

Response to **Introduction**:

1. *This paper presents some ideas on how to evaluate models of volatility (conditional standard deviation) or “variability” (the conditional mean of the squared return). The author takes very seriously the fact that we cannot directly observe the conditional variance..* These are the first two sentences of referee report 1, and they suggest referee 1 has fundamentally misunderstood the way I define key concepts of the paper. Large portions of the essay is devoted to these concepts, and interestingly referee 1 even seems to acknowledge the difference in our definitions: “. . . I think I also cover an important difference in how the author and I view ‘volatility’ or ‘variability’” (p. 1). So I am quite puzzled by the fact that referee 1 consistently mis-portrays me (deliberately?) throughout the report. In particular, the suggestion that I equate variability with the *conditional mean* of squared return, and the statement that “I take seriously the fact that we cannot observe directly the conditional variance” are misleading. Already in the introduction (third paragraph, p. 2) I say that “return variability is defined as squared return”—not as the conditional mean of squared return—and that it is observable. In contrast, the conditional expectation (or mean) of variability is a prediction or model of variability and at the heart of the simulations in section 3. Similarly, volatility is a model of the (unsigned) error term e_t (the portion left unexplained by the mean specification) and so the “unobservability” of volatility is virtually irrelevant for the objectives of my paper.

The starting point of my paper is the view that it is inappropriate to evaluate explanatory models of financial variability by comparing their forecasts against high-frequency estimates based on continuous time theory. It is not the first time that I with co-authors express similar views and ideas, see Bauwens et al. (2006), Sucarrat (2006), Sucarrat (2007), Bauwens and Sucarrat (2008). In the course of this period we have received numerous comments and responses at conferences, from referees, from editors and from other colleagues. The reactions we have been met with range from “you are obviously right, why do you even bother expressing the view?” on the one hand, to “you are obviously wrong” on the other.¹ From the referee report it is crystal clear that referee 1 belongs to the group associated with the second view. My experience is that it is extremely difficult to convince people that holds the second view (possibly the same holds the other way), so the only thing I can hope for is to identify why we disagree. My own hypothesis on why we disagree is twofold. First, I do not believe people of the second view have fully understood the ideas and implications of econometric reduction theory. The whole of section 2 is devoted to the reduction theoretical implications in the current context, and the fact that referee 1 does

¹When people of the first view realise how widespread the second view is, then they immediately become more sympathetic to our venture.

not address any of these issues I interpret in support of my hypothesis. Second, I do not believe people of the second view are aware of the philosophical issues—nor the associated economic consequences produced by continuous time models, in particular when continuous time models are applied in contexts that gives rise to temporal aggregation issues. I think it is fair to say that most philosophers of time, most philosophers of mathematics, most philosophers of language, most philosophers of mind and most social philosophers would agree that mathematics in general and real numbers in particular are not capable of accurately representing neither objective nor subjective continuous time in social analysis.² For many purposes the representation error is unimportant. When temporal aggregation issues are involved, however, as is the case in explanatory financial variability modelling, it can lead to seriously erroneous conclusions. This is exactly why we need a framework and procedures to evaluate continuous time and discrete time models against each other without treating either as more fundamental. For simplicity I deliberately keep philosophical jargon and references in the paper to a minimum, and focus instead on the economic aspects. For example, three of the five reasons I list in the second paragraph of the introduction (pp. 1-2) can be given deeper, philosophical justifications. However, these reasons, which are further discussed in section 2 (most explicitly in the last paragraph of subsection 2.2 on pp. 5-6), are somewhat arrogantly dismissed by the referee: “. . . by the end of the paper I was at a loss to name the issues that the author is so concerned about” (p. 3). Also, the only place I allude to some philosophically related concepts, namely at the end of subsection 2.2 on page 6, referee 1 finds that the argument does not appear “to be relevant” (p. 4). So, in my view, there is absolutely no evidence that referee 1 has understood the underlying philosophical issues or their economic consequences, and there is even less evidence that s/he has tried to.

2. Some of the choices made and conclusions drawn by the author suggest to me that his framework for thinking about this problem is not ideal. Rather than presenting simulation evidence and hoping that the conclusions apply more generally, theory from the existing literature shows that some of the conclusions do hold, whilst others do not. Two papers in particular are relevant to the author's study, Hansen and Lunde (2006) and Patton (2006); the author cites both of these papers, but he either does not agree with their results or does not see how they relate to this problem. First, I do not cite “Patton (2006)” but Patton (2007), which is an updated version of “Patton (2006)”. Second, both Patton and Hansen and Lunde are interested in predicting volatility (as if volatility is given objectively and not determined by the modeller) and not variability the way I define it, so it is not evident that their arguments and conclusions apply to the current context. This, I believe, I state

²The root of the problem comes from the so-called principle of extensionality, that is, the axiom that two sets (or elements of a sets) are equal if and only if they are the same. This axiom is needed by mathematics in order to avoid contradictions caused by self-referential paradoxes, say, the Liar's Paradox, and its consequence is essentially that mathematics is “discrete” and that the notion of continuity has to be approximated. Typically the axiom of infinity plays an important role in such approximations, and a common example of a mathematical structure that is used in order to approximate the idea of continuity are subsets of the real numbers. See Tiles (1989) for an introduction to the philosophy of mathematics that takes continuity issues as its organising theme.

quite clearly in the introduction to section 3 (p. 7):

“A substantive number of studies have contributed directly or indirectly to the understanding of these questions within the paradigm of volatility being *given* (as opposed to *determined* by the explanatory information included), see amongst others...Hansen and Lunde (2005, 2006), and Patton (2007). Here, by contrast, the aim is to shed light on variability model evaluation within the paradigm of volatility *not* being given, but a result of the information included in the mean and variance specifications. As a consequence, loss will be conceived in terms of the forecast error of squared returns, since squared returns is a measure of the total variation in return or price variability...”

Response to **Comments**:

1. *It is not generally true that MAE is better than MSE in small samples.* The argument of referee 1 is based on the assumption that the magnitude I am interesting in forecasting is $E(r_t^2|\mathcal{I}_t)$. As mentioned above (in point 1 in the introduction part), this is not the case. The magnitude I am interesting in forecasting is r_t^2 , so the argument of referee 1 is erroneous. Also, referee 1 incorrectly claims that my simulations shows that MAE provides incorrect rankings as $T \rightarrow \infty$:

... MAE will *not* provide correct rankings as $T \rightarrow \infty$... and this is also what is found by the author in his simulations (p. 2)

Contrary to what referee claims, my simulations (see tables 5-8) suggest MAE provides correct rankings as $T \rightarrow \infty$ for model 1, the simulations are inconclusive for models 2 and 3 (the difference in percentage point going from $T = 500$ to $T = 1000$ is 0 and 1, respectively, and could therefore be due to simulation error), whereas for model 4 the non-monotonic evolution in probabilities from $T = 500$ to $T = 1000$ could indeed be interpreted as MAE not providing correct ranking as $T \rightarrow \infty$. However, the results are more likely due to incongruence issues (see point 5 in the specific remarks part in my response to referee 2) rather than the explanation suggested by referee 1. Finally, one may at any rate argue that in explanatory modelling the case $T \rightarrow \infty$ is not necessarily of interest. As I write in the introduction (p. 2):

“Appropriate understanding in finite samples is crucial since explanatory data typically is available at lower frequencies only...”

This is reflected in my conclusions (subsection 3.6 on p. 17 and section 5 on p. 23) where I suggest that MAE can be more useful than MSE in *small* samples, which in my simulations means up to about 100 observations.

2. *Kurtosis performs terribly and should be excluded from the paper; include QLIKE measures instead.* It is true that Kurtosis does not perform very well and this might provide

reason for not including in. Also, I agree with the explanation given by referee 1 for why Kurtosis performs badly (but see also point 1 in the specific remarks part in my response to referee 2). Nevertheless, I do not share the general pessimism of referee 1. I think Kurtosis might perform quite well in selecting between parsimonious models obtained through GETS simplification search of a general unrestricted variance-congruent model. But, of course, that remains to be investigated. The problem with QLIKE loss-functions of the type proposed by referee 1 is that, strictly speaking, their validity is limited to a certain distribution or a certain class of distribution. Referee 1 says the suggested QLIKE loss-function is likely to work well for non-normal distributions, but no argument nor references are provided. One might guess that referee 1 has Patton's (2007) discussion of QLIKE distributions in mind. But, as mentioned above, Patton's argument is conducted in a fundamentally different set-up. So it is not straightforward that Patton's conclusions apply in the current context.

3. *Conditional variances depend on the DGP and not a model. Obtaining an estimate of the conditional variance usually requires a model, but a direct estimate is not always needed: Hansen and Lunde (2006b)³ and Patton (2006) present methods for evaluating and comparing volatility models that avoid having to directly model the conditional variance.* First, as stated numerous times already, my aim is to compare models of variability and not models of volatility. In my setup volatility is one out of the two components that make up conditional variability, that is, a prediction of variability, and so the works by Hansen, Lunde and Patton are not directly relevant. Second, nowhere do I state that the conditional variance does not depend on the DGP, nor that conditional variances cannot be discussed without writing down a specification of the conditional variance. What I do say, however, is that the specification(/model) of the conditional variance, that is, the choice of conditioning information and *how* the conditioning information is used, determines the value and distributional properties of the standardised residual. This is simply an implication from reduction-theory and has important implications in explanatory modelling of financial variability. These issues are *not* accommodated by the work of Hansen, Lunde and Patton, nor in the works by Andersen and Bollerslev with co-authors, etc. For further discussion on the notion of a DGP, see my response to comment 1 (point 2 in the comments part).

4. *The chosen data for the empirical illustration are inappropriate because there is not much volatility clustering in the returns, and this might affect the conclusions.* The statement that the data are not well chosen I find surprising. In explanatory financial variability modelling one would typically work with daily, weekly or monthly data, or data at even lower frequencies. It is well-known that financial returns at these frequencies exhibit little or no volatility persistence. So, contrary to the view of referee 1, a series with little volatility persistence seems to me to be a very adequate choice for the illustration. As to whether the conclusions of the empirical *ex post* and *ex ante* evaluation of the models would change in the presence of more volatility persistence, this is irrelevant since it would

³I assume referee 1 is referring to Hansen and Lunde (2006).

depend on the empirical accuracy of each model. As stated on p. 19, several alternative specifications were explored in the derivation of ECON. These alternative specifications contained a GARCH structure in the variance equation, and variables that sometimes account for volatility persistence (order flow, volume amongst others). So my guess is that, no, the conclusions of the empirical evaluation would not change in the presence of more volatility persistence in returns, because then the volatility persistence would have been adequately modelled by ECON. But this is at any rate a counterfactual, empirical question that is not relevant for the theoretical considerations of the paper.

5. *By the end of the paper I was at loss to name the issues that the author is so concerned about.* As stated previously, I believe this is partly due to the fact that referee 1 misportrays me throughout, and partly because referee 1 does not bother with the economic and reduction-theoretic arguments given in the introduction and in section 2. It is true that there is a large body of continuous time literature that addresses the effects of microstructure issues. However, the weakness with this body of literature is analogous to the weakness of microfoundationalism: Empirically uncheckable (and unrealistic) continuous time structures are posited and effectively assumed true, and then their “microstructure effect fixes” are derived without any possibility to check all the assumptions contained in the continuous time structure. Indeed, since any empirical check of continuous time structures *must* be undertaken in discrete time, and since—philosophically—modelling social processes in mathematical continuous time is questionable, it only underlines the objective of the paper: We need a framework and procedures for evaluating continuous time and discrete time estimates against each other without assuming that either is more fundamental.

Madrid, 9 July 2008
Genaro Sucarrat

References

- Bauwens, L., D. Rime, and G. Sucarrat (2006). Exchange Rate Volatility and the Mixture of Distribution Hypothesis. *Empirical Economics* 30, 889–911.
- Bauwens, L. and G. Sucarrat (2008). General to Specific Modelling of Exchange Rate Volatility: A Forecast Evaluation. Forthcoming in the *International Journal of Forecasting*. Working Paper version available as <http://www.eco.uc3m.es/sucarrat/research/gets.pdf>.
- 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, 873–889.
- Hansen, P. R. and A. Lunde (2006). Consistent ranking of volatility models. *Journal of Econometrics* 131, 97–121.

Patton, A. J. (2007). Volatility Forecast Evaluation and Comparison Using Imperfect Volatility Proxies. Available as http://www.economics.ox.ac.uk/members/andrew.patton/patton_robust_aug07.pdf.

Sucarrat, G. (2006). *Essays in the Study and Modelling of Exchange Rate Volatility*. Ph. D. thesis, Department of Economics, Université catholique de Louvain, Louvain la Neuve, Belgium. Thesis Number 508/2006. Available as http://edoc.bib.ucl.ac.be:81/ETD-db/collection/available/BelnUcetd-09152006-141220/unrestricted/these_sucarrat.pdf.

Sucarrat, G. (2007). Econometric Reduction Theory and Philosophy: The First Stage in Hendry's Reduction Theory Revisited. Available as <http://www.eco.uc3m.es/sucarrat/research/hendry.pdf>.

Tiles, M. (1989). *The Philosophy of Set Theory*. New York: Dover Publications, Inc.