Bayesian Specification Analysis in Econometrics

John Geweke and William McCausland

All econometric models are wrong, but some are useful. They mirror only certain aspects of reality, and these imperfectly. However, in complex situations they can provide structure and clarity that improve decisions. The process of discovering the reliability with which various aspects of reality are accommodated in an econometric model is known as specification analysis. It is a vital step in learning about the properties of a given model, in determining whether to use a model in actual decision making, and in improving models, and thereby, decisions.

This article outlines the essentials of Bayesian specification analysis, as practiced using state of the art simulation methods. This approach is especially pertinent to models used for decision making, because Bayesian inference is the econometric cornerstone of decision making within the expected utility theory on which virtually all modern economics is constructed (Friedman and Savage; Savage). It is attractive in practical work because it is straightforward to apply and can be used to study the congruence of any complete econometric model with relevant features of the data.

This topic has received close attention in the Bayesian mathematical statistics literature but is less well known among practicing applied economists. There are two approaches, sometimes combined in various ways. The first is to ask: having expressed a model, what are its predictions for observables, before using data for inference? How consistent are these predictions with the actual data? The classic exposition of this approach is Box, and a recent treatment using modern computational techniques is Geweke (1999b). Combining the terminology of this literature and that of econometrics, we use the term *predictive specification analysis* for this approach here.

The second approach is to ask: having expressed a model, and having used the observed data for inference about its parameters, what would we predict would happen in an independent replication of the observables? (In a prediction problem this is similar to asking what would happen in the next T observations, where T is the size of the original sample.) The event that the replications are quite different from what was actually observed, for some interesting aspect of the data, constitutes the notion of surprise: this idea has been developed in a series of important studies including Bayarri and Berger, and Guttman. We use the term *postpredictive specification analysis* for this approach.

Bayesian Analysis

The concept of a *complete model* is central to Bayesian specification analysis. A complete model specifies the distribution of $T \times 1$ vector of observable random variables, $\mathbf{y} \in Y$, by means of a $k_A \times 1$ vector of unknown parameters $\boldsymbol{\theta}_A \in \boldsymbol{\Theta}_A$. Conditional on $\boldsymbol{\theta}_A$ the probability density of the observables is

$$(1) \quad p(\mathbf{y} \mid \boldsymbol{\theta}_A, A) = \prod_{t=1}^{T} p(y_t \mid y_1, \ldots, y_{t-1}, \boldsymbol{\theta}_A, A).$$

This function is familiar from non-Bayesian analysis—after $\mathbf{y}$ is observed, and the observed value $\mathbf{y}^o$ replaces the argument $\mathbf{y}$, (1) becomes the likelihood function.

A complete model includes a prior density for the unobservable parameters, $p(\boldsymbol{\theta}_A \mid A)$. The combination of the prior density and the density for observables (1) provides the joint density of parameters and observables:

$$(2) \quad p(\boldsymbol{\theta}_A, \mathbf{y} \mid A) = p(\boldsymbol{\theta}_A \mid A) p(\mathbf{y} \mid \boldsymbol{\theta}_A, A).$$

John Geweke holds the Harlan E. McGregor Chair in economic theory at the University of Iowa. William McCausland is assistant professor at the University of Montreal. Partial financial support from national Science Foundation grants SBR-9731037 and SES-9996332 is gratefully acknowledged. George Chang provided help with the computations and useful discussions. The usual disclaimers apply.

This article was presented in a principal paper session at the AAEA annual meeting (Chicago, IL, August 2001). The articles in these sessions are not subjected to the Journal's standard refereeing process.

Amer. J. Agr. Econ. 83 (Number 5, 2001): 1181–1186

This joint distribution is the key to Bayesian inference and specification analysis. The marginal density for $\mathbf{y}$,

$$(3) \quad p(\mathbf{y} \mid A) = \int_{\Theta_A} p(\boldsymbol{\theta}_A, \mathbf{y} \mid A)\, d\boldsymbol{\theta}_A$$

is the model predictive density of the observable $\mathbf{y}$. Conditional on observed $\mathbf{y} = \mathbf{y}^o$,

$$(4) \quad p(\boldsymbol{\theta}_A \mid \mathbf{y}^o, A) = p(\boldsymbol{\theta}_A, \mathbf{y}^o \mid A)/p(\mathbf{y}^o \mid A)$$

the posterior density of the unknown parameter vector $\boldsymbol{\theta}_A$.

The third and final component of a complete model is a vector of interest $\boldsymbol{\omega}$. Elements of $\boldsymbol{\omega}$ may include transformations of parameters $\boldsymbol{\omega} = g(\boldsymbol{\theta}_A; A)$, for example the value of returns to scale in a production function. They also include observables whose values are not yet known, for example $\boldsymbol{\omega}' = (y_{T+1}, y_{T+2}, y_{T+3})$ in a forecasting problem. A complete model provides the density $p(\boldsymbol{\omega} \mid \boldsymbol{\theta}_A, \mathbf{y}, A)$,

$$(5) \quad p(\boldsymbol{\omega} \mid A) = \int_{\Theta_A} \int_{Y} p(\boldsymbol{\theta}_A, \mathbf{y} | A) \times p(\boldsymbol{\omega} \mid \boldsymbol{\theta}_A, \mathbf{y}, A)\, d\mathbf{y}\, d\boldsymbol{\theta}_A,$$

and the posterior density of the vector of interest is

$$(6) \quad p(\boldsymbol{\omega} | \mathbf{y}^o, A) = \int_{\Theta_A} p(\boldsymbol{\theta}_A | \mathbf{y}^o, A) \times p(\boldsymbol{\omega} | \boldsymbol{\theta}_A, \mathbf{y}^o, A) d\boldsymbol{\theta}_A.$$

Expressions like (6) are not immediately useful in applications, because the integral on the right side typically cannot be obtained in closed form. In modern Bayesian analysis this problem is usually obviated by means of a *posterior simulator*. This is an algorithm that produces random vectors $\boldsymbol{\theta}_A^{(m)}$ $(m = 1, 2, \ldots)$ whose distribution corresponds to the posterior density (4).[1] Simulation from the other three distributions, corresponding to the prior density $p(\boldsymbol{\theta}_A \mid A)$, the observables density $p(\mathbf{y} \mid \boldsymbol{\theta}_A, A)$, and the vector of interest density $p(\boldsymbol{\omega} \mid \boldsymbol{\theta}_A, \mathbf{y}, A)$ is typically straightforward. Taken together, these simulators make it possible to generate synthetic random vectors from each of the above distributions.

[1] These algorithms use varied methods, some quite sophisticated, to achieve this correspondence. Moreover, the nature of the correspondence varies with the simulator. These details are beyond the scope of this article, but there are quite a few accessible surveys and texts, including Chib and Greenberg (1996), Gelman, Carlin, Stern, and Rubin (1995), and Geweke (1999a).

The *t*-GARCH Model

The modeling of returns to financial assets has emerged as a challenging problem of considerable practical importance in recent years. The importance stems from such decision problems as the pricing of financial derivatives like options and strategies for avoiding risk such as hedging. In all these situations, the distribution of asset returns, and in particular the conditional distribution of future asset returns, is central to rational decision making. The problem is interesting, because the distributions in question are clearly not normal and do not appear to be of any other simple form. There is also strong evidence that the spread of distributions, and possibly the shape, changes as conditioning information evolves.

A leading model that captures some of these characteristics for a single asset return is the generalized autoregressive conditional heteroskedasticity model (Bollerslev) with Student-*t* shocks (*t*-GARCH). Denote the observable asset return from period $t-1$ to period t by y_t and the variance of y_t at time t by h_t. Given h_t,

$$(7) \quad y_t \sim t(\mu, h_t; \nu)$$

and variance evolves as

$$(8) \quad h_t = \alpha + \gamma(y_{t-1} - \mu)^2 + \delta h_{t-1}.$$

So long as $|\gamma + \delta| < 1$ the return series $\{y_t\}$ is stationary and displays periods of both high volatility (large values of $(y_t - \mu)^2$) and low volatility (small $(y_t - \mu)^2$) relative to its unconditional variance of $\alpha/(1 - \gamma - \delta)$. The unconditional distribution of y_t is nonnormal—it displays excess kurtosis, and the accompanying "fat tails" in its unconditional probability density relative to the normal.

Predictive Specification Analysis

The *t*-GARCH distribution of observables does not, alone, constitute a complete model. It remains to specify prior distributions, as well as a vector of interest. The process of predictive specification analysis consists of three steps: (1) choosing a vector of interest that summarizes interesting aspects of the data, $\boldsymbol{\omega}$; (2) selecting a trial prior distribution; and (3) examining the implications of the prior distribution for the vector of interest $\boldsymbol{\omega}$.

Table 1. Definition of Vector of Interest

Preliminary Statistics: $\bar{y}_T = \sum_{t=1}^{T} y_t/T$; $\bar{y}_T^{(2)} = \sum_{t=1}^{T} y_t^2/T$		$S_T = \sum_{t=1}^{T}(y_t - \bar{y}_T)^2/T$; $S_T^{(2)} = \sum_{t=1}^{T}(y_t^2 - \bar{y}_T^{(2)})^2/T$
ω_1	First-order volatility	$\sum_{t=1}^{T-1}(y_t^2 - \bar{y}_T^{(2)})(y_{t+1}^2 - \bar{y}_T^{(2)})/(T s_T^{(2)})$
ω_2	Twentieth order volatility	$\sum_{t=1}^{T-20}(y_t^2 - \bar{y}_T^{(2)})(y_{t+20}^2 - \bar{y}_T^{(2)})/(T s_T^{(2)})$
ω_3	Volatility decay	ω_2/ω_1
ω_4	Excess kurtosis	$\sum_{t=1}^{T}(y_t - \bar{y}_T)^4/T(s_T)^2 - 3$
ω_5	Quantile ratio	$(y_{(T)} - y_{(1)})/(y_{(3T/4)} - y_{(T/4)})$
ω_6	Skewness	$\sum_{t=2}^{T}(y_t - \bar{y}_T)^3/T(s_T)^{3/2}$
ω_7	Leverage ahead	$\sum_{t=1}^{T-1}(y_t - \bar{y}_T)(y_{t-1}^2 - \bar{y}_T^{(2)})/T(s_T \cdot s_T^{(2)})^{1/2}$
ω_8	Leverage behind	$\sum_{t=2}^{T}(y_t - \bar{y}_T)(y_{t+1}^2 - \bar{y}_T^2)/T(s_T \cdot s_T^{(2)})^{1/2}$
ω_9	Standard deviation	$(S_T)^{1/2}$

Steps (2) and (3) may be repeated, experimenting with different prior distributions. This process is best described by illustration.

The outstanding statistical characteristics of returns to financial assets tend to be changing but persistent volatility, excessive leptokurtosis relative to the normal distribution, and the "leverage" phenomenon in which extreme negative returns are more likely to presage high volatility than similarly extreme positive returns. For some return series, the distribution of asset returns is also skewed to the left. These characteristics can all be captured through transformations $\boldsymbol{\omega}$ of the observable returns $\mathbf{y}$. The transformed observables used in this study, which constitute the vector of interest, are detailed in table 1.

To illustrate Bayesian specification analysis, we shall apply the t-GARCH model described in the previous section to the daily returns of the Standard and Poors 500 index used in Ryden, Terasvirta, and Asbrink. It extends from January 3, 1928, through April 29, 1991, a total of 17,052 observations. Returns are formed as $y_t = \log(p_t/p_{t-1})$, where p_t is the daily index (see Ryden, Terasvirta, and Asbrink for complete details). Table 2 provides, in the first column, the observed values $\boldsymbol{\omega} = \boldsymbol{\omega}^o$ for this period.

To interpret these values, consider a model for returns that is much simpler than those described in the previous section: $y_t \sim N(\mu, \sigma^2)$ i.i.d. The sampling distribution of $\omega_1, \ldots, \omega_8$ does not depend on the value of μ or σ^2. Therefore the model predictive distribution of each of ω_1 through ω_8, for a sample of size 17,052 (or any other size) will be the same no matter what the prior distribution of the unknown parameters μ and σ^2. Some quantiles of this distribution are also indicated in table 2.[2]

The failure of the normal model is strikingly evident. None of the characteristics of the data captured in the functions of interest ω are consistent with an i.i.d. normal model. The observed values of volatility ($\omega_1^o, \omega_2^o, \omega_3^o$), the thickness of the tails of the distribution (ω_4^o, ω_5^o), skewness (ω_6^o) and leverage (ω_7^o, ω_8^o) are so improbable as to be impossible for practical purposes.

In the t-GARCH model, as well as in the other models considered in this article, the model predictive distribution of $\boldsymbol{\omega}$ depends on the prior density $p(\boldsymbol{\theta}_A \mid A)$ as well as the data density (1). This is generally the case, as is evident from (2) and (5). It is therefore necessary to specify a proper prior distribution for the parameters μ, α, γ and δ of the t-GARCH model. Given the prior density $p(\boldsymbol{\theta}_A \mid A)$, the simulation $\tilde{\boldsymbol{\theta}}\bullet_A \sim p(\boldsymbol{\theta}_A \mid A)$ followed by $\tilde{\mathbf{y}} \sim p(\mathbf{y} \mid \tilde{\boldsymbol{\theta}}\bullet_A, A)$ and followed by the computation of $\boldsymbol{\omega}$ as indicated in table 1 (with $\tilde{y}_t$ in place of y_t) produces a single drawing from the predictive distribution with density (5). Repetition of this process many times provides quantiles for the predictive density $p(\boldsymbol{\omega} \mid A)$.

The t-GARCH model does not fail in this way. Table 3 provides quantiles for the predictive density of $\boldsymbol{\omega}$ for a t-GARCH model

[2] These quantiles and those in tables 3 and 4 are based on 1,000 simulations.

Table 2. Predictive Distribution of Vector of Interest Gaussian i.i.d. Model

		Data	Median	(25%, 75%)	(1%, 99%)
ω_1	First-order volatility	0.218	0.000	(−0.005, 0.005)	(−0.018, 0.018)
ω_2	Twentieth order volatility	0.083	0.000	(−0.005, 0.005)	(−0.018, 0.018)
ω_3	Volatility decay	0.382	0.000	(−1.00, 0.982)	(−31.3, 31.9)
ω_4	Excess kurtosis	22.4	−0.001	(−0.026, 0.024)	(−0.083, 0.090)
ω_5	Quantile ratio	39.9	5.874	(5.686, 6.089)	(5.301, 6.761)
ω_6	Skewness × 100	−0.373	0.000	(−0.010, 0.010)	(−0.030, 0.030)
ω_7	Leverage ahead	0.0312	0.0000	(−0.0052, 0.0053)	(−0.0177, 0.0179)
ω_8	Leverage behind	−0.0752	0.0000	(−0.0051, 0.0052)	(−0.0180, 0.0176)
ω_9	Standard deviation	0.0115			

Table 3. Predictive Distribution of Vector of Interest *t*-GARCH Model

		Data	Median	(25%, 75%)	(1%, 99%)
ω_1	First-order volatility	0.218	0.334	(−0.187, 0.459)	(0.004, 0.718)
ω_2	Twentieth order volatility	0.083	0.001	(−0.002, 0.012)	(−0.014, 0.244)
ω_3	Volatility decay	0.382	0.004	(−0.007, 0.052)	(−0.458, 0.935)
ω_4	Excess kurtosis	22.4	17.2	(3.23, 196)	(0.634, 3607)
ω_5	Quantile ratio	39.9	29.88	(14.6, 113)	$(7.98, 6.0 \times 10^5)$
ω_6	Skewness	−0.004	0.000	(−0.002, 0.002)	(−0.204, 0.211)
ω_7	Leverage ahead	0.0312	0.000	(−0.027, 0.026)	(−0.343, 0.344)
ω_8	Leverage behind	−0.0752	0.000	(−0.028, 0.027)	(−0.350, 0.352)
ω_9	Standard deviation	0.0115	0.008	(0.003, 0.025)	(0.001, 2266)

with the prior distribution

(9) $\log(\alpha) \sim N(-12, 2.2)$

(10) $(\gamma, \delta, 1-\gamma-\delta) \sim Beta(1,1,1)$

(11) $v - 4 \sim \chi^2(4)$.

The distribution (10) corresponds to a "flat" prior for γ and δ defined on the unit simplex $\{\gamma > 0, \delta > 0, \gamma + \delta < 1\}$. Alternatives to (10)—for example, Beta distributions with somewhat different values of the parameters—have distinct but small effects on the predictive distribution of $\boldsymbol{\omega}$. The restriction $v > 4$ in (11) ensures the existence of population conditional fourth moments; without this restriction, the sample measures of volatility and excess kurtosis, for $T = 17{,}052$, can become so large as to overflow computer floating point representation. Finally, (9) was chosen to bring ω_9^o well within its support. The quantiles shown in table 3 indicate that *t*-GARCH can account for all the observed ω_i^o.

This elucidation of the prior density $p(\boldsymbol{\omega} \mid A)$ corresponding to a prior distribution $p(\boldsymbol{\theta}_A \mid A)$ is key in choosing prior distributions; indeed, it is hard to see how else subjective prior distributions can be elucidated.[3] It also indicates whether the model at hand (data density combined with prior density) can account for individual $\boldsymbol{\omega}_i^o$. But it says nothing about whether the model is consistent with the *entire* observed vector of interest. To examine this question it is necessary to move to the posterior density (4) and the corresponding density of the vector of interest (6).

Postpredictive Specification Analysis

Consider the following, conceptual experiment. We have observed data $\mathbf{y}^o$, and the corresponding vector of interest $\boldsymbol{\omega}^o$, in an experiment that can be repeated. Then, given the complete model A and the data, the predicted distribution of the vector of interest $\boldsymbol{\omega}$ over future experiments is indicated by the density (6). The observed $\boldsymbol{\omega}^o$, in the context of this distribution, tells us much about the model A. An extreme case is $p(\boldsymbol{\omega}^o \mid \mathbf{y}^o, A) = 0$. Because $\boldsymbol{\omega}^o$ is computed directly from $\mathbf{y}^o$,

[3] There is a substantial literature on elicitation of prior distributions that builds on this fact (see for example Kadane and Wolfson 1998).

Table 4. Postpredictive Distribution of Vector Interest t-GARCH Model

		Data	Median	(25%, 75%)	(1%, 99%)
ω_1	First-order volatility	0.218	0.474	(0.350, 0.586)	(0.101, 0.777)
ω_2	Twentieth order volatility	0.083	0.000	(−0.0004, 0.003)	(−0.002, 0.132)
ω_3	Volatility decay	0.382	0.000	(−0.001, 0.007)	(−0.005, 0.256)
ω_4	Excess kurtosis	22.4	1312	(712, 2388)	(184, 7683)
ω_5	Quantile ratio	39.9	1031	(558, 2124)	(217, 22940)
ω_6	Skewness	−0.004	−0.003	(−0.086, 0.082)	(−0.522, 0.468)
ω_7	Leverage ahead	0.0312	−0.001	(−0.155, 0.146)	(−0.534, 0.510)
ω_8	Leverage behind	−0.0752	0.006	(−0.178, 0.164)	(−0.544, 0.553)
ω_9	Standard deviation	0.0115	0.028	(0.018, 0.054)	(0.010, 0.504)

this can happen only if $p(\mathbf{y}^o \mid \boldsymbol{\theta}_A, A) = 0$ for all $\boldsymbol{\theta}_A$ in the support of the prior density $p(\boldsymbol{\theta}_A \mid A)$: the observed $\mathbf{y}^o$ is impossible, and hence, $p(\mathbf{y}^o \mid A) = 0$. A less extreme outcome is one in which the observed $\boldsymbol{\omega}^o$ is implausible, in the sense that $p(\boldsymbol{\omega}^o \mid \mathbf{y}^o, A)$ is quite small, or in the sense that $\boldsymbol{\omega}_i^o$ lies in the extreme tails of $p(\omega_i \mid \mathbf{y}^o, A)$; the two are usually equivalent.[4] Such an outcome is a surprise, in the sense that the probability of the observed event occurring again in a great many repetitions of the experiment is quite low.

Of course time series are not repeated experiments. But for a long stationary time series with T observations, a nearly equivalent conceptual experiment is to ask about the next T observations instead of the next experiment, and the T observations after those, and so on.

Carrying out a postpredictive analysis requires little additional effort, given the draws $\tilde{\boldsymbol{\theta}}_A^{(m)} \sim p(\boldsymbol{\theta}_A \mid \mathbf{y}^o, A)$ from a posterior simulator. We simply repeat the exercise of Section 4, using these $\tilde{\boldsymbol{\theta}}_A^{(m)}$ in place of the draws from the prior distribution.

The results of this exercise for the t-GARCH model are shown in table 4. From the postpredictive quantiles of ω_1, ω_2, and ω_3 we see that the slow rate of decay in the autocorrelation function of y_t^2 is inconsistent with the t-GARCH specification. (The first autocorrelation is predicted to be higher, the twentieth lower, and the ratio of the twentieth to the first is predicted to be much lower.) From the postpredictive quantiles of ω_4 and ω_5, it is evident that the t-GARCH model implies tails in the unconditional distribution of y_t that are in general too thick. (The distribution of the excess kurtosis lies well above that observed, and the observed quantile ratio is in the bottom 1% quantile of the postpredictive distribution.) Observed skewness and leverage are well within the support of the postpredictive distribution, but that distribution implies that, in general, we should expect sample skewness and leverage much greater in magnitude than is, in fact, the case.

The analysis in table 4 presents a series of specification problems with t-GARCH. (A full exploration of these problems is beyond the scope of this article.) An important clue is the unsuccessful effort of these models to accommodate the slow decay in volatility correlations. This leads to high values of $\gamma + \delta$: the postpredictive interquartile range for the t-GARCH model is $(0.9948, 0.9975)$, and the centered postpredictive 98% interval is $(0.9913, 0.9998)$. Such models are close to being integrated GARCH (IGARCH) models; in these models, the observed value of y_t eventually collapses about its unconditional mean, but in the intervening period, very large values of $|y_t|$ typically arise (Geweke 1986). This characteristic is evident in the postpredictive distribution of the largest absolute return in the series; it is 0.228 in the dataset, whereas for the t-GARCH model the median of the postpredictive distribution of the largest absolute return is 0.929, the interquartile range is $(0.480, 1.60)$, and the 98% centered interval is $(0.298, 87.0)$. Some resolution of these difficulties is provided by fractionally integrated GARCH (FIGARCH) models (Baillie, Bollerslev, and Mikkelsen (1996)), whose consideration is beyond the scope of this article.

---

[4] The distinction is far from innocuous, however; see Bayarri and Berger (1999).

Conclusion

This example illustrates how Bayesian specification analysis can be used to capture the implications of models for observables. The goal of this analysis is to highlight inconsistencies between the models and the observed data, thus increasing our understanding of models and sowing the seeds for the development of better models and improved decision making.

References

Baillie, R.T., T. Bollerslev, and H.O. Mikkelsen. "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity." *J. Econometrics* 74(September 1996):3–30.

Bayarri, M.J., and J.O. Berger. "Quantifying Surprise in the Data and Model Verification." *Bayesian Statistics 6: Proceedings of the Sixth Valencia International Meeting.* J.M. Bernardo, et al., eds., pp. 53–82. Oxford: Oxford University Press, 1999.

Bollerslev, T. "Generalized Autoregressive Conditional Heteroskedasticity." *J. Econometrics* 31 (April 1986):307–27.

Box, G.E.P. "Sampling and Bayes Inference in Scientific Modeling and Robustness." *J. Royal Stat. Soc.* Ser. A, 143(4)(1980):383–430.

Chib, S., and E. Greenberg. "Markov Chain Monte Carlo Simulation Methods in Econometrics." *Econometric Theory* 12(August 1996):409–31.

Friedman, M., and L.J. Savage. "The Utility Analysis of Choices Involving Risk." *J. Polit. Econ.* 56(August 1948):279–304.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. *Bayesian Data Analysis.* London: Chapman & Hall, 1995.

Geweke, J. "Comment on 'Modelling the Persistence of Conditional Variances'." *Econometric Rev.* 5(1986):57–61.

——. "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication." (with discussion and rejoinder). *Econometric Rev.* 18(1999a): 1–126.

——. "Simulation Methods for Model Criticism and Robustness Analysis." In *Bayesian Statistics 6: Proceedings of the Sixth Valencia International Meeting*, J.M. Bernardo, et al., eds., pp. 53–82. Oxford: Oxford University Press, 1999b.

Guttman, I. "The Use of the Concept of a Future Observation in Goodness-of-fit Problem." *J. Royal Stat. Soc.* Ser. B 29(1967): 83–100.

Kadane, J.B. and L.J. Wolfson. "Experiences in Elicitation." *J. Royal Stat. Soc.* Ser. D, 47(1998): 3–19.

Ryden, T., T. Terasvirta, and S. Asbrink. "Stylized Facts of Daily Return Series and the Hidden Markov Model." *J. Appl. Econometrics* 13(May–June 1998):217–44.

Savage, L.J. "Bayesian Statistics." *Recent Developments in Information and Decision Processes.* R.E. Machol and P. Gray, eds., New York: Macmillan, 1962.