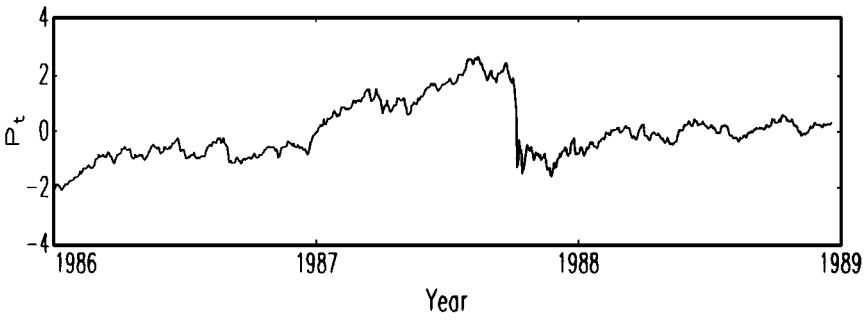
$$f(u_t) = \exp\left[\theta_{1,t} u_t + \theta_{2,t}\left(\frac{u_t^2}{2}\right) - u_t^4/4 - \eta_t\right], \tag{25}$$

$$\theta_{1,t} = \pi_1 d_t, \tag{26}$$

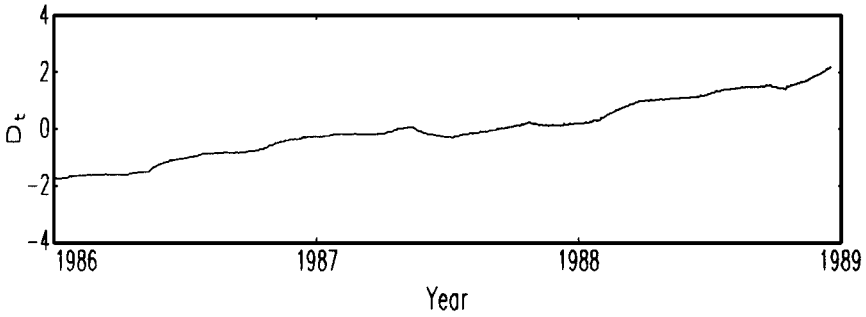
$$\theta_{2,t} = \pi_2 d_t. \tag{27}$$

Equation (24) shows that the mean of the stock price is a function of the market fundamental variable, dividend payments D_t, with parameter λ_1. This constitutes the present value formulation, which represents a direct link between prices and dividend payments. The error term is given by u_t, with corresponding standard deviation υ, which for simplicity is assumed to be a constant. The distribution of

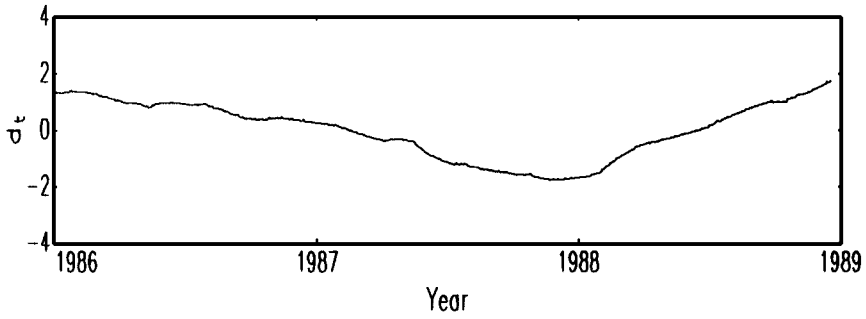
https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press



A) CRSP Value-weighted index (standardized)



B) Dividend payments (standardized)



C) Dividend yields (standardized)

FIGURE 3. U.S. stock market data (standardized).

u_t has the generalized normal form given by (25), where η_t is the normalizing constant. The shape of this distribution over time is determined by $\theta_{1,t}$ and $\theta_{2,t}$ in equations (26) and (27), which are expressed as linear functions of the dividend yield d_t. This represents an indirect link between the stock price and market fundamentals as d_t affects P_t through the distribution of u_t. Moreover, it is this link that creates the nonlinear relationship between P_t and the market fundamental variables.

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

Table 2. U.S. stock market total sample period: Maximum likelihood parameter estimates of equations (24)–(27)

Variable	Parameter	Linear[a]	Nonlinear[a]
D_t	λ_1	0.322	0.419
		(9.321)	(15.232)
d_t	π_1		−0.281
			(7.154)
d_t	π_2		−1.496
			(17.961)
	υ	0.946	1.081
			(59.532)

[a] Absolute t-statistics are in parentheses.

As with the simulation experiment above, linear functional forms are used in (26) and (27) to show the ability of the model to detect jumping behavior. Further, the model is summarized by four parameters. This would be the same number if a dummy variable model with both shift and slope dummies was used to capture the jump observed in Figure 3. Of course, the main difference is that the dummy variable model can be used only to determine the timing of the jump ex post and cannot be used to forecast the jump. In contrast, the model specified by (24) and (27) can capture not only the timing and magnitude of jumps, but potentially it can be used to forecast jumps, even when a jump has not occurred during the sample period.

5.2. Results

Maximum likelihood parameter estimates of the nonlinear model are given in Table 2 for the full sample period.[2] For comparison, ordinary least squares (OLS) estimates of the linear model (24) also are given.[3] The results show that there are statistically significant linkages between P_t and D_t (the direct link) and between P_t and d_t (the indirect link). This contrasts with the linear results in which only dividend payments have a significant effect on the stock price index. Finding evidence of a significant relationship between d_t and P_t is important because it constitutes a necessary first step for finding an endogenous mechanism for modeling the large movement in the stock price index over the sample period.

A test of the properties of the linear and nonlinear models is given in Figure 4, which presents the within-sample predictions of the two models for the total period. The delay convention has been used in the case of the nonlinear model. The most striking observation to make about the predictions of the nonlinear model is that it is able to capture arbitrarily well both the timing of the crash as well as the magnitude of its fall. This contrasts with the linear present value model which fails to capture the crash at all.

The nonlinear model performs so well because it predicts that the distribution of the stock price is bimodal on the days leading up to the crash. This is shown by the negative values of Cardan's discriminant, δ_t, from (19), in Figure 5A. Confidence

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

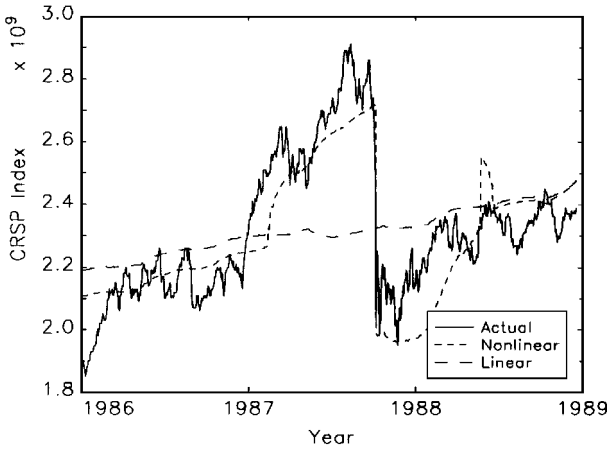
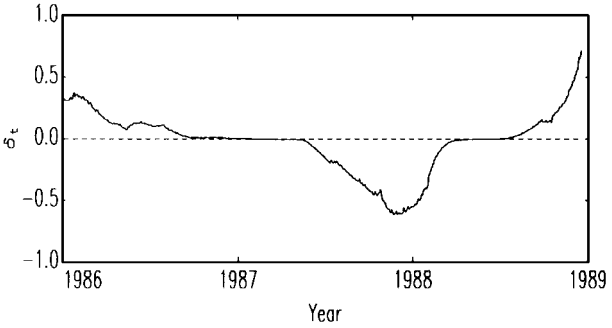
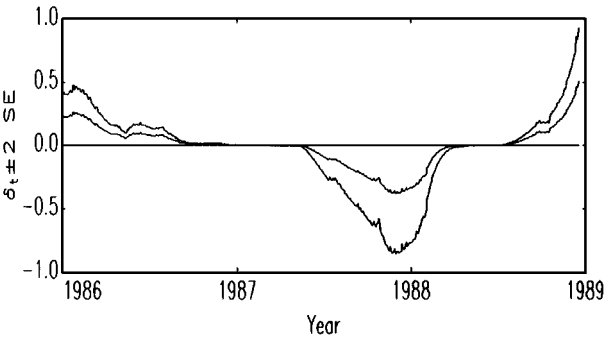


FIGURE 4. Linear and nonlinear predictions using the total sample period.



A) Measure of bimodality: δ_t



B) Confidence interval of bimodality: $\delta_t \pm 2$ SE

FIGURE 5. Measures of bimodality: δ_t.

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

intervals of the delta function are plotted in Figure 5B for the sample period. For the above-estimated model, the standard error at each point in time is computed as the square root of

$$\left(0 \quad 0 \quad \frac{\hat{\pi}_1 d_t^2}{2} \quad -\frac{\hat{\pi}_2^2 d_t^3}{9}\right)\hat{H}^{-1}\left(0 \quad 0 \quad \frac{\hat{\pi}_1 d_t^2}{2} \quad -\frac{\hat{\pi}_2^2 d_t^3}{9}\right)', \tag{28}$$

where $\hat{H}^{-1}$ is the covariance matrix of the unconstrained model evaluated at the MLE parameter estimates reported in Table 2. Note that the measure of bimodality is statistically significant around the time of the crash because both confidence interval bands are less than zero.

The bimodal properties of the stock index distribution are further highlighted in Figure 6, which gives temporal snapshots of the price distribution for the weeks before, during, and after the crash. These results provide strong support for the dual-view hypothesis that there was more than one view of the equilibrium stock price at around the time of the stock-market crash.

The negative values of δ_t in Figure 5 suggest that this dual view began approximately in the last week of April 1987, and ended approximately one year later when the market was once again united about its view of the share price. This property of the price distribution is illustrated further in Figure 7, which combines the weekly snapshots of the estimated distribution for the total sample period.

5.3. Forecasting

To investigate its forecasting properties, the model is estimated over three subsamples. The three subperiods all begin January 2, 1986, and end 12, 7, and 2 trading days prior to the crash. For each subperiod, a one-step-ahead forecast of the distribution is computed. If the distribution is bimodal, the probability that the area to the left of the antimode, that is, the area around the lower mode, is calculated. In computing the one-step-ahead forecast density, information on future dividends is obtained from the forecasts of 10th-order autoregressions for dividend payments and yields estimated over the same subperiods as the price distribution.

Figure 8 shows that the estimate of δ_t with the sample period ending 12 days before the crash is significantly negative before the crash, thereby suggesting that there is significant bimodality before the crash. Similar results are obtained for the other two subperiods. These results are significant in that they show that the bimodality results obtained for the total period are not completely conditioned on the large fall in the stock price on October 19, 1987. In Table 3, the MLE estimates are given for the subsample that ends 12 days before the crash.

For the subperiod corresponding to 12 trading days prior to the crash, the one-step-ahead prediction of the upper mode is 2.85 trillion, which is the same as the recorded value of 2.85 trillion. What is more significant is that the one-step-ahead prediction of the lower mode is 1.88 trillion, which is close to the share price index of 2.03 trillion when it actually crashed.

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

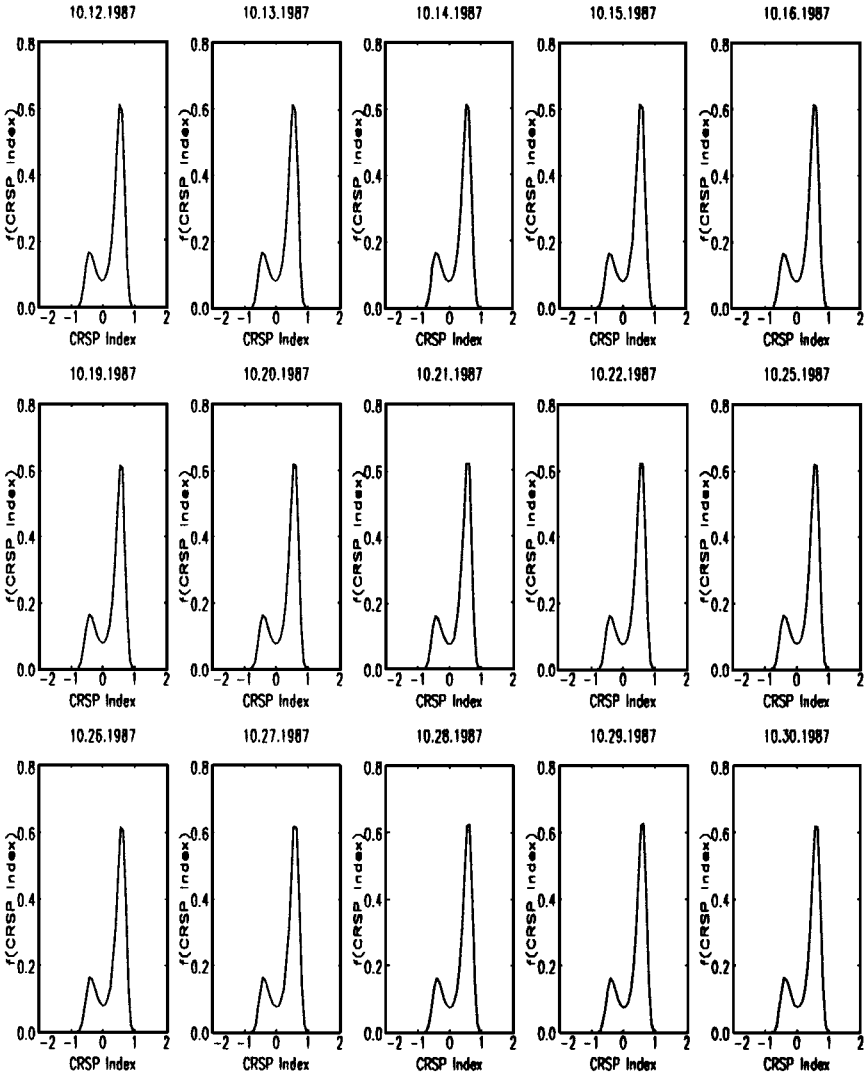


FIGURE 6. Snapshots of estimated distributions from October 12 to 30, 1987 (standardized).

To gauge the probability mass around the lower mode, and hence the likelihood of a crash occurring, a sensitivity analysis is performed on the one-step-ahead forecast distribution. The approach is to shock the parameters, π_1 and π_2 in (26) and (27), respectively. The results of shocking π_1 and π_2 by $\pm 3\hat{\sigma}(\pi_1)$ and $\pm 3\hat{\sigma}(\pi_2)$ respectively, where $\hat{\sigma}(\pi_i)$, $i = 1, 2$, are the associated MLE's of the standard deviations, are shown in Figure 9 for the subsample ending 12 days prior to the actual crash. For comparison the forecast distribution without a shock also is shown.

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

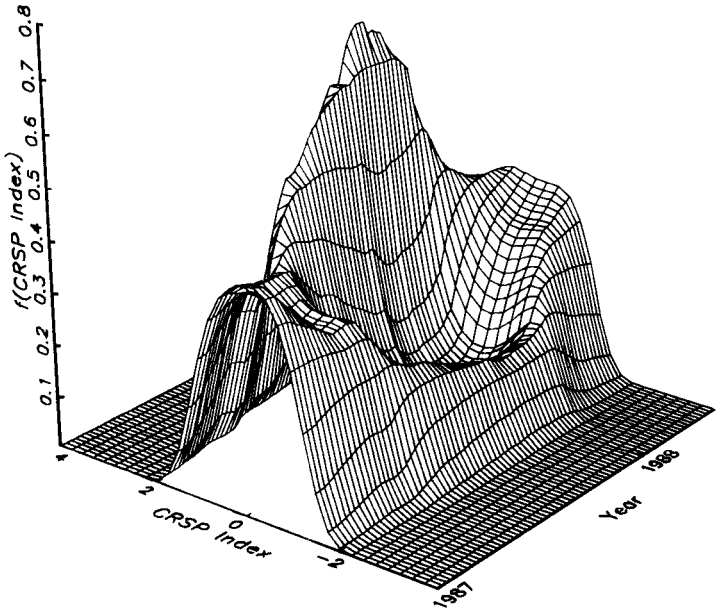


FIGURE 7. Weekly snapshots from January 1987 to May 1988 (standardized).

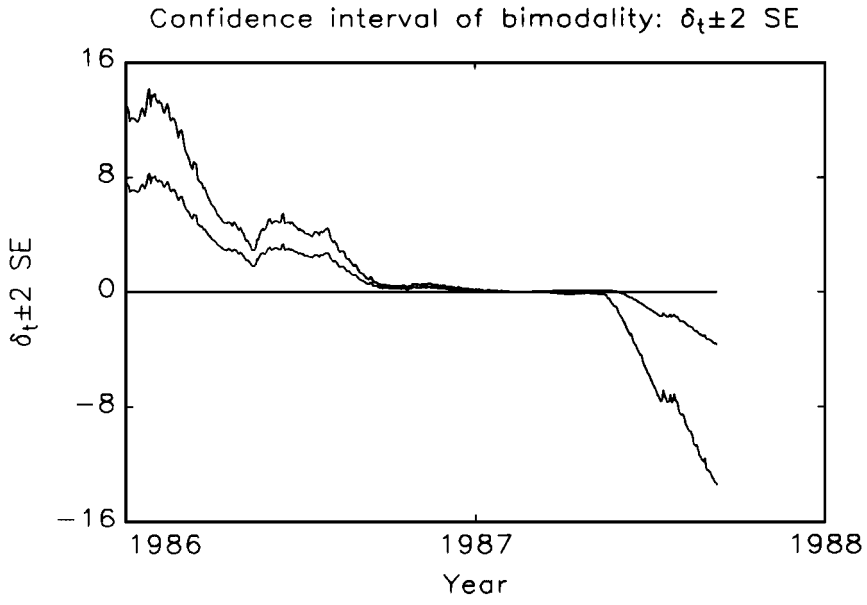


FIGURE 8. Measure of bimodality computed 12 days before the crash.

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

TABLE 3. U.S. stock market 12 days before the crash: Maximum likelihood parameter estimates of equations (24)–(27)

Variable	Parameter	Linear[a]	Nonlinear[a]
D_t	λ_1	0.629	0.519
		(10.961)	(8.075)
d_t	π_1		−2.023
			(4.225)
d_t	π_2		−4.530
			(11.515)
	υ	1.069	0.835
			(45.882)

[a] Absolute t-statistics are in parentheses.

Inspection of Figure 9 shows that, for a three-standardization negative shock, the probability around the lower mode is 0.369. The lower modal value is 2.085 trillion, which again compares favorably with the value of the stock price index on the day of the crash.

5.4. Alternative Specification

In specifying the model in (24) and (27), linear expressions for θ_{jt} were adopted so as to highlight the nonlinearity properties arising from the underlying distribution. However, from the theoretical model derived in Section 2, the relationships between θ_{jt} and d_t generally are highly nonlinear. To capture some of these nonlinearities, the model is extended by generalizing (26) and (27) to[4]

$$\theta_{1,t} = \pi_1 d_t + \pi_2 e_t + \pi_3 e_t^2, \tag{29}$$

$$\theta_{2,t} = \pi_4 d_t + \pi_5 e_t + \pi_6 e_t^2, \tag{30}$$

where e_t are the residuals from a regression of dividend yields on a constant and a linear time trend. The results of estimating (24), (25) and (29), (30) by MLE for the full sample are given in Table 4. Figure 10 shows the snapshots of the estimated daily densities for the weeks before, during, and after the crash. Predictions of the extended nonlinear model as well as the linear model are given in Figure 11, with the combined weekly snapshots of the price distribution for the total period given in Figure 12. Just as in the initial specification, the extended model is able to capture the jump. Both he timing and the magnitude of the estimated jump are close to the actual price. The extended specification represents an improvement over the initial model specification because the probability of a crash, as measured by the area around the lower mode, is now increased.

The one-step-ahead forecasting performance of the extended model is investigated over the same three subsamples, namely, estimation periods ending 12, 7, and 2 trading days prior to the crash. The forecast distribution is bimodal for

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

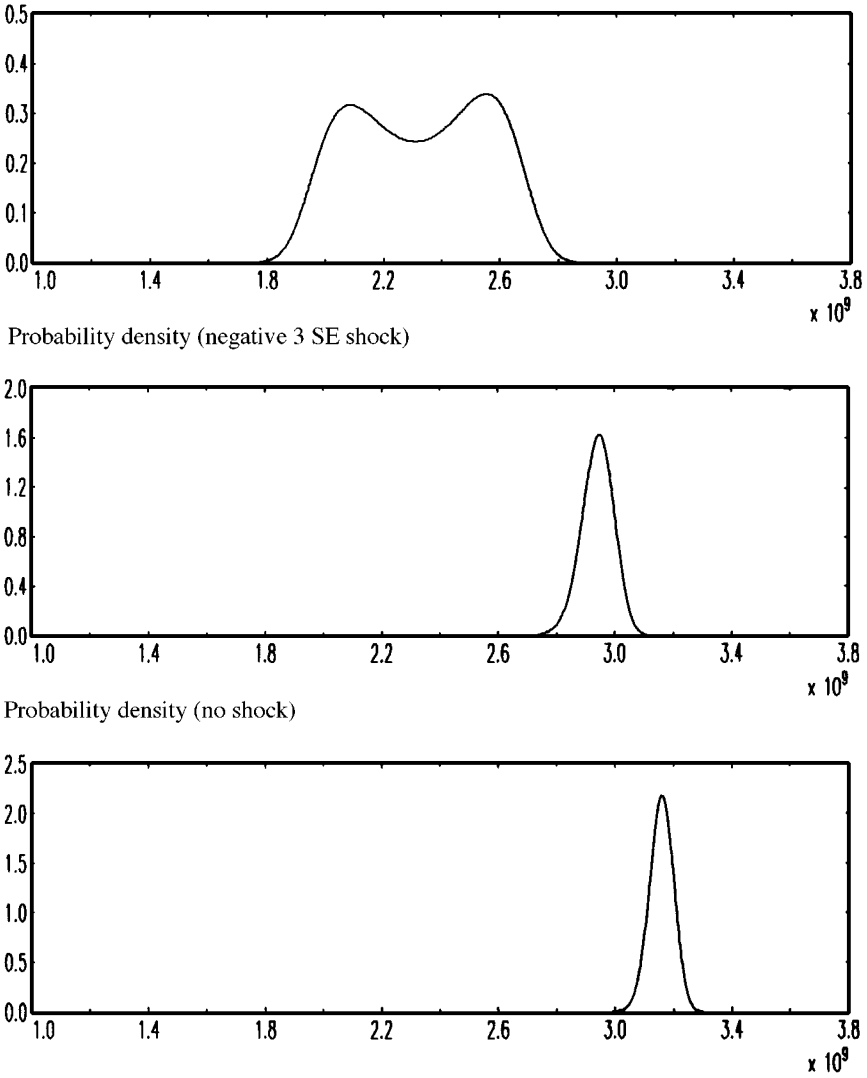


FIGURE 9. One-step-ahead prediction of price distributions for alternative shocks 12 days before crash (standardized).

all three subperiods. For the one-step-ahead forecast for the first subperiod the prediction of the higher mode is 2.83 trillion, which compares favorably with the actual value of 2.85 trillion. The prediction of the corresponding lower mode is 1.93 trillion, which is of a magnitude similar to the actual value of 2.03 trillion at the time of the crash in 11 days' time.

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

TABLE 4. U.S. stock market alternative model specification: Maximum likelihood parameter estimates of equations (24)–(27)

Variable	Parameter	Linear[a]	Nonlinear[a]
D_t	λ_1	0.629	0.539
		(10.961)	(14.556)
d_t	π_1		b4.808
			(7.008)
e_t	π_2		−8.442
			(9.398)
e_t^2	π_3		−2.193
			(13.054)
d_t	π_4		3.667
			(3.860)
e_t	π_5		−9.378
			(8.371)
e_t^2	π_6		−2.469
			(12.394)
	υ	1.069	0.9373
			(64.051)

[a] Absolute t-statistics are in parentheses.

6. CONCLUSIONS

This paper has provided a new approach for predicting the timing and the magnitude of stock-market crashes. An important advantage of this approach is that the timing of the crash is determined endogenously, thereby overcoming problems associated with the use of dummy variables to explain the crashes because these variables are determined *ex post* and hence cannot be used for forecasting. The model also is shown to be able to predict sudden changes in asset prices even when the estimation period coincides with a period of tranquility.

The approach was applied to modeling the U.S. stock-market crash of October 1987 using daily data on share prices and dividends. The main result of the empirical analysis was that the distribution of the stock price was bimodal around the time of the crash, reflecting two views of the market equilibrium price, and that the large fall in the stock price was the result of the price jumping between equilibria. This period of bimodality was estimated to exist for about a year beginning around April 1987.

NOTES

1. The stock price index is the Center for Research in Security Prices (CRSP) value-weighted portfolio. The dividend yield is the difference between the daily return of the value-weighted CRSP portfolio with and without dividends. This figure is aggregated over the past 253 days to yield an annual figure. Dividend payments are calculated by multiplying the daily dividend yield by the total value of

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

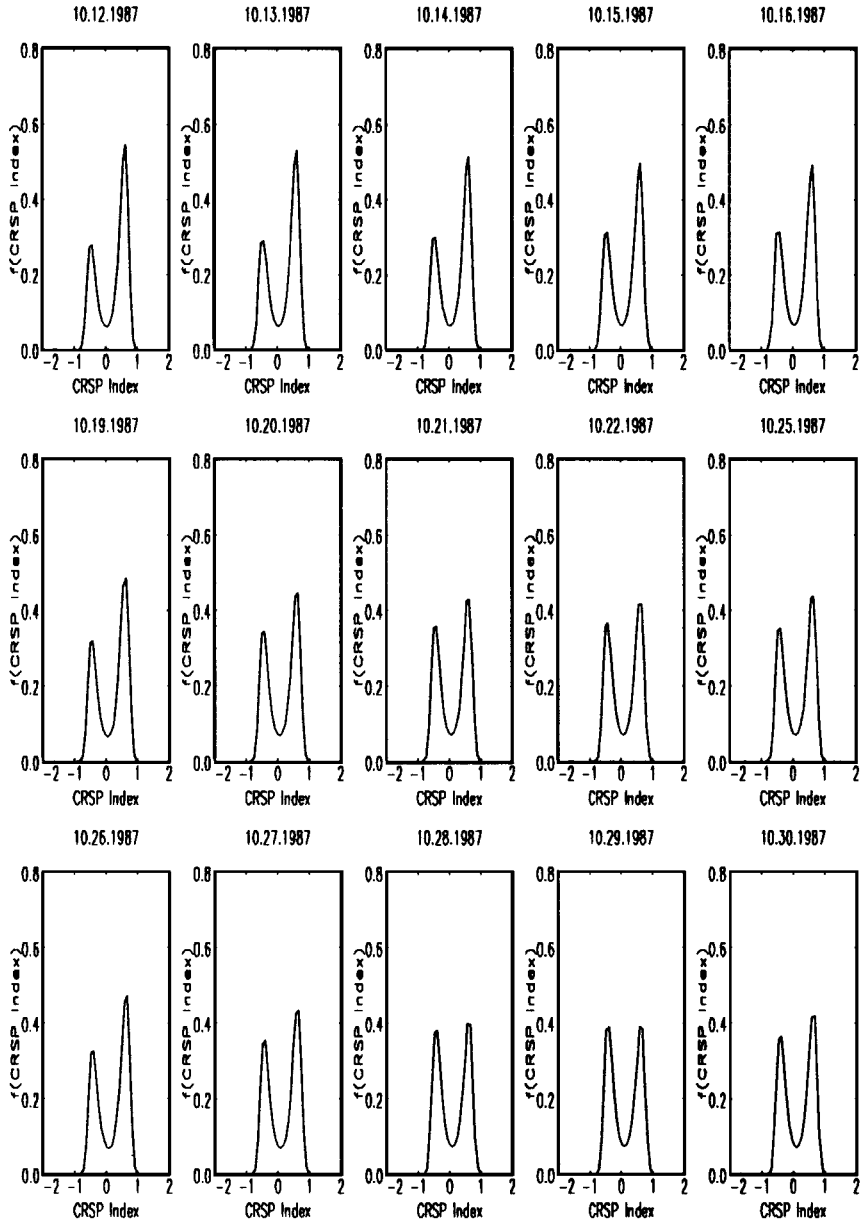


FIGURE 10. Alternative specification: Snapshots of estimated distributions from October 12 to 30, 1987 (standardized).

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

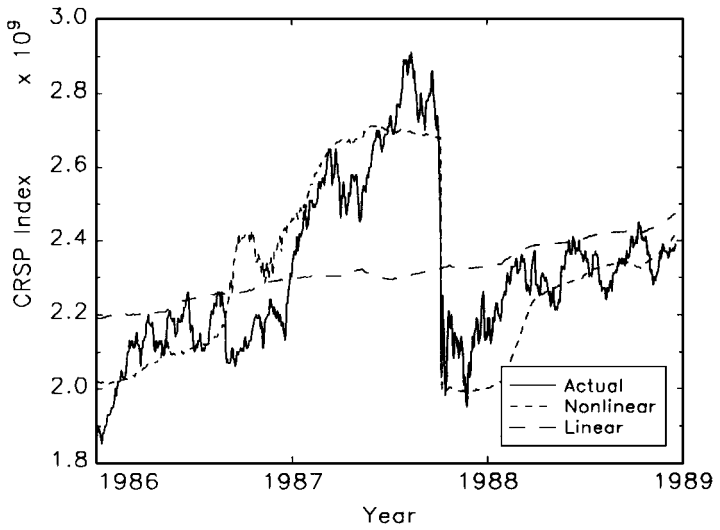


FIGURE 11. Alternative-specification predictions.

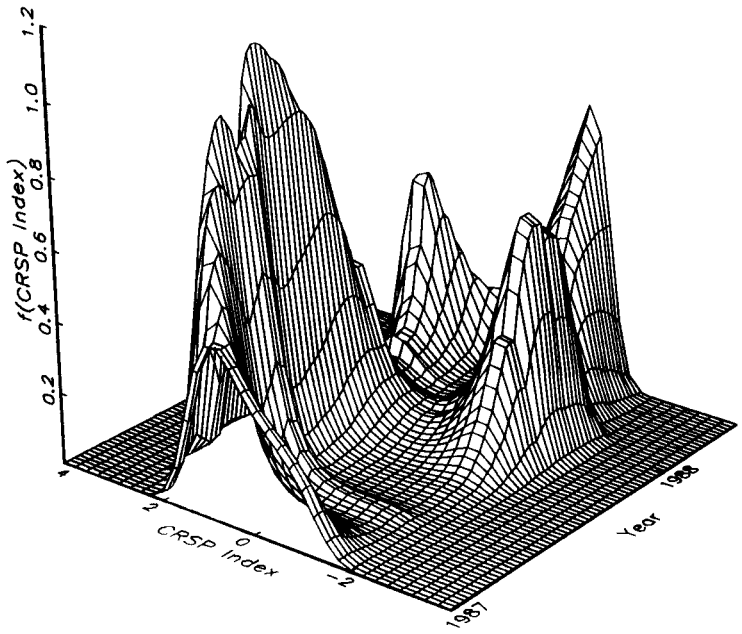


FIGURE 12. Weekly snapshots for alternative specification.

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

the portfolio, lagged one period. An annual figure then is calculated by aggregating the daily values over the past 253 trading days. This annual series is smoothed by a moving average of the past seven trading days. This smooths out spurious random fluctuations in the data arising from daily observations and generates smoothed market fundamentals which is important for the empirical analysis below. However, the results are still similar when data that are not smoothed are used.

2. The data were initially tested for unit roots using both the ADF and Phillips Z_t unit root tests. The results, which are not reported to save space, show that the null hypothesis of a unit root cannot be rejected at the 5% significance level for all three series. Co-integration tests, also not reported, based on the ADF and the Phillips Z_t tests show that both P_t and D_t are not co-integrated. Given that the length of the data is only three years, one has to be careful when interpreting the results of these tests. See also Craine (1993) for a discussion of the interpretation of unit root tests of these series.

3. The Froot and Obstfeld (1991) intrinsic bubble model also was estimated, but no significant nonlinearity was found.

4. Another alternative model specification that was tried was to allow the expression for θ_{jt} to be functions of both dividend yields and daily changes in dividend payments. However, the changes in dividend payments were not statistically significant, and more important, were found to have no influence on the predictive power of the model.

REFERENCES

Ahn, C.M. & H.E. Thompson (1988) Jump-diffusion processes and term structure of interest rates. *Journal of Finance* 43, 155–174.

Akgiray, V. & G.G. Booth (1988) Mixed diffusion—jump process modeling of exchange rate movements. *Review of Economics and Statistics* 70, 631–637.

Barsky, R.B. & J.B. De Long (1993) Why does the stock market fluctuate? *Quarterly Journal of Economics* 108, 291–311.

Black, F. (1988) An equilibrium model of the crash. *NBER Macroeconomics Annual* 3, 269–275.

Campbell, J.Y. & R.J. Shiller (1988) Stock prices, earnings, and expected dividends. *Journal of Finance* 43, 661–676.

Cobb, L., P. Koppstein & N.H. Chen (1983) Estimation and moment recursion relations for multimodal distributions of the exponential family. *Journal of the American Statistical Association* 78, 124–130.

Cochrane, J.H. (1992) Explaining the variance of price dividend ratios. *Review of Financial Studies* 5, 243–280.

Cox, D.R. & H.D. Miller (1984) *The Theory of Stochastic Processes*. London: Chapman and Hall.

Craine, R. (1993) Rational bubbles: A test. *Journal of Economic Dynamics and Control* 17, 829–846.

Creedy, J., J.N. Lye & V.L. Martin (1996) A non-linear model of the real US/UK exchange rate. *Journal of Applied Econometrics* 11, 669–686.

Das, S.R. (1993) Jump-Hunting Interest Rates. Department of Finance working paper, New York University.

Dezhbakhsh, H. & A. Demirguc-Kunt (1990) On the presence of speculative bubbles in stock prices. *Journal of Financial and Quantitative Analysis* 25, 101–112.

Diba, B.T. & H.I. Grossman (1988) Explosive rational bubbles in stock prices? *American Economic Review* 78, 520–530.

Flood, R.P. & P.M. Garber (1980) Market fundamentals versus price level bubbles: The first tests. *Journal of Political Economy* 89, 745–770.

Frankel, J.A. & K.A. Froot (1986) Understanding the US dollar in the eighties: The expectations of chartists and fundamentalists. *Economic Record* Suppl. 24–38.

French, K.R. (1988) Crash-testing the efficient market hypothesis. *NBER Macroeconomics Annual* 3, 277–285.

Froot, K.A. & M. Obstfeld (1991) Intrinsic bubbles: The case of stock prices. *American Economic Review* 81, 1189–1214.

Gennotte, G. & H. Leland (1990) Market liquidity, hedging, and crashes. *American Economic Review* 80, 999–1021.

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press

LeRoy, S.F. & R. Porter (1981) The present value relation: Tests based on implied variance bounds. *Econometrica* 49, 555–574.
Lye, J. & V.L. Martin (1993) Robust estimation, non-normalities and general exponential distributions. *Journal of the American Statistical Association* 88, 253–259.
Mankiw, N.G., D. Romer & M. Shapiro (1985) An unbiased reexamination of stock market volatility. *Journal of Finance* 40, 677–687.
Mankiw, N.G., D. Romer & M. Shapiro (1991) Stock market forecastability and volatility: A statistical appraisal. *Review of Economic Studies* 58, 455–477.
Merton, R.C. (1976) Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3, 125–144.
Pindyck, R.S. (1984) Risk, inflation, and the stock market. *American Economic Review* 74, 335–351.
Rubinstein, M. & Leland, H. (1981) Replicating options with positions and cash. *Financial Analysts* 37, 63–72.
Shiller, R.J. (1981) Do stock markets move too much to be justified by subsequent changes in dividends? *American Economic Review* 71, 421–436.
Shiller, R.J. (1988) Portfolio insurance and other investor fashions as factors in the 1987 stock market crash. *NBER Macroeconomics Annual* 3, 287–295.
Shiller, R.J. (1990) Speculative prices and popular models. *Journal of Economic Perspectives* 4, 55–65.
Stiglitz, J.E. (1990) Symposium on bubbles. *Journal of Economic Perspectives* 4, 13–18.
Tucker, A.L. (1991) *Financial Futures, Options & Swaps*. New York: West.
Turner, C.M., R. Startz & C.R. Nelson (1989) A Markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economics* 25, 3–22.
West, K.D. (1987) A Specification test for speculative bubbles. *Quarterly Journal of Economics* 102, 553–580.
West, K.D. (1988) Bubbles, fads and stock price volatility tests: A partial evaluation. *Journal of Finance* 43, 639–656.

https://doi.org/10.1017/S1365100598007044 Published online by Cambridge University Press