Advancements in structural break testing

Valorisation addendum

When investigating motives and patterns of human action and interaction, statistics are an important source of information. Because they are important, statistical results then steer human decision making, both implicitly and explicitly. Statistics are therefore deeply ingrained in human behavior as an interactive component.

Understandably, people are fascinated by the patterns of their own behavior and are determined to understand them. Scientists investigate the configurations of human interaction tirelessly. Our general curiosity for the topic can also explain the popularity of such books as Thinking, Fast and Slow by Kahneman (2011). We recognize the importance of being aware of rules and implications of statistics, of probabilities and heuristics. We acknowledge that many conclusions are better drawn having observed a lot of data and looking for significant effects, by Thinking, Slow. Statisticians are guides for learning from data and navigating common problems. The scientific contributions to the field of statistics can help us better understand ourselves and each other, and help us make decisions more carefully.

Structural break testing is one explicit form of statistical information processing. It is a form of model diagnostics. Models are used to reduce the complexity of an interaction between variables, between human actions or other forces. We observe patterns and may be able to model a process that can explain and predict what we see within limited margins of error. But do we have reason to assume that the same process holds for a full range of observations, or must we admit that there has been a change at some point? Are the patterns that we observe the outcome of a stable model or do we have to consider a structural break in the process?

The error of any model can be reduced by incorporating a structural break, as it increases the adaptability of the model to the given data. If a structural break reduces the model error significantly, the break itself can be interpreted, but the separation of data at the break also ensures that prediction of additional data is based on an appropriately estimated model. Therefore, structural break testing can serve a wide range of purposes in unlimited application areas.

Often, as in this book, structural break tests are applied to time series to find changes in macroeconomic relations, between financial series or other records of human behavior such a administrative records. Evidently, periods of financial and economic crisis such as the Global Financial Crisis and the European sovereign debt crisis impact our lives drastically. All chapters of this book illustrate how structural break tests can contribute to the evaluation of such events and thus help find appropriate policy responses. Structural break tests are, however, much more widely applicable, if variables are ordered. They can, for example, be used to find a significant change in a health indicator for a group of people sorted by age or another indicator. Another relevant function of structural break

testing is studying geospatial data, to find pattern changes in satellite images. One could also imagine a structural break test being used to detect a significant composition change in a probe of ice taken from a glacier. As soon as time is added as a dimension again, the application possibilities multiply. A test could be used to find a significant change in the area fluctuations of an ice sheet. The recording of a structural break for light sensor data would help identify movement of objects in the detection field of the sensor. Therefore, any advancement of methods used in structural break testing can have a substantial impact just by the large number of application possibilities.

Applied statistics in general can augment and improve the analysis of observations we make. Modern technology allows us to conveniently collect and track a large amount of data about ourselves, from calorie intake to time spent clicking a computer mouse and recording the locations we have been to. Structural break tests, specifically, can be employed to find changes of our behavior in these vast sets of data. Apps are created with this goal in mind. Computers enable us to reap the benefits of statistical analysis in a quick and convenient way. But a confident analysis of the suggested results requires some insight into the statistical methods being used. An understanding of structural break testing could help people make more sense of their behavior and how it has possibly changed.

It must be discussed, whether applicants of structural break testing programs understand what it means to have a significant structural break result. Structural break tests have become somewhat standard practice, a development that has led to and has been accelerated by the inclusion of structural break testing methods in popular econometric software such as Eviews and in open source distributions such as R. As with other statistical tests, the convenience of having a scalar test result with an associated p-value can also pose a thread. First, the p-value is not to be interpreted as the probability of the structural break hypothesis to be true. Secondly, structural break tests typically compare an abrupt change in the model to the no-change assumption, but they have power against much more general forms of change and instability. When a structural break is detected, all that is shown is that the assumed model without break is unlikely to have produced the observed data, but nothing is said about the base probabilities of no-break and break models and nothing is proven about the way in which the parameters of the model may have changed.

Given a positive structural break test result, the sum of likelihood ratio tests offers a way to investigate the following hypothesis that parameters of the model have changed abruptly at the estimated structural break date. The test incorporates more general forms of change as potential sources of the significance of a structural break hypothesis. This book's Chapter 5 has been written to describe the test and its concept. But further, the chapter is also intended to highlight and discuss the limited implications of a significant structural break test result in practice.

The descriptions of the tests of this book do not rely on a large amount of formal descriptions and derivations. As a relatively formal science, statistics is not a very accessible field of study for many people. When presenting the bootstrap, Efron and Tibshirani (1994) comment that the road to statistical knowledge is often blocked by a wall of mathematics. The bootstrap, a computer-based method of statistical inference, requires few formulas but answers real statistical questions. Generally, simulation based research may lower the barrier between statistics and the people interested in it.

Simulation based research also conforms to our relatively straightforward understanding of controlled experiments, and how the frequency of an observation can indicate the relevance of a pattern.

Not adding theoretical proof, when showing the validity of the discussed concepts of this book empirically, is an unusual choice in this area of research. Regardless, the concepts of this book have been explored and researched carefully and are worth sharing and discussing and exploring further. They are shown to be relevant, by simulation based means. The book is focused on conveying ideas in a illustrative and convincing way rather than their mathematical justification, similarly to the argument in Efron and Tibshirani. Ideally, the presentation will broaden the interest in the proposed concepts by making them accessible to a larger group of people.

Practicality and accessibility are central themes in this book. The intention of the sum of likelihood ratios of Chapter 5 is to make a more informed statement about the kind of change that has triggered a positive structural break test result. The new test uses data that is already available during structural break testing to test the next hypothesis: Can the change that has been discovered be assumed to have happened abruptly, or is there evidence against this assumption? Simulated or resampled observations that fulfill the abruptness assumption are used to conclude. The setup of the test is therefore very pragmatic and relatively straightforward.

Similarly, the rolling window algorithm of Chapter 4 has been developed to provide the benefit of multiple structural break testing at low computation cost. Time efficiency can be a concern in structural break testing, especially when data sets are large and when results must be presented quickly. The proposed algorithm offers a solution in such circumstances. Tradeoffs, however, have to be accepted when a heuristic such as the rolling window algorithm is used, as the resulting test is conservative and may only find a locally optimal result.

Diversity in research is a central principle of doctoral programmes. The methods of this book connect the topic of structural break testing to a range of other areas of science. They seek to incorporate insight from areas of science other than econometrics or applied statistics to enrich the literature on structural break testing. The idea to describe the rolling window algorithm of Chapter 4 has been developed when studying search algorithms, a topic of computer science. The heuristic can be discussed from the perspective of optimization theory. The need to provide an initial guess in several optimization techniques has inspired the combined method presented in the chapter, where the rolling window estimates serve as starting points for finding the global optimum. A feature selection technique has been adopted in Chapter 6, where cross-validated error estimates are used to find a structural break. Here, observations are excluded from the sample to test the existence of a structural break, a concept typically used in Machine Learning when building neural networks. Finally, topics of interest for macroeconomics and finance are discussed in all empirical examples of the presented tests, when studying the European sovereign debt crisis.

Parts of this book offer insight into a range of important aspects of structural break testing in general. The sequential testing procedure of Chapter 3 is designed specifically to capture financial contagion events in time series data. As such, it is developed to detect structural breaks that involve different types of parameters, which can pose problems of nuisance if not properly addressed. A solution that is incorporated in the se-

quential testing procedure is iterative testing, where break testing is applied conditional on previous results, and repeated until convergence of the estimates. A common break test hypothesis is included in the sequential testing procedure to account for estimation uncertainty and the possibility of joint structural breaks. Uncertainty of break estimation results is also explicitly addressed in other parts of the book. The cross-validated error estimation method of Chapter 6 uses random resampling of the observed data and thus directly incorporates a method of demonstrating the estimation uncertainty. Finally, since structural break detection test statistics typically have non-standard distributions, the book can also be used as a guide to implementing bootstrap empirical resampling methods to derive critical values of testing.

Jointly, the methods of this book can be combined into a powerful and comprehensive multiple structural break testing procedure. Cross-validated error estimation would be the central test statistic, a modern detection strategy that incorporates estimation uncertainty. The statistic would be calculated in a sequential, iterative procedure to account for nuisance, if several types of parameters had to be tested for structural breaks individually. Optionally, rolling windows of data could be used, if time-efficiency is a concern, possibly to derive an initial guess of a globally optimal solution. In the end, a summation of test statistic outcomes within specified ranges would offer additional insight into the transition duration of detected changes in the model. The product could be a package or library that is made available to a broad community of researchers using R syntax, and python scripts could be shared on open source platforms such as github.

The field of statistics can be compelling to a student of economics because its concepts defy time. Economic analysis is an exciting subject to a large degree because it depends on current definitions of economic concepts and on current understanding of economic value. This, however, makes results and conclusions of economic research somewhat fragile, and dependent on context. The methods of this book are invariant to context. The presented tests are shown to have certain properties in standard data simulations, regardless of the times in which we live.

This thesis combines timeless contributions to statistical concepts with modern topics of economics as well as modern issues of statistical testing. The rapidly increasing size and complexity of collected data imply that questions related to time efficiency in computation are a major concern. But also, rapid increases in computer processing power, Monte Carlos techniques and simulation based research enable an analysis that uses radically different statistical models than a few decades ago. Kroese et al. (2014) write that statistical analysis can be based on any probabilistic model that can be simulated on a computer. Data science is looking to significantly shape the production processes in the next years and decades. While interest in the specific contributions to statistical analysis will fluctuate, developments in statistics are certain to remain valid and widely applicable.

So far, the research of this thesis has led to a publication in an academic journal and has been discussed at four international conferences on applied econometrics and finance. Given a broad target audience, there are vast opportunities for the testing concepts of this thesis to be used to further knowledge creation in many areas of societal interest, and to inspire continued research on the specific topics and similar advancements in other areas of statistical research.