Essays on Structural Change Tests under Long Memory

Der Wirtschaftswissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des akademischen Grades

Doktor der Wirtschaftswissenschaften
— Doctor rerum politicarum —

genehmigte Dissertation

von

M.Sc. Kai Rouven Wenger
geboren am 08.05.1993 in Braunschweig

2020

Referent: Prof. Dr. Philipp Sibbertsen, Leibniz Universität Hannover

Koreferent: Prof. Dr. Michael Massmann, WHU - Otto Beisheim School of Management

Tag der Promotion: 23.01.2020

Acknowledgements

Many people provided support for this dissertation, and I owe all of them a debt of gratitude.
First, I want to thank my advisor and co-author Prof. Dr. Philipp Sibbertsen for his guidance over the last years. I highly appreciate his support, our interesting conversations, and the laughter we shared.
I also thank the members of my dissertation committee, Prof. Dr. Massmann, Prof. Dr. Schöndube, and Dr. Fabian Hollstein, for their suggestions.
Furthermore, I am grateful to all of my other co-authors, namely Dr. Christian Leschinski, Dr. Georg Ullmann, Janis Becker, and Benjamin Fritzsch. In particular, I want to thank Christian who helped me a lot in the beginning of my dissertation by sharing his experience and enthusiasm with me.
Many thanks are also due to all of my colleagues and former colleagues for creating such a productive and friendly working atmosphere. In addition to the already mentioned Christian Leschinksi, three of them became very dear to my heart: Tristan Hirsch and Janis Becker, with whom I became very close friends, and Dr. Michelle Voges, who accompanied me in my whole academic life from the start of my undergraduate studies until our time as PhD students.
Thanks must also go to Dr. Joachim Schnurbus, Dr. Annika Betken, and attendees of numerous conferences (Statistical Weeks 2016-2019, SNDE 2018, CFE 2018, 2019) that have provided critical remarks on the work in this dissertation.
Finally, I want to thank my brother Yannick Wenger who has proofread nearly all of my written work since the beginning of my studies in 2011 providing many helpful and often sarcastic comments. Lastly, I want to thank my parents Petra and Michael Wenger for their permanent, outstanding support.

Abstract

This thesis contains six essays on tests for structural change under long memory with applications to economic time series. Chapter 1 introduces structural change tests and the concept of long memory. The main problem examined in Chapters 2 to 5 is that standard change-in-mean tests are invalid in long-memory time series.

Chapter 2 reviews the literature on long-memory robust extensions of standard testing principles for a change in mean. Apart from giving a systematic review, an extensive Monte Carlo study is conducted to compare the relative performance of the introduced methods. Special attention is paid to the interaction the test results have with the estimation of the long-memory parameter. Furthermore, it is shown that the power of self-normalized test statistics can be improved considerably by using an estimator that is robust to mean shifts.

Chapter 3 introduces a simple test on structural change in long-memory time series. In contrast to the testing principles introduced in the previous chapter, it is much easier to implement in statistical software and has a limiting distribution that does not depend on the degree of memory. The test is based on the idea that the test statistic of the standard CUSUM test retains its asymptotic distribution if it is applied to fractionally differenced data. It is proven that the approach is asymptotically valid if the memory is estimated consistently under the null hypothesis. Therefore, the well-known CUSUM test can be used on the differenced data without any further modification. In simulations, the proposed CUSUM test on the differenced series is compared with a CUSUM test on structural change that is specifically constructed for long-memory time series. It is observed that the new approach performs reasonably well.

The two chapters described previously find that self-normalized tests on change in mean are robust against size distortions in finite samples and that the CUSUM testing principle tends to be the most powerful. Based on these results, Chapter 4 proposes a new family of self-normalized CUSUM tests for structural change under long memory. The test statistics apply non-parametric kernel-based long-run variance estimators and have well-defined limiting distributions that only depend on the long-memory parameter. A Monte Carlo simulation shows that these tests provide finite sample size control while outperforming the competing procedures, which are presented in the previously described chapters, in terms of power.

Chapter 5 presents the memochange package which offers implementations of all methods reviewed and proposed in Chapters 2 to 4 in the programming language R. The package

is complemented with implementations of the most prominent change-in-persistence tests and estimation methods.
Conversely to standard structural change tests which are invalid in long-memory time series, inference on the memory order is also biased (upwards) by the presence of structural change and other so called 'low-frequency contaminations'. Since the presence of long memory invalidates standard inference about structural breaks and vice versa, it is not clear for many economic time series - such as asset volatilities - what their actual degree of memory is.
In Chapter 6, a comprehensive analysis of the memory in volatilities of international stock indices and exchange rates is provided. On the one hand, it is found that the volatility of exchange rates is subject to spurious long memory and the true memory parameter is in the higher stationary range. Stock index volatilities, on the other hand, are free of low-frequency contaminations and the memory is in the lower non-stationary range. These results are obtained using state-of-the-art local Whittle methods that allow consistent estimation in presence of perturbations or low-frequency contaminations.
In Chapter 7 standard time series models are combined with search query data to examine whether they can be helpful in predicting sales. Search volume of company as well as product related keywords provided by Google Trends are included as new predictors in models to forecast sales on a product level. Using weekly data from January 2015 to December 2016 of two products of the audio company Sennheiser, evidence is found that using Google Trends data can enhance the prediction performance of conventional models.

Keywords: Fixed-bandwidth Asymptotics · Fractional Integration · Google Econometrics Forecasting · High-frequency Data · Long Memory · Perturbations · Realized Volatility Search Query Data · Spurious Long Memory · Structural Breaks

Contents

Chapter 1

Introduction

For modeling and forecasting a time series, it is of major importance to adequately capture its dependency structure. Commonly, this is achieved using (short-memory) ARMA models. A characteristic of these models is that the impact of shocks is short-lived and dies out quickly. Hence, the autocorrelation function of these models decays exponentially. However, many time series seem to be more persistent than it can be captured by short-memory models. Examples of economic series that typically exhibit high persistence are, among many others, inflation, interest, exchange, and unemployment rates, as well as real output, volatilities, and consumption (e.g. Diebold and Rudebusch (1989), Cheung (1993), Hassler and Wolters (1995), Baillie et al. (1996), Tsay (2000), Van Dijk et al. (2002), Martens et al. (2009), Gil-Alana et al. (2010), Chiriac and Voev (2011)).

Therefore, modeling and forecasting these series is commonly done using long-memory models. In the time domain, long memory is defined by a slow hyperbolical decay of the autocorrelation function $\gamma(k)$ at large lags, such that $\gamma(k) \sim G_\gamma |k|^{2d-1}$ as $k \to \infty$, where G_γ is a finite constant, k is the lag, and d is the long-memory parameter that relates to the persistence of the series. Higher values of d indicate higher persistence of the series. In the frequency domain, long memory can be equivalently defined by a pole in the spectral density $f(\lambda)$ local to the zero frequency, such that it obeys the power law $f(\lambda) \sim G_f |\lambda|^{-2d}$ as $\lambda \to 0$, where G_f is a finite constant, λ is the frequency, and d denotes again the long-memory parameter (cf. Beran et al. (2016)).

Long-memory processes are parametrically modeled predominantly by ARFIMA processes proposed by Granger and Joyeux (1980) and Hosking (1981). Since consistent maximum likelihood estimation of the model parameters requires a specification of the unknown ARMA components, semiparametric estimators are often applied. They rely on the above specification of the spectral density local to the origin to estimate d and thus the short-run dynamics must not be estimated. Popular semiparametric memory estimators are local Whittle based estimators by Künsch (1987), Robinson (1995), and Shimotsu and Phillips (2005) and the log-periodogram regression estimator by Geweke and Porter-Hudak (1983).

The key features of long-memory processes, an asymptotically hyperbolically decaying autocorrelation function as well as poles in the spectrum local to the zero frequency, can also be generated by other processes. When a short-memory time series is contaminated by a so called low-frequency contamination, for example a mean shift, long memory is falsely detected by standard memory estimation approaches. In the literature, this is referred to as 'spurious long memory' (e.g. Diebold and Inoue (2001), Granger and Hyung (2004), or Mikosch and Stărică (2004)). Conversely, traditional testing procedures for a mean shift are similarly invalidated when the series exhibits long memory. This is due to the fact that they are developed for independent or serially correlated series but not for long-range dependent data. Wright (1998) and Krämer and Sibbertsen (2002) even show that these tests reject the null hypothesis of a constant mean asymptotically with a probability of one if $d > 0$.

Therefore, a lot of research has recently focused on robust estimation methods and testing procedures when both features, long memory and structural change, could be present in a time series.

Tests to distinguish pure long-memory time series from series contaminated by low-frequency contaminations are proposed by Qu (2011) and Sibbertsen et al. (2018), among others. Robust estimation methods for the long-memory parameter are developed, for example by McCloskey and Perron (2013) and Hou and Perron (2014).

Chapters 2 to 5 examine tests for a change in mean under long memory. Thus, the focus lies on testing the null hypothesis of a constant unconditional mean against the alternative of a change in mean at an unknown time. Three standard procedures for this testing problem, which are valid when the series exhibits short memory, are CUSUM tests originally proposed by Brown et al. (1975), Wilcoxon-type rank tests (e.g. Bauer (1972)), and sup-Wald type tests by Andrews (1993). Since they suffer from the issue that they spuriously reject the null in a pure long-memory time series, as discussed above, numerous modification of these tests have been proposed over the past years (e.g. Wang (2008), Shao (2011), Dehling et al. (2013), Iacone et al. (2014), Betken (2016)).

Chapter 2 reviews all relevant change-in-mean tests for long-memory time series. They are categorized into two dimensions: First, which type of testing principle they are based on (CUSUM, Wilcoxon, sup-Wald). Second, if the long-run variance is estimated by a consistent estimator or self-normalization approach. Apart from giving a systematic review, the relative finite sample performance of all tests is evaluated in an extensive Monte Carlo study. Up to this point, the newly proposed procedures have been compared at most pairwise in simulations. Hence, it was not possible to make any recommendation which method should be selected in practice. Chapter 2 closes this gap by showing that the CUSUM test tends to be the most powerful testing principle, but that the self-normalized Wilcoxon test of Betken (2016) and the fixed-b sup-Wald test of Iacone et al. (2014) offer the best tradeoff between size control and power. Furthermore, the effect of estimating the memory parameter that appears in the test statistics and/or the limit distributions is examined by simulation. It is demonstrated that substantial power gains can be realized if the modified local Whittle estimator of Hou and Perron (2014) is applied to estimate the long-memory parameter.

Chapter 3 proposes a new test on change in mean in a long-memory time series. The idea of the test is that the test statistic of the standard CUSUM test retains its asymptotic distribution if it is applied to fractionally differenced data. Therefore, the well-known CUSUM test can be used on the differenced data without any further modification. Compared to the testing principles reviewed in Chapter 2, this makes the new test computationally much easier to implement. Furthermore, in contrast to all the reviewed modified tests, the memory parameter does not appear in the limiting distribution of the test statistic. This has the advantage that the critical values of the test do not depend on the long-memory parameter. Hence, the CUSUM test on the differenced data can be applied without performing additional simulations to determine the correct critical values for the estimated value of d. The chapter is completed by simulations which show that the new approach performs reasonably well.

Based on the observation of the two previous chapters that CUSUM tests tend to be the most powerful and that self-normalized tests offer the best size control, Chapter 4 proposes a new family of self-normalized CUSUM tests for structural change under long memory. The test statistics apply non-parametric kernel-based long-run variance estimators, and it is proven that they have a well-defined limiting distribution. Using simulations, the proposed family of self-normalized CUSUM tests is compared with the tests that were shown to offer the best size-power trade-off in the Monte Carlo section of Chapter 2, i.e. the self-normalized Wilcoxon test of Betken (2016) and the fixed-b sup-Wald test of Iacone et al. (2014). First, it is found that the self-normalized CUSUM tests provide finite sample size control in contrast to the CUSUM test of Horváth and Kokoszka (1997) and Wang (2008) that applies a consistent long-run variance estimator. Second, it is observed that the tests outperform the competing procedures in terms of power. Therefore, it is suggested to consider one of the self-normalized CUSUM tests when testing for a change in mean under long memory.

Chapter 5 presents the memochange package which offers implementations of all methods reviewed and proposed in Chapters 2 to 4 in the programming language R. The package is complemented with implementations of the most important change-in-persistence tests and estimation methods.

In Chapter 6, the memory of realized volatility of international stock indices and exchange rates is analyzed. In literature, estimated memory parameters for such series range roughly between 0.4 and 0.6 - that is from the higher stationary to the lower non-stationary region (c.f. Andersen et al. (2003), Hurvich and Ray (2003), Martens et al. (2009), among others). However, the presence of measurement error in realized volatility and the potential of spurious long memory is not considered in most of these publications. We provide a more comprehensive analysis by using the aforementioned state-of-the-art methods that allow consistent estimation of the memory parameter in presence of perturbations or low-frequency contaminations (e.g. Frederiksen et al. (2012), McCloskey and Perron (2013), Hou and Perron (2014), Sibbertsen et al. (2018)). On the one hand, it is found that the volatility of exchange rates is subject to spurious long memory and the true memory parameter is in the higher stationary range. On the other hand, stock index volatilities are free of low-frequency contaminations and the memory is in the lower non-stationary range.

Chapter 7 discusses the usage of online search query data to improve sales forecasts. Nowadays, online search engines provide new sources of data about real-time economic activity (Ettredge et al. (2005)). On a product level, it is examined whether including Google Trends data as new predictors in standard time series models improves forecasting sales. Due to companies not making sales data openly available, the literature on forecasting sales on a product level is scant (Cui et al. (2018), Boone et al. (2018)). The research up to now has focused on data provided by online retailers, which directly sell their products to the customer. Since these companies fully execute their business online, the link between using a search engine and buying the products is tight.

In contrast, Chapter 7 analyzes a unique set of weekly data from January 2015 to December 2016 of two products of the audio company Sennheiser. The principal business of Sennheiser is to sell their products on the spot through department stores. Therefore, the link between searching for a product and buying it is not as close as for an online retailer and it is not clear whether the results obtained in the literature carry over to the products investigated. Considering the search volume of company as well as product related keywords provided by Google Trends as new predictors, evidence is found that using search query data can enhance the prediction performance of conventional models.

Chapter 2

Change-in-Mean Tests in Long-memory Time Series: A Review of Recent Developments

Co-authored with Christian Leschinski and Philipp Sibbertsen.

Published in Advances in Statistical Analysis (2019), 103(2), 237-256.

https://doi.org/10.1007/s10182-018-0328-5

Chapter 3

A Simple Test on Structural Change in Long-Memory Time Series

Co-authored with Christian Leschinski and Philipp Sibbertsen.
Published in Economics Letters (2018), 163, 90-94.

https://doi.org/10.1016/j.econlet.2017.12.007

Chapter 4

Fixed-Bandwidth CUSUM Tests Under Long Memory

Co-authored with Christian Leschinski.
Published in Econometrics and Statistics (forthcoming).

https://doi.org/10.1016/j.ecosta.2019.08.001

Chapter 5

memochange: An R Package for Estimation and Tests in Persistent Time Series

Co-authored with Janis Becker.
Published in The Journal of Open Source Software.

https://doi.org/10.21105/joss.01820

Chapter 6

The Memory of Volatility

Co-authored with Christian Leschinski and Philipp Sibbertsen.

Published in Quantitative Finance and Economics (2018), 2(1), 137-159.

https://doi.org/10.3934/QFE.2018.1.137

Chapter 7

Can Google Trends improve Sales Forecasts on a Product Level?

Co-authored with Benjamin Fritzsch, Philipp Sibbertsen, and Georg Ullmann.
Published in Applied Economics Letters (forthcoming).

https://doi.org/10.1080/13504851.2019.1686110

Bibliography

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Andrews, D. W. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica: Journal of the Econometric Society*, pages 821–856.

Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3–30.

Bauer, D. F. (1972). Constructing confidence sets using rank statistics. *Journal of the American Statistical Association*, 67(339):687–690.

Beran, J., Feng, Y., Ghosh, S., and Kulik, R. (2016). *Long-Memory Processes.* Springer.

Betken, A. (2016). Testing for change-points in long-range dependent time series by means of a self-normalized wilcoxon test. *Journal of Time Series Analysis*, 37(6):785–809.

Boone, T., Ganeshan, R., Hicks, R. L., and Sanders, N. R. (2018). Can google trends improve your sales forecast? *Production and Operations Management*.

Brown, R. L., Durbin, J., and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 149–192.

Cheung, Y.-W. (1993). Long memory in foreign-exchange rates. *Journal of Business & Economic Statistics*, 11(1):93–101.

Chiriac, R. and Voev, V. (2011). Modelling and forecasting multivariate realized volatility. *Journal of Applied Econometrics*, 26(6):922–947.

Cui, R., Gallino, S., Moreno, A., and Zhang, D. J. (2018). The operational value of social media information. *Production and Operations Management*, 27(10):1749–1769.

Dehling, H., Rooch, A., and Taqqu, M. S. (2013). Non-parametric change-point tests for long-range dependent data. *Scandinavian Journal of Statistics*, 40(1):153–173.

Diebold, F. X. and Inoue, A. (2001). Long memory and regime switching. *Journal of econometrics*, 105(1):131–159.

Diebold, F. X. and Rudebusch, G. D. (1989). Long memory and persistence in aggregate output. *Journal of monetary economics*, 24(2):189–209.

Ettredge, M., Gerdes, J., and Karuga, G. (2005). Using web-based search data to predict macroeconomic statistics. *Communications of the ACM*, 48(11):87–92.

Frederiksen, P., Nielsen, F. S., and Nielsen, M. Ø. (2012). Local polynomial whittle estimation of perturbed fractional processes. *Journal of Econometrics*, 167(2):426–447.

Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4(4):221–238.

Gil-Alana, L. A., Loomis, D., and Payne, J. E. (2010). Does energy consumption by the us electric power sector exhibit long memory behavior? *Energy Policy*, 38(11):7512–7518.

Granger, C. W. and Hyung, N. (2004). Occasional structural breaks and long memory with an application to the s&p 500 absolute stock returns. *Journal of empirical finance*, 11(3):399–421.

Granger, C. W. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15–29.

Hassler, U. and Wolters, J. (1995). Long memory in inflation rates: International evidence. *Journal of Business & Economic Statistics*, 13(1):37–45.

Horváth, L. and Kokoszka, P. (1997). The effect of long-range dependence on change-point estimators. *Journal of Statistical Planning and Inference*, 64(1):57–81.

Hosking, J. R. (1981). Fractional differencing. *Biometrika*, pages 165–176.

Hou, J. and Perron, P. (2014). Modified local whittle estimator for long memory processes in the presence of low frequency (and other) contaminations. *Journal of Econometrics*, 182(2):309–328.

Hurvich, C. M. and Ray, B. K. (2003). The local whittle estimator of long-memory stochastic volatility. *Journal of Financial Econometrics*, 1(3):445–470.

Iacone, F., Leybourne, S. J., and Taylor, R. A. (2014). A fixed-b test for a break in level at an unknown time under fractional integration. *Journal of Time Series Analysis*, 35(1):40–54.

Krämer, W. and Sibbertsen, P. (2002). Testing for structural changes in the presence of long memory. *International Journal of Business and Economics*, 1(3):235–242.

Künsch, H. R. (1987). Statistical aspects of self-similar processes. *Proceedings of the first world congress of the Bernoulli Society*, 1:67–74.

Martens, M., Van Dijk, D., and De Pooter, M. (2009). Forecasting S&P 500 volatility: Long memory, level shifts, leverage effects, day-of-the-week seasonality, and macroeconomic announcements. *International Journal of Forecasting*, 25(2):282–303.

McCloskey, A. and Perron, P. (2013). Memory parameter estimation in the presence of level shifts and deterministic trends. *Econometric Theory*, 29(6):1196–1237.

Mikosch, T. and Stărică, C. (2004). Nonstationarities in financial time series, the long-range dependence, and the igarch effects. *Review of Economics and Statistics*, 86(1):378–390.

Qu, Z. (2011). A test against spurious long memory. *Journal of Business & Economic Statistics*, 29(3):423–438.

Robinson, P. M. (1995). Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics*, pages 1630–1661.

Shao, X. (2011). A simple test of changes in mean in the possible presence of long-range dependence. *Journal of Time Series Analysis*, 32(6):598–606.

Shimotsu, K. and Phillips, P. C. (2005). Exact local whittle estimation of fractional integration. *The Annals of Statistics*, 33(4):1890–1933.

Sibbertsen, P., Leschinski, C., and Busch, M. (2018). A multivariate test against spurious long memory. *Journal of Econometrics*, 203(1):33 – 49.

Tsay, W.-J. (2000). Long memory story of the real interest rate. *Economics Letters*, 67(3):325–330.

Van Dijk, D., Franses, P. H., and Paap, R. (2002). A nonlinear long memory model, with an application to us unemployment. *Journal of Econometrics*, 110(2):135–165.

Wang, L. (2008). Change-in-mean problem for long memory time series models with applications. *Journal of Statistical Computation and Simulation*, 78(7):653–668.

Wright, J. H. (1998). Testing for a structural break at unknown date with long-memory disturbances. *Journal of Time Series Analysis*, 19(3):369–376.