

VU Research Portal

Panta rhei, measurement and discovery of change in financial markets

Zamojski, M.J.

2017

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Zamojski, M. J. (2017). *Panta rhei, measurement and discovery of change in financial markets*. Tinbergen Institute.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Chapter 1

Introduction

‘[People] want to believe that somebody really knows. A world in which nobody really knows can be frightening.’

— Alfred Cowles, as quoted by *Bernstein (1993)*

Heraclitus of Ephesus (535–475 BC), a pre-Socratic Greek philosopher, is widely believed to have said that ‘ever-newer waters flow on those who step into the same rivers’. The ever-changing nature of the world is not limited to the physical layer alone as the literal reading would suggest. It also reflects the evolution of our own convictions—both as individuals and as societies. The ability to adapt and change one’s behaviour is an intrinsic quality of all successful living organisms, embedded in us through millennia of evolution. Our adaptability to the changing environment underpins the so-called [Lucas \(1976\)](#) critique. [Lucas](#) argues that static econometric models run a risk of being uninformative if the modelled quantity is affected by actions of economic agents who dynamically change their behaviour in response to (or in anticipation of) market, regulatory, or technological changes. The question is not whether time-variation occurs, but whether (i) we have the data necessary to show evidence of the changes, (ii) tools to surface them, and (iii) if these changes are economically and statistically significant.

In this spirit, this thesis consists of four self-contained essays in finance and financial econometrics. Each of the chapters which follow this introduction attempts to contribute

towards a better understanding of sources and consequences (or profit opportunities) of the continuous changes in the financial markets. The nature of the changes differs throughout the thesis. Chapter 2 looks at time-varying interest rates and documents an equally time-varying ability of simple, mis-specified models to fit data generated by possibly complicated processes. Chapter 3 provides tools to investigate, among others, stability of the relative risk aversion. In Chapter 4, I propose a method to measure statistical significance of such changes. Finally Chapter 5 documents the evolution of a whole segment of the financial industry and the various economic motives which have been driving the changes.

Score-driven dynamic Nelson-Siegel models

Chapter 2, based on [Koopman, Lucas, and Zamojski \(2016\)](#), is devoted to modelling the term-structure of interest rates. The chapter starts with a detailed discussion of observation- and parameter-driven methods. In this way, it also sets the stage for Chapters 3 and 4 where I work further with observation-driven methods. The distinction between the approaches is originally due to [Cox \(1981\)](#) who defines the two ways a time-varying parameter can be included in an econometric model. In fact, the names of the approaches identify the source of randomness which ‘drives’ the dynamic parameter, i.e., what determines the changes in parameter values between two adjacent points in time. In parameter-driven models, both observations and latent variables can evolve independently. The latter are thus ‘parameter’-driven as it is necessary to impose additional distributional assumptions on them as well. A typical example of this class of models is the stochastic volatility model for which [Shephard \(2005\)](#) provides an overview. In contrast, there is only a single source of error in observation-driven models. The dynamic parameters are tied to the distribution which drives the observations. A typical example of an observation-driven model is the family of generalised autoregressive conditional heteroskedasticity (GARCH) models originated by [Engle \(1982\)](#) and [Bollerslev \(1986\)](#). Among other things, the two approaches take a clear stance on what kind of behaviour is possible for the dynamic parameter. For instance, a parameter-driven

volatility can remain low even after a relatively long series of large price changes while such behaviour would be highly unlikely in the observation-driven framework. An advantage of observation-driven models is their computational simplicity. This is because values of dynamic parameters are perfectly predictable one-step-ahead given past information which makes evaluation of a chosen criterion function (e.g., likelihood function or inner product of moment conditions) straightforward.

In Chapter 2, we look at the performance of parameter-driven models and score-driven models in explaining the evolution of the yield curve in time. First proposed by Creal, Koopman, and Lucas (2013) and Harvey (2013), score-driven models are a subclass of the observation-driven family. The dynamic parameters are driven by a scaled first derivative of the local criterion function. The chapter focuses on the parsimonious yield curve model of Nelson and Siegel (1987) which is extended within a dynamic setting by Diebold and Li (2006) and Diebold, Rudebusch, and Aruoba (2006). We investigate in-sample and out-of-sample performance for two data sets and for simulated data. In a univariate setting, score-driven models were shown to offer competitive performance to parameter-driven models in terms of in-sample fit and quality of out-of-sample forecasts but at a lower computational cost (Koopman, Lucas, and Scharth, 2016). We investigate whether this performance and the related advantages extend to more general and higher-dimensional models.

Although conceptually simple, the dynamic Nelson-Siegel models we consider in this chapter involve time-varying parameters in means, variances, and covariances. Furthermore, apart from having disturbances from a multivariate Gaussian distribution, we also allow disturbances to come from fat-tailed distributions, namely from the multivariate Student's t distribution. The computational simplicity of observation-driven models allows us to implement a wide range of models. We are able to allow for specifications whose counterparts in the parameter-driven framework are, in terms of computing power, too costly to consider. Based on an extensive Monte Carlo study, we show that in multivariate settings the advantages of score-driven models can even be more pro-

nounced than in the univariate setting. We find that allowing for heteroskedasticity in pricing errors produces the highest incremental gain in performance. Furthermore, the inclusion of more features in the model appears to benefit mostly the short maturity yields. Finally, Chapter 2 proposes a novel way of obtaining multi-period forecasts for both the dynamic factors and the interest rates. The method uses previous, normalised score adjustments. This may improve performance when the model is mis-specified. As the forecast horizon is extended, we observe a gradual decay in relative performance of the simple models with the short-term maturities being, again, especially affected.

Generalised autoregressive Method of Moments

Chapter 3, based on Creal, Koopman, Lucas, and Zamojski (2015), continues with the score-driven models and generalises the approach to the method of moments framework. The chapter provides a unified framework for modelling parameter instability in a context where the model and its parameters are only specified through (conditional) moment conditions. The new estimation method extends the Generalised Method of Moments (GMM) of Hansen (1982) to settings where a subset of the parameters vary over time with unknown dynamics. Unlike other attempts to extend GMM estimators in this manner, the GaMM does not require simulation. Estimation is similarly straightforward, easy to implement, and computationally inexpensive as the likelihood based score-driven methods in Chapter 2. Furthermore, because the GaMM falls entirely within the standard set-up of GMM estimation, inference can be based on the asymptotic normality and consistency results which were previously established for these estimators. We label the approach as the Generalised autoregressive Method of Moments (GaMM).

We include three different example applications of the GaMM that are increasing in complexity. The models we consider highlight settings where traditional approaches are either extremely difficult to implement or no alternative technique is readily available. We start with the estimation of stable distributions with time-varying scale parameters. In this case, a closed-form expression for the density function is generally not known

which renders likelihood based estimation challenging. In contrast, GaMM can be directly applied by leveraging the fact that the characteristic function—which is known for this family of distributions—can be rewritten as a set of moment conditions. Results of an extensive Monte Carlo study indicate that GaMM is able to successfully track the unknown path of the time-varying parameter, despite being partially mis-specified. We also apply the model to a real-world time-series of daily returns. We use S&P 500 data from the period 1988–2013 and we benchmark our results with two likelihood based models—GARCH and generalised autoregressive score model with Student’s t disturbances (t-GAS) of [Creal, Koopman, and Lucas \(2011\)](#). The filtered path of the time-varying scale shares similar dynamic properties with the paths in the benchmark models. It is, however, considerably smoother as the GaMM is more cautious in ascribing the realisation of a large positive or negative return to an increase in the time-varying scale parameter. The smoothness is, in part, also due to the low, estimated value of the tail index that suggests returns are drawn from heavy-tailed conditional distributions even after correcting for changes in scale.

For the second application of the GaMM, we consider a standard consumption-based asset pricing model in which we allow the relative risk aversion to change over time. Without the added flexibility to accommodate changes in this parameter, results of many studies suggest that estimates of a static relative risk aversion parameter are typically too high when compared to results obtained from experimental data. Furthermore, the static model fails in explaining the cross-sectional variation in stock returns, see, e.g., [Ludvigson \(2011\)](#) for a recent summary of the literature. The poor performance of the static model led to other modelling approaches, many of which replace the constant relative risk aversion (CRRA) utility function in the standard model with more complicated functions. A popular choice is to allow for habit formation which, on the one hand, adds an additional source of variation to the stochastic discount factor, see, e.g., [Campbell and Cochrane \(1999\)](#). On the other hand, it explicitly allows the relative risk aversion to be a time-varying quantity. GaMM allows us to easily implement such time variation also in the standard utility function. Similarly to the previous example,

we provide simulation results which testify to the GaMM's ability to filter out the path of the time-varying parameter. We also bring the model to U.S. data from the period 1947–2015. There, we find that the estimates of the discount factor and the relative risk aversion are plausible and consistent with both common sense and experimental results. Furthermore, the results suggest that the relative risk aversion has indeed been changing in the U.S. There are two components that contribute to the dynamics. First, a short-term cycle which appears to follow the business cycle. More interestingly, we see a long-term, decreasing trend in the risk aversion.

The third example application of GaMM is presented in Appendix 3.D to Chapter 3. We look at the estimation of linear regression models with time-varying coefficients when a subset of the covariates are endogenous. As a benchmark model, we use an ordinary state space model which we estimate using standard Kalman Filter methods. In a static context, this would be close to comparing performance between an instrumental variables (IV) estimator and an ordinary least squares (OLS) estimator. We find that the trade-off between the Kalman Filter and the GaMM approach seems to mirror the differences between OLS and IV estimation. The Kalman Filter produces results which are biased but with low sampling variability, whereas paths estimated by GaMM appear to be unbiased, but at the cost of a higher sampling variance.

Finally, even though GaMM performs well in the three examples, it is possible to improve the results in scenarios where the true time-varying parameter exhibits sudden, structural changes. To that end, we introduce a penalty to the criterion function. The functional form of the penalty is reminiscent of the least absolute shrinkage and selection operator (LASSO) of Tibshirani (1996). In contrast to that method, we do not apply the penalty to static parameters but to the scores of the criterion function which drive the dynamic parameter. This approach favours dynamics that adjust the value of the time-varying parameter to a lesser extent at every point in time, not only on average. On the one hand, penalisation induces smoother paths and prevents over-fitting. On the other hand, the estimated path may react more quickly after large changes in the

value of the parameter. The penalisation seems to lead to estimates of the time-varying paths that have much lower root mean square error (RMSE) at a cost of a slight bias.

Filtering With Confidence

The issue of modelling changes in time-varying parameters without imposing rigid distributional assumptions, which is addressed in Chapter 3, is not the only challenge in the literature. A big disadvantage of the observation-driven approach is the lack of precision measures for the time-varying parameter. The perfect predictability of the time-varying parameters' values one period ahead, makes it seemingly not relevant to think about uncertainty around the path. However, an observation-driven model is unlikely to be the true data-generating process—at least to the extent of correctly specifying the transition dynamics. Instead, one could look at the method as filter, i.e., a clever way of approximating and extracting the unknown path. In this case, multiple sources of uncertainty are involved. Chapter 4, based on [Zamojski \(2016\)](#), is the first to propose a robust method of computing in-sample confidence bands for time-varying parameters estimated with misspecified observation-driven models. As an example of this class, I look at the family of GARCH models which are used to estimate time-varying variances and covariances. I propose a novel bootstrap procedure and a new moving-window resampler which together generate confidence bands around estimated volatility paths. The procedure is labelled as Local In Time (LITE) bootstrap and resampler.

Due to the assumed mis-specification of the model, the estimated path of the time-varying parameter can be expected to lag behind the true path creating random misfit. Thus, a special attention needs to be given to finding valid replicates for innovations in the bootstrap world. The role of the LITE resampler is to select, at every point t in time, replicating residuals which suffer from a similar misfit. The LITE bootstrap is then able to generate valid bootstrap samples conditional on the data, by explicitly including the estimated path of the time-varying parameter. In contrast, traditional

approaches would discard the estimated path and simulate new paths of the time-varying parameter in every bootstrap sample unconditionally from the observation-driven model, effectively treating it as a correctly specified data-generating process. The LITE is able to generate pseudo time-series which sufficiently recreate the true—unobserved—dependence structure in the data. In this way, the bootstrap world poses the same challenges for the filter as the real world.

The approach accounts for various sources of uncertainty, including parameter and filtering uncertainty. I illustrate the method by applying it to S&P 500 returns. Moreover, I investigate finite sample properties of the confidence bands and their convergence in a range of simulation experiments. I find that the average coverage is close to the nominal level in finite samples and that it converges to the nominal level as the sampling frequency is increased. The confidence bands proposed in Chapter 4 behave remarkably similar to those which could have been obtained in Monte Carlo experiments, i.e., when both the true path of the time-varying parameter and the true conditional densities are known. As in the case of GaMM, the new method is easy to implement and does not significantly increase the computational burden.

Hedge Fund Innovation

One common feature of Chapters 2 to 4 is the focus on *modelling the changing market* environment with time-varying parameter models. In contrast, Chapter 5, based on Siegmann, Stefanova, and Zamojski (2014), is devoted to finding out how institutional investors *react to the changing market* conditions. In particular, we investigate the hedge fund industry which promises high, absolute returns in exchange for hefty fees. Due to light regulation, hedge funds are able to invest in instruments and regions which are not available to other investors. Furthermore, they are able to quickly change exposures in reaction to (or in anticipation of) changing market conditions, see, e.g., Criton and Scaillet (2011) or Patton and Ramadorai (2013). They can be seen as an innovative force in seeking out new profit opportunities and achieving diversification for investors.

In Chapter 5, we are interested in identifying factors that motivate fund managers to innovate. We also look at whether some of the benefits of innovation are shared with investors. To that end, we study first-mover advantages in the hedge fund industry. The benefits of early entry are a typical object of study in the literature on industrial organisation and there is already a strand of literature which focuses on the financial industries, see, e.g., [Tufano \(1989\)](#); [Herrera and Schroth \(2011\)](#); [Lounsbury and Crumley \(2007\)](#); [Makadok \(1998\)](#); [Lopez and Roberts \(2002\)](#). There, the findings are that first-movers are able to acquire a higher market share, but do not necessarily obtain higher margins or higher fees. A common feature in past research is that innovation is identified based on the public information. This is not the case for hedge funds, which are secretive about their strategies. Hedge funds disclose little information to potential investors, enough for the purpose of advertising, but (ideally) not enough for less talented managers to easily copy them and remove their advantage by increasing market efficiency. The data that is available to us identifies the type of assets and instruments hedge funds trade in; sector, geographical, and investment focus; as well as other fund details. In particular, we do not directly observe the innovation or the underlying changes in the market or regulatory environment which prompt the managers to innovate. It is thus necessary to identify the newly created product categories through inference. The high dimensionality and the nature of the data require us to use a custom clustering algorithm, which we call the Fast Binary Clustering. Existing algorithms are either not suitable to binary data or exhibit problems with the high dimensionality.

Our results show that the reported characteristics can be used to infer strategy-related information and suggest that specific first-mover advantages do exist in the hedge fund industry. Early entry is associated with higher excess returns, longer survival, higher incentive fees, and lower management fees compared to funds that arrive later. Managers' motives to innovate seem to be clear—by creating new product categories (strategies) they are able to not only earn higher returns but also to keep a bigger portion of these for themselves. Investors may prefer innovative funds as they offer better performances than funds which imitate established strategies.