Daniel B. Nelson, 1959–1995

With profound sadness, we report that Dan Nelson died on May 4, 1995, at the age of 36. Dan served as an Associate Editor of the *JBES* from 1991 until his death. Dan's remarkable research legacy is summarized in the following survey by Tim Bollerslev and Peter E. Rossi.

Dan leaves his wife, Therese, and three young children, Carolyn, Scott, and Allen. An educational fund has been established for his children. If you would like to contribute, please make checks payable to "Nelson Children College Fund" and mail them to Beth Fama, 5553 S. Kenwood Avenue, Chicago, IL 60637-1775 (pefama@gsbphd.uchicago.edu).

George Tauchen, Ruey Tsay, and Mark Watson

Dan Nelson Remembered

After a long battle with cancer, Dan Nelson passed away on May 4, 1995, at the age of 36. With his untimely death, time series econometrics and asset pricing finance have lost a great scholar, and many of us have lost a great friend and colleague. Dan will be sorely missed. One characterization of the ideal scholar is someone who identifies an important class of unresolved problems, sets about on a systematic attack on this set of problems, and after a period of time achieves closure on many if not all of the issues. This is a very good description of Dan's decade-long research agenda on autoregressive conditional heteroscedasticity (ARCH) models. Dan started his research on ARCH models while a graduate student at MIT. It is important to recognize that he crafted this research agenda by himself and that he emerged as one of the most prominent researchers independently of others in the field. In 1985, when Dan first began his work, there were basically two approaches for modeling time-varying volatility in financial markets. The ARCH class of models, introduced by Robert F. Engle only three years earlier, was already starting to enjoy considerable empirical success. At the same time most of the theoretical literature on the pricing of risky assets was, and still is, based on continuous-time stochastic volatility-type models. In spite of their theoretical

appeal, this latter class of models is much more difficult to analyze empirically, and Dan was persuaded that ARCH-type models were the wave of the future. He also recognized the importance of forging a closer link between the ARCH and stochastic volatility class of models, however, along with several important holes and deficiencies in the existing ARCH literature.

Dan's first major contribution to the ARCH literature was the now classic exponential generalized ARCH (EGARCH) model originally developed in his 1988 MIT thesis and published in 1991 [6]. To formally define this model, let $\{z_t\}$ denote a sequence of iid random variables with $E(z_t) = 0$ and $E(z_t^2) = 1$. Moreover, denote the innovations to the return process by $\epsilon_t = \sigma_t z_t$, where σ_t is a function of the time $t - 1$ information set so that the conditional variance of the return at time t equals σ_t^2. In the earlier ARCH(q) and GARCH(p, q) class of models, σ_t^2 was parameterized as a function of past squared innovations, ϵ_{t-i}^2, and lagged conditional variances, σ_{t-i}^2, only. Although these formulations capture the thick-tailed distributions and volatility clustering phenomena that are characteristic of high-frequency asset returns, the ARCH/GARCH models are not well suited to accommodate any asymmetric effects in the evolution of the

volatility process. In the EGARCH model, $\ln(\sigma_t^2)$ is parameterized as an autoregressive moving average (ARMA) model in the absolute size and the sign of the lagged residuals. In particular, for the AR(1)–EGARCH model,

$$\ln(\sigma_t^2) = \omega + \beta \ln(\sigma_{t-1}^2) + \theta z_{t-1} + \gamma[|z_{t-1}| - E|z_{t-1}|].$$

Thus, for $\gamma, \beta > 0$, large price changes are still followed by large price changes, but with $\theta < 0$ this effect is accentuated for negative price changes, a stylized feature of equity returns often referred to as the "leverage effect." For instance, a stock-market crash is typically followed by a period of much higher volatility than a corresponding upward run in the prices.

The EGARCH model was an instant hit, and in the short time since its introduction more than 100 empirical studies have already employed the model. EGARCH is now also available as a standard procedure in several commercial statistical software packages.

Although the univariate EGARCH model has enjoyed considerable empirical success, many interesting questions in financial economics necessarily call for a multivariate modeling approach. The formulation of multivariate ARCH models poses several practical problems, however, including parameter parsimony and positive definiteness of the conditional covariance matrix estimators. The bivariate version of the EGARCH model of Braun, Nelson, and Sunier [20], designed explicitly to capture any "leverage effects" in the conditional betas of equity returns, represents a particularly elegant solution to both of these problems.

In addition to the asymmetry, the ARMA formulation for $\{\ln(\sigma_t^2)\}$ in the EGARCH model also overcomes many of the difficulties in the probabilistic structure with the ARCH(q) and GARCH(p, q) class of models. For instance, consider the GARCH(1, 1) model,

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

$$= \sigma_0^2 \prod_{i=1}^{t} (\beta + \alpha z_{t-i}^2) + \omega\left[1 + \sum_{j=1}^{t-1}\prod_{i=1}^{j}(\beta + \alpha z_{t-i}^2)\right],$$

where the last equality follows by repeated substitution for $t \geq 2$. In estimating this model with high-frequency data, $\widehat{\alpha} + \widehat{\beta}$ are often found to be very close to 1. From a forecasting perspective, the corresponding IGARCH model with $\alpha + \beta = 1.0$ behaves like a random walk; that is, $E_t(\sigma_{t+s}^2) = \sigma_{t+1}^2 + (s - 1)\omega \to \infty$ almost surely. This finding caused a lot of confusion and concern in the early ARCH literature. As Dan showed in publication [5], however, the persistence of shock to the conditional variance should be very carefully interpreted. The behavior of a martingale can differ markedly from the behavior of a random walk. From the preceding expression, strict stationarity and ergodicity of the GARCH(1, 1) model requires geometric convergence of $\{\beta + \alpha z_t^2\}$, or $E[\ln(\beta + \alpha z_t^2)] < 0$, which is a much less stringent condition than the arithmetic convergence, or $E[\beta + \alpha z_t^2] < 1$, required for the model to be covariance stationary. This important insight also underlies subsequent work on the consistency and asymptotic normality

of maximum likelihood-based estimators for GARCH-type models. Another complication that arises in the GARCH class of models, which is elegantly avoided by the EGARCH formulation for $\ln(\sigma_t^2)$, concerns the parameter restrictions required for the conditional variance function to be positive almost surely. In fact, these necessary and sufficient conditions for the GARCH(p, q) model were first developed by Nelson and Cao [9]. Dan's other seminal contributions to the ARCH literature are embodied in a series of seven or more papers on the relationship between ARCH models and the types of stochastic volatility models typically used in the mathematical finance literature. The method of analysis and the theoretical results in these papers have had a major impact on the thinking in the area. The foundation for these investigations were laid in publication [4], which establishes weak convergence results for sequences of stochastic difference equations (e.g., ARCH models) to stochastic differential equations as the length of the sampling interval between the observations diminishes. For instance, consider the sequence of GARCH(1, 1) models observed at finer and finer time intervals h with conditional variance parameters $\omega_h = \omega hh, \alpha_h = \alpha(h/2)^{1/2}$, and $\beta_h = 1 - \alpha(h/2)^{1/2} - \theta h$, and conditional mean $\mu_h = hc\sigma_t^2$. Under suitable regularity conditions, the diffusion limit of this process equals

$$dy_t = c\sigma_t^2 dt + \sigma_t dW_{1,t}$$
$$d\sigma_t^2 = (\omega - \theta\sigma_t^2)dt + \alpha\sigma_t^2 dW_{2,t},$$

where $W_{1,t}$ and $W_{2,t}$ denote independent Brownian motions. Similarly, the sequence of AR(1)–EGARCH models with $\beta_h = 1 - \beta h$ and the other parameters as defined by Nelson [4] converges weakly to the

$$d(\ln(y_t)) = \theta\sigma_t^2 dt + \sigma_t dW_{1,t}, \quad d(\ln(\sigma_t^2))$$
$$= -\beta(\ln(\sigma_t^2) - \alpha) dt + dW_{2,t}$$

diffusion process commonly employed in the theoretical options-pricing literature.

Even when misspecified, however, appropriately defined sequences of ARCH models may still serve as consistent estimators for the volatility of the true underlying diffusion, in the sense that the difference between the true instantaneous volatility and the ARCH filter estimates converges to 0 in probability as the length of the sampling frequency diminishes. Although the formal proofs for this important result as developed by Nelson [8] are somewhat complex, as with most important insights, the intuition is fairly straightforward. In particular, suppose that the sample path for the instantaneous volatility process, $\{\sigma_t^2\}$, is continuous almost surely. Then for every $\xi > 0$ and every $t > 0$, there exists a $\delta > 0$ such that $\sup_{t-\delta \leq \tau \leq t}|\sigma_\tau^2 - \sigma_t^2| < \xi$. Now, divide this interval into N equal pieces. Given the drift, $\mu_{t-\delta}$, and the instantaneous volatility, $\sigma_{t-\delta}^2$, the N increments $(y_{t-(i-1)\delta/N} - y_{t-i\delta/N})$, $i = 1, 2, \ldots, N$, are then approximately iid normally distributed with mean $N^{-1}\delta\mu_{t-\delta}$ and variance $N^{-1}\delta\sigma_{t-\delta}^2$. A natural estimate for σ_t^2 is, therefore,

$$\widehat{\sigma}_t^2(\delta, N) = \delta^{-1}(y_{t-(i-1)\delta/N} - y_{t-i\delta/N})^2,$$

which under suitable moment conditions converges in probability to σ_t^2 as $\delta \rightarrow 0$ and $N \rightarrow \infty$ by a law of large numbers. Note that the drift term is of second-order importance, so that a failure to account for the drift does not affect the consistency. Now consider the GARCH(1, 1) model,

$$\widehat{\sigma}_t^2 = \omega + \alpha\widehat{\epsilon}_{t-1}^2 + \beta\widehat{\sigma}_{t-1}^2$$

$$= \omega(1 - \beta)^{-1} + \sum_{i=0}^{\infty} \alpha\beta^i\widehat{\epsilon}_{t-1-i}^2.$$

When applied at increasingly higher sampling frequencies, the corresponding GARCH(1, 1) filter with $\omega_h = \omega h$, $\alpha_h = \alpha(h/2)^{1/2}$, and $\beta_h = 1 - \alpha(h/2)^{1/2} - \theta h$ effectively achieves consistency in the same manner, by estimating σ_t^2 as the average of an increasing number of squared residuals close to time t. The AR(1)–EGARCH diffusion approximation discussed previously and many other ARCH filters share this smoothing property. Note that, the autoregressive coefficients in this consistent GARCH(1, 1) filter, $\alpha_h + \beta_h = 1 - \theta h$, and the AR(1)–EGARCH filter, $\alpha_h = 1 - \beta_h$, both converge to unity as $h \rightarrow 0$. This feature of the consistent filters is important, and provides a possible explanation for the widespread empirical findings of apparent IGARCH-type behavior with high-frequency financial data.

Of course, not all consistent ARCH filters will perform equally well in a given situation so that issues of efficiency become important. The basic framework of continuous-record asymptotics developed by Nelson and Foster [12] and Nelson [19] allows for a formal analysis along these lines. The results in both of these papers rely on a Markov assumption. This assumption is dropped in the alternative continuous-record asymptotics developed by Nelson and Foster [17], which analyzes the rolling-regression type of estimators of conditional variances often used by investment professionals. It is impossible to do any justice to these path-breaking approaches in a short summary like this. A useful illustration is provided, however, by Dan's comment [13] on the estimation of stochastic volatility models. In particular, consider the stochastic volatility model

$$y_{t+h} = y_t + h^{1/2}\sigma_t z_{t+h}, \quad \ln(\sigma_{t+h}^2) = \ln(\sigma_t^2) + h^{1/2}\lambda v_{t+h},$$

where z_t and v_t are assumed to be iid normally distributed with correlation ρ. The linear Kalman filter advocated by several researchers provides a simple way of analyzing this model by transforming the measurement equation to $\ln(y_{t+h} + y_t)^2 = \mu + \ln(\sigma_t^2) + [\ln(z_{t+h}) - E(\ln(z_{t+h}))]$. In fact, Dan was the first to employ this approach in his 1988 MIT thesis. Of course, because the measurement error is nonnormal, the linear Kalman filter is suboptimal. In particular, from the results of Nelson and Foster [12], the asymptotically optimal linear Kalman filter achieves an asymptotic variance for $h^{-1/4}[\ln(\widehat{\sigma}_t^2) - \ln(\sigma_t^2)]$ as $h \rightarrow 0$ equal to $\lambda\Psi(1/2)^{1/2}$, where $\Psi(x) = d[\ln(\Gamma(x)]/dx$ denotes the psi function. Meanwhile, the asymptotically optimal ARCH filter,

$$\ln(\widehat{\sigma}_{t+h}^2) = \ln(\widehat{\sigma}_t^2) + \rho\lambda(y_{t+h} - y_t)\widehat{\sigma}_t^{-1} + \lambda(1 - \rho^2)^{1/2}$$

$$\times \left[\Gamma(1/2)^{-1/2}\Gamma(3/2)^{1/2}|y_{t+h} - y_t|\widehat{\sigma}_t^{-1} - 2^{-1/2}\right],$$

yields an asymptotic variance for $h^{-1/4}[\ln(\widehat{\sigma}_t^2) - \ln(\sigma_t^2)]$ as $h \rightarrow 0$ of $2^{1/2}\lambda(1 - \rho^2)^{1/2}$. Comparing these two asymptotic variances, it follows that the efficiency loss from using the suboptimal linear Kalman filter in this context may be quite substantial. Several full-information approaches for estimating stochastic volatility models have recently been proposed in the literature. As illustrated by this example, Dan's work on continuous-record asymptotics is likely to provide invaluable insight on the formulation of filters and moment-generating functions in the future analyses of such models.

Although many different sequences of misspecified ARCH models may provide consistent filtered volatility estimates, the forecasting performance across these filters is likely to be very different. This issue was addressed by Nelson and Foster [19]. Although the higher-order moments can be ignored in the continuous time limit, correctly modeling the first two conditional moments turns out to be crucial for generating accurate volatility forecasts. Thus in practice some experimentation with alternative flexible functional forms designed to capture the relevant stylized facts for the data set at hand will be important for successful forecasting.

Of course, once an ARCH model is viewed merely as an approximation to the true data-generating process, there is no need to restrict the in-sample estimation of the true volatility process to rely on lagged values of the process only. Dan's forthcoming 1995 *Econometrica* article [18], on which he started work shortly before he was diagnosed with cancer, takes up this issue of smoothing and the optimal use of both lagged and led residuals in the estimation of the true volatility. This work represents a natural closure on Dan's decade-long research agenda on the use of ARCH models in approximating continuous time models, and it was a great source of satisfaction to him that he managed to finish the latest revisions on this paper shortly before his death.

Although Dan will be best remembered for his work on ARCH models, his research interests were wide. Most notably, he wrote a well-respected paper on the question of whether the Great Depression was anticipated. This paper, which involved an extensive amount of research into old magazine and newspaper articles, was published as a lead article in *Research in Economic History* [7].

We can only wonder what Dan's next research agenda would have been and what further contributions he would have made. Given his talent and creativity, it is a great shame that we will never know. We do know, however, that the fields of time series econometrics and asset pricing are much richer because of the important research contributions that he did make. It has been said that as long as a person is read and talked about he continues to live. Dan's papers on ARCH models will certainly be read for a long time to come.

Tim Bollerslev
Kellogg School of Management, Northwestern University
Peter E. Rossi
Graduate School of Business, University of Chicago

PUBLISHED AND FORTHCOMING ARTICLES

1989

[1] "Commentary: Price Volatility, International Market Links, and their Implications for Regulatory Policies," *The Journal of Financial Services Research*, 3, 247–254.

[2] "Modeling Stock Market Volatility Changes," *Proceedings of the Business and Economic Statistics Section, American Statistical Association*, 93–98.

1990

[3] "Simple Binomial Processes as Diffusion Approximations in Financial Models" (with Krishna Ramaswamy), *Review of Financial Studies*, 3, 393–430.

[4] "ARCH Models as Diffusion Approximations," *Journal of Econometrics*, 45, 7–38 (lead article).

[5] "Stationarity and Persistence in the GARCH(1, 1) Model," *Econometric Theory*, 6, 318–334.

1991

[6] "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347–370.

[7] "Was the Deflation of 1929–1930 Anticipated? The Monetary Regime as Viewed by the Business Press," *Research in Economic History*, 13, 1–65 (lead article).

1992

[8] "Filtering and Forecasting With Misspecified ARCH Models I: Getting the Right Variance With the Wrong Model," *Journal of Econometrics*, 25, 61–90.

[9] "Inequality Constraints in the Univariate GARCH Model" (with Charles Cao), *Journal of Business & Economic Statistics*, 10, 229–235.

[10] "ARCH Models" in *The New Palgrave Dictionary of Money and Finance, Volume 1*, eds. P. Newman, M. Milgate, and J. Eatwell, London: MacMillan Press Ltd., pp. 56–57.

[11] "Filtering With ARCH Models: A Monte Carlo Investigation" (with Boaz Schwartz), *Proceedings of the Business and Economic Statistics Section, American Statistical Association*, 1–6.

1994

[12] "Asymptotic Filtering Theory for Univariate ARCH Models" (with Dean Foster), *Econometrica*, 62, 1–41 (lead article).

[13] Comment on Jacquier, Polson, and Rossi's "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business & Economic Statistics*, 12, 403–406.

[14] "ARCH Models," Chap. 49 in *The Handbook of Econometrics, Volume 4*, eds. R. F. Engle and D. L. McFadden (with Tim Bollerslev and Robert Engle).

Forthcoming 1995

[15] "Filtering and Forecasting With Misspecified ARCH Models II: Making the Right Forecast With the Wrong Model" (with Dean Foster), forthcoming in *The Journal of Econometrics*.

[16] "A Note on the Normalized Errors From ARCH and Stochastic Volatility Models," forthcoming in *Econometric Theory*.

[17] "Continuous Record Asymptotics for Rolling Sample Variance Estimators" (with Dean Foster), forthcoming in *Econometrica*.

[18] "Asymptotic Smoothing Theory for ARCH Models," forthcoming in *Econometrica*.

[19] "Asymptotic Filtering Theory for Multivariate ARCH Models," accepted with minor revisions at *the Journal of Econometrics*.

[20] "Good News, Bad News, Volatility and Betas" (with Phillip Braun and Alain Sunier), forthcoming in *Journal of Finance*.