Code Sharing Policy – Frequently Asked Questions

What is the purpose of this policy?
Reproducability of research findings is a basic principle of science and a pre-requisite for replicability. Many researchers have undoubtedly had the experience of trying, unsuccessfully or with great difficulty, to reproduce the results of a published paper. In most cases, the authors of the published paper cannot be blamed for these difficulties. Many empirical and computational research projects these days are complex. Files are merged from many data sets, data manipulation is extensive, and empirical techniques have become more and more sophisticated. It has become virtually impossible to catalog all the precise data definitions and specifications in the paper or an online appendix ("Which combinations of Computstat items go into this variable?"; "How is the clustering done?"; "How do they treat stocks in their delisting month?", etc.). This is where code sharing can help.

What are the benefits for authors?
We strive for the *JF* to be a journal that meets the highest standards of quality, reproducibility and replicability. By agreeing to share the code, authors submitting to the *JF* signal that they, too, are committed to these high standards. With code transparency, research findings will be more credible and they will, ultimately, have greater impact.

Will authors have to respond to queries about the code?
Our code disclosure policy is clear that authors cannot be expected to provide support and assist with queries regarding the code. At the same time, authors can facilitate the use of their code by providing proper documentation. Just like clear exposition and writing in the published article, proper documentation of code can help authors enhance the reproducibility of their results and the impact of their research.

Should the code reproduce the research findings based on the pseudo-data?
It is not required that the code produces meaningful results based on the pseudo-data set. But pseudo-data should illustrate the format of the data that are read by the programs (e.g., dimension of the data, numbers or strings, etc.) to help a user understand the code. The pseudo-data should not include any data that is protected by copyright or confidentiality agreements.

Why not require posting of data as well?
This is a good question, but one that raises a host of difficult issues because much of the data used in finance research is proprietary and licensed. As the code sharing policy states, authors are welcome to post data along with code as long as doing so is not in violation of copyright or confidentiality agreements.