The Story of GARCH: A Personal Odyssey

This version: September 6, 2022

Tim Bollerslev*
Duke University, NBER and CREATES

Abstract

I provide a brief history of the origins of the GARCH model and my 1986 paper published in the *Journal*, along with a discussion of how the GARCH model and applications thereof have flourished since then. I also briefly highlight connections to the more recent realized volatility literature.

Keywords: GARCH; volatility clustering; volatility modeling and forecasting; realized volatility.

JEL: C22, C32, C53, C58, G1

*I would like to thank Torben Andersen, Frank Diebold, Andrew Patton and George Tauchen for their comments and suggestions.
Address: Department of Economics, Duke University, Durham, NC 27708, USA; 919-660-1846; boller@duke.edu.

I am grateful to the editors of the *Journal* for this opportunity to offer some reflections on my 1986 paper and the GARCH model (Bollerslev (1986)), which unwittingly shaped much of my research for several years thereafter and in turn serendipitously ended up playing such an important role for my professional career more generally. I will begin with a personal anecdote of how the model was conceived, followed by a discussion of the successes and widespread empirical usage of the model. I will also briefly touch on the relation to more recent so-called realized volatility measures, before concluding with a few final thoughts.

1. Making of the GARCH Model

Trying to recall the birth of the GARCH model, it must have been sometime in late Fall 1984 or Winter 1985, when I was a second year Ph.D. student at UCSD. I was fortunate to have just started working as a research assistant for Rob Engle. My job assignment at the time was to help complete the estimation and testing for the empirical portion of Rob's paper with David Lilien and Russ Robins on the ARCH-M model (Engle, Lilien, and Robins (1987)). One of the issues that Rob and his coauthors were looking at was how to enter the conditional variance in the conditional mean equation to best capture the risk-return tradeoff relationship (in the finally preferred model reported in the published paper, they settled on using the log of the variance). Another equally important empirical issue concerned the form of the ARCH conditional variance equation and how to most succinctly capture the observed dynamic dependencies (after a lot of experimentation with the estimation of different lag lengths and zero-restrictions, the finally preferred model reported in the published paper ended up using the same linearly declining lag structure used in Rob's original ARCH paper, Engle (1982)). I distinctly remember sitting in Rob's office poring over stacks of computer outputs (this was before the days of PCs, at least for graduate students) discussing the estimation results, when David Hendry (who was visiting UCSD at the time) stepped in. Although I can't remember exactly how the discussion transpired, I do remember that at some point David questioned whether the ARCH model should not be interpreted as an MA model rather than an AR model for variances, and whether it wouldn't be possible to think about more general ARMA type models? I went home piqued by David's question, and came back to the graduate student computer lap the next day to program up the very first GARCH model. To my great delight the model seemed to fit the data very well, with the simplest version of the model in which the conditional variance only depends on its own lag and the lagged squared innovation,

$$\sigma_{t+1}^2 \;=\; \omega \;+\; \alpha\varepsilon_t^2 \;+\; \beta\sigma_t^2,$$

performing particularly well.

Having formulated and successfully estimated the model, I remember contemplating what to name the model? ARMACH might have seemed like the obvious choice, but I didn't like the ring of that. Plus, the analogy to ARMA models for second order moments was not entirely complete and worked out yet. Looking back at my notes, another contender was the LARCH model, as in Long-memory ARCH. With the benefit of hindsight, I am obviously happy that I didn't go with that misleading name. Of course, Generalized ARCH is not an entirely appropriate name for the model either. While it is true that the GARCH(p,q) model represents a generalization of the ARCH(q) model, under appropriate regularity conditions it is still a special case of the more general ARCH(∞) model, in much the same way that a covariance stationary ARMA(p,q) model may alternatively be expressed in terms of its infinite MA Wold type representation. In retrospect, and perhaps somewhat unfortunate, my choice of the name GARCH might also inadvertently have helped stir the proliferation of acronyms associated with the many variations of the basic model subsequently proposed in the literature (more on this below). I don't think that the name ARMACH would have lend it itself quite so naturally to as many tweaks and tongue-twisters as did GARCH.

I first submitted the GARCH paper to the *Journal* for possible publication in May, 1985. I received two referee reports back in November 1985, along with a letter from Dennis Aigner, the Co-Editor in charge of the submission, saying that "Our readers agree that your paper is interesting and an important contribution to the literature. I am therefore pleased to accept it for publication in the *Journal* subject to appropriate revision." I resubmitted the revised version in February, 1986, which was then immediately accepted for publication. Little did I know that the publication process is not always as smooth and painless as that! One of the two referees, who later revealed himself, was Sastry Pantula, an Assistant Professor of Statistics at North Carolina State University (NCSU) at the time, and later Distinguished Professor and President of the American Statistical Association (ASA). Sastry had a number of very helpful comments related to the ARMA interpretation of the model, which clearly helped sharpen and improve the paper. To this date, I have no idea who the second referee was. Much on point, however, in her/his summary of the paper the second referee did note that "the justification for considering the models is largely pragmatic, i.e., they appear to fit well, especially in situations involving uncertainty."

2. The World According to GARCH

The motivating empirical illustration in the 1986 paper, and the first ever published empirical application of the GARCH model, concerned uncertainty of inflationary expectations, specifically the conditional variances of 1948-1983 US quarterly inflation rates (incidently the motivating empirical application in Rob Engle's Nobel Prize winning 1982 paper on the ARCH model also pertained to inflation, albeit 1958-1977 UK quarterly inflation rates). However, there is no question that the great empirical success of the GARCH model primarily stems from its ability to succinctly capture volatility clustering in financial rates of returns.

It is difficult to exactly pinpoint the historically first recorded observation that the volatility in financial markets changes through time. Mandelbrot (1963) famously notes that when looking at price changes: "... large changes tend to be follow by large changes - of either sign - and small changes tend to be followed by small changes ... ." But that quote appears in the very last section of the paper, which is otherwise concerned with characterizing unconditional return distributions. The working paper by Rosenberg (1972), which remained unpublished until its inclusion in the collection *Stochastic Volatility: Selective Readings* edited by Neil Shephard (Shephard (2005)), contains the maybe first systematic empirical evidence that volatility is predictable, as evidenced by regressions of squared monthly stock market returns on sums of lagged squared monthly returns. These simple regressions are also motivated by the implications of informal stochastic volatility type models reminiscent of ARCH and GARCH type formulations. Nonetheless, most of the discussion in the paper centers on explaining excess kurtosis of returns by non-stationary volatility, rather than actually modeling and forecasting volatility per se. Another oft cited early study in support of volatility clustering is Black (1976). The empirical analyses in that short note are also quite remarkable more generally in terms of foreshadowing many of the subsequent developments in the GARCH literature, asymmetric GARCH models explicitly designed to capture the so-called "leverage effect" included. At the same time, the closing two sentences in Black (1976) are quite dismissive about the use of more formal model-based procedures: "I don't dare write down any sort of formal model of the process by which volatilities change. I'm not sure I ever will."

Meeting Black's challenge head on, the GARCH models for exchange rates in Engle and Bollerslev (1986) and the GARCH models for equity index returns in Bollerslev (1987) and French, Schwert, and Stambaugh (1987) were arguably among the very first formal such models.[1] As a newly minted PhD, I also still remember how happy I was to see someone not directly associated with UCSD use the GARCH model, and that usage being published in one of the leading finance journals no less. I am also convinced

[1] As a sidebar and a testament to the different times, the weekly exchange rate series analyzed in Engle and Bollerslev (1986) were *hand-collected* and kindly provided to us by Frank Diebold, who had previously estimated ARCH models for the same data in his 1986 PhD dissertation (Diebold (1986)).

that this early well-published and insightful paper by Ken French, Bill Schwert and Rob Stambaugh played a very fortuitus role in "spreading the word" outside econometrics and popularizing the model more widely in asset pricing finance. In fact, it wasn't long thereafter that empirical applications of GARCH to financial data truly started to abound. My 1992 paper with Ray Chou and Ken Kroner, two former UCSD classmates, published in the *Journal* (Bollerslev, Chou, and Kroner (1992)) provides a review of these first-generation empirical finance applications.

Commensurate with this rapid growth in empirical applications, a parallel and equally large stream of more theoretically oriented econometric research seeking to clarify the theoretical properties of the GARCH model also started to emerge. My 1994 *Handbook of Econometrics* chapter coauthored with Rob Engle and Dan Nelson provides a survey of that early literature (Bollerslev, Engle, and Nelson (1994)). Some of the key developments summarized therein include quasi-maximum likelihood and robust estimation of GARCH models, specification testing and model selection with GARCH models, stationarity and ergodicity properties of GARCH models, temporal aggregation of GARCH models, diffusion approximations of GARCH models, the use of GARCH models as filters and forecasters, along with multivariate extensions of the GARCH model. All of these developments clearly helped to put GARCH models on firmer theoretical grounds.

I personally remember finding the work by Dan Nelson on the connection between discrete-time GARCH models and continuous-time diffusions to be particularly innovative and stimulating, and in retrospect arguably also years ahead of its time (e.g., Nelson (1990, 1992)). As forcefully demonstrated by Dan, even formally misspecified GARCH models can under the "right" asymptotic scheme still generate consistent (for increasingly finer sampled observations) nonparametric volatility estimates, as long as the underlying true latent process is sufficiently well-approximated by a continuous-time diffusion. Dan sadly passed away in 1995 at the early age of 36 (Bollerslev and Rossi (1995) provides a brief summary of some of Dan's seminal works), and one can only imagine the scope and depth of the theoretical contributions that he might have made to the burgeoning realized volatility literature that emerged years later (more on this below).

The growing empirical use of the GARCH model combined with the theoretical foundational work discussed above in turn spurred somewhat of an "arms race" and the formulation of ever more complex GARCH models designed to account for specific features. Along with that race also came a proliferation of acronyms associated with the various models, including models specifically designed to capture the aforementioned "leverage effect" and asymmetry in the way in which the conditional variance responds to positive and negative innovations (e.g., AGARCH, Aug-GARCH, APARCH, EGARCH, GJR-GARCH, GQARCH, HGARCH, NGARCH, TGARCH, VGARCH), models designed to

Figure 1: **Select GARCH Acronyms**

better describe longer-run dynamic dependencies (e.g., CGARCH, FIGARCH, HARCH, HYGARCH, IGARCH, LMGARCH), models based on non-normally distributed innovations (e.g., ARCD, EVT-GARCH, GARCHS, GARJI, GED-GARCH, SGARCH, SPARCH, t-GARCH), multivariate models (e.g., BEKK-GARCH, CCC-GARCH, DCC-GARCH, Flex-GARCH, GO-GARCH, MGARCH, OGARCH, PC-GARCH), and continuous-time models (e.g., COGARCH, ECOGARCH, SQR-GARCH), to name but a few. My chapter in the book *Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle* (Bollerslev (2010)) provides a guide to this perplexing “alphabet soup” of names. To be clear, even though I am personally responsible for some of these acronyms, that is not to say that I ever endorsed the practice of associating each and every variation of the basic GARCH model with its own unique name. Fortunately, this practice also seems to have ceased more recently, and I don’t believe that there are too many new names to add to that earlier 2010 glossary.

This long list of specialized GARCH models notwithstanding, the simple GARCH(1,1) model defined in the equation above has proven surprisingly difficult to convincingly beat on an out-of-sample basis. Indeed, more formal tests for superior predictive ability often find that the GARCH(1,1) model performs on par with, or sometimes even better, than many of these more complicated models (see, e.g., Hansen and Lunde (2005)). As such, even though the GARCH(1,1) model is obviously not the *true* model, it remains the *go-to* volatility model for many applied researchers, not just in finance but also more generally.

Along these lines, while the vast majority of empirical applications of GARCH to date have been in the area of finance, there is a more recent growing interest in macroeconomics concerning the role played by time-varying risk and uncertainty for explaining business cycle fluctuations (recall, the first published applications of the ARCH and GARCH models also pertained to inflationary uncertainty). In addition, GARCH models have also found wide-ranging applications in many other areas of research outside economics, where different notions of "risk" are similarly important, including biomedical science (e.g. drug effectiveness), climatology (e.g., rainfall and temperature), earth science (e.g. earthquakes), hydrology (e.g., drought conditions), medicine (e.g., mortality rates), neuroscience (e.g., ultrasound imaging), operations research (e.g., network capacity and machine failures), and political science (e.g., electorial outcomes), to name some.

3. Beyond GARCH

Despite the widespread empirical use of GARCH models in the late 1990s, many financial researchers remained skeptical about the model and its ability to accurately capture volatility clustering. Much of this skepticism was rooted in the results from Mincer-Zarnowitz type forecast evaluation regressions, in which the ex-post squared returns, interpreted as a proxy for the true volatility, were regressed on the ex-ante forecasts from GARCH models. Although the estimated intercept and slope coefficients from such regressions typically didn't deviate significantly from zero and unity, respectively, as would be implied by unbiased GARCH variance forecasts, the R^2s were typically very low, and for daily horizons in particular close to zero, suggesting to many that GARCH models didn't actually do a very good job in terms of explaining the temporal variation in the true volatility.

My 1998 paper with my longtime friend and frequent coauthor Torben Andersen (Andersen and Bollerslev (1998)) tried to rectify this misconception and convince the GARCH skeptics. Intuitively, while daily squared returns provide (approximately) unbiased estimates of the true daily latent volatilities, the estimates are fraught with measurement error. Instead, building on the earlier continuous-time representations developed in the work by Dan Nelson, much more accurate (and under ideal conditions, for increasing sampling frequencies also formally consistent) ex-post measurements of the true latent daily integrated volatilities may be constructed by summing squared intraday returns. Correspondingly, Mincer-Zarnowitz type regressions based on regressing these more accurate ex-post integrated volatility measures on GARCH-type forecasts also tell a very different story, with R^2s as high as fifty percent at the daily horizon. In short, a classical errors-in-variables faux pas.[2]

[2]General model-free and easy-to-implement adjustment procedures to account for the errors in the

Perhaps ironically, while the Andersen and Bollerslev (1998) paper was originally construed as a defense of GARCH, it ended up providing somewhat of a pivot away from the use of GARCH models (and other parametric volatility models) to nonparametric procedures. The seeds for that pivot were planted at a 1997 NBER Asset Pricing Program meeting, where Frank Diebold served as the discussant of the "skeptics" paper. Paraphrasing part of Frank's discussion: "if you have these more accurate ex-post volatility measures, why not use them in place of the lagged squared returns in the formulation of even better GARCH models?" This idea was also subsequently explored empirically by Engle (2002) among many others. However, Frank also further commented: "if you really think that you have an almost noiseless measure of the true volatility, why not simply model and forecast that measure using standard ARMA type models?" In a sequence of joint papers with Frank and Torben, and Frank's Ph.D. student Paul Labys (Andersen, Bollerslev, Diebold, and Labys (2000, 2001, 2003)), we laid the theoretical foundation for doing just that and went on to empirically demonstrate the superiority of this new approach to volatility forecasting (see also the related contemporaneous work by Barndorff-Nielsen and Shephard (2002)). Other early applications of the realized volatility idea (e.g., Andersen, Bollerslev, Diebold, and Ebens (2001)) also helped bring into sharper focus various distributional features, including among others that logarithmic volatilities and returns stadardized by their volatilities both appear close to unconditionally Gaussian distributed, and that the dynamic dependencies in volatilities appear to be well described by long-memory type processes.

Importantly, the realized volatility concept also helped to formally embed the econometric analyses of return volatility within the vast probability and statistics literatures on Itô semimartingales and the theory of quadratic variation, thereby affording a more rigorous justification for the approach based on no-arbitrage assumptions and theoretical in-fill asymptotic arguments. The extensive realized volatility literature that has emerged over the past two decades, has in turn benefitted from simultaneous developments by applied econometricians and financial economists making use of the new tools for improved risk management and asset pricing predictions, financial engineers and mathematicians developing new pricing formulas based on nonparametric realized volatility type measures, and statisticians and probabilists further refining the measures and deepening the theory related to the separation of jump and diffusive components and semimartingales contaminated by "noise" (the introductory chapter to the collection of seminal volatility papers in Andersen and Bollerslev (2018) provides a more thorough summary of these developments). My recent SoFiE presidential address (Bollerslev (2022)) discusses some

volatility measurements were subsequently presented in Andersen, Bollerslev, and Meddahi (2005).

of the directions that I personally consider to be especially promising in this still very active area of research.

4. Some Final Thoughts

The field of financial econometrics has had a glorious run during the life span of the *Journal*. Starting from nascent in the 1970s, slowly coming into existence in the 1980s, truly taking off in the 1990s, in turn accounting for a substantial fraction of the papers published in the *Journal* during the 2000s. Although this trend may now have slowed down or even reversed, there are still routinely many first-rate papers in financial econometrics being published in the *Journal*. Much of this growth in the field of financial econometrics was fueled by research related to time-varying volatility, GARCH models included. Meanwhile, research on GARCH per se has arguably long since reached diminishing returns to scale, being supplanted by analyses related to realized volatility type measures and related procedures. At the same time, the availability of reliable high-frequency data, or the lack thereof, invariably restricts the practical applicability of such procedures to actively traded financial assets. The realized volatility concept also doesn't easily lend itself to applications outside finance and the modeling and forecasting of "risks" in other situations, where the idea of increasingly finer sampled observations over fixed time intervals is difficult, if not impossible, to mimic in practice. As such, having been validated as the near-optimal tool for recovering volatility from more coarsely sampled data, GARCH models will likely remain in empirical use in a many different areas for years to come, and I hope keep up the fortuitous citation count for my 1986 *Journal* paper.

References

Andersen, T. G., and T. Bollerslev (1998): "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39(4), 885–905.

——— (2018): "Volatility: Introduction," in *Volatility*, ed. by T. G. Andersen, and T. Bollerslev, pp. xiiv–xliii. Edward Elgar Publishing, Cheltenham, UK.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001): "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, 61(1), 43–76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2000): "Great Realizations," *Risk*, 13, 105–108.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001): "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96(453), 42–55.

——— (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71(2), 579–625.

Andersen, T. G., T. Bollerslev, and N. Meddahi (2005): "Correcting the Errors: Volatility Forecast Evaluation Using High-Frequency Data and Realized Volatilities," *Econometrica*, 73(1), 279–296.

Barndorff-Nielsen, O. E., and N. Shephard (2002): "Econometric Analysis of Realised Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society: Series B*, 64(2), 253–280.

Black, F. (1976): "Studies of Stock Price Volatility Changes," in *Proceedings of the 1976 Meetings of the American Statistical Association*, pp. 177–181.

Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31(3), 307–327.

——— (1987): "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics Statistics*, 69(3), 542–547.

——— (2010): "Glossary to ARCH (GARCH)," in *Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle*, ed. by T. Bollerslev, J. R. Russell, and M. Watson, chap. 8, pp. 137–163. Oxford University Press, Oxford, UK.

——— (2022): "Realized Semi(Co)variation: Signs That All Volatilities are Not Created Equal," *Journal of Financial Econometrics*, 20(2), 219–252.

Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992): "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 52(1), 5–59.

Bollerslev, T., R. F. Engle, and D. B. Nelson (1994): "ARCH Models," in *Handbook of Econometrics, Volume 4*, ed. by R. F. Engle, and D. McFadden, chap. 49, pp. 2959–3038. Elsevier, Amsterdam.

Bollerslev, T., and P. E. Rossi (1995): "Dan Nelson Remembered," *Journal of Business and Economic Statistics*, 13(2), 361–364.

Diebold, F. X. (1986): "ARCH models of exchange rate fluctuations," PhD Dissertation, University of Pennsylvania.

Engle, R. (2002): "New Frontiers for ARCH Models," *Journal of Applied Econometrics*, 17(5), 425–446.

Engle, R. F. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50(4), 1987–1007.

Engle, R. F., and T. Bollerslev (1986): "Modelling the Persistence of Conditional

Variances," *Econometric Reviews*, 5(1), 1–50.

Engle, R. F., D. M. Lilien, and R. P. Robins (1987): "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model," *Econometrica*, 55(2), 391–407.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987): "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19(1), 3–29.

Hansen, P. R., and A. Lunde (2005): "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?," *Journal of Applied Econometrics*, 20(7), 873–889.

Mandelbrot, B. (1963): "The Variation of Certain Speculative Prices," *Journal of Business*, 36(4), 394–419.

Nelson, D. B. (1990): "ARCH Models as Diffusion Approximations," *Journal of Econometrics*, 45(1), 7–38.

——— (1992): "Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model," *Journal of Econometrics*, 52(1), 61–90.

Rosenberg, B. (1972): "The Behavior of Random Variables with Nonstationary Variance and the Distribution of Security Prices," Working Paper No.11, University of California Berkeley, Institute of Business and Economic Research.

Shephard, N. (2005): *Stochastic Volatility: Selective Readings*. Oxford University Press, Oxford, UK.