Erasmus School of Economics

Dependences between long-run covariability and data frequencies

Jorian van der Kuilen

Student number 427233

July 7, 2019

Supervisor: dr. (M.) Grith

First reader: dr. (A.A.) Naghi

Abstract

In order to deal with paucity of sample inference and the problem that inference critically depends on the long-run persistence of the data, Müller and Watson (2018) propose models that use low-frequency weighted averages to construct asymptotically efficient confidence intervals for long-run covariability parameters. The (A, B, c, d) model constructs these confidence intervals for a wide range of persistence patterns. The accurateness of the constructed confidence intervals using the (A, B, c, d) model is demonstrated by the simulaton section. Furthermore, this paper elaborates on Müller and Watson (2018) by measuring the performance of their models on data sets with higher frequencies on the basis of the relationship between the S&P 500 Index and the S&P 500 Futures. Further knowledge on the performance of the methods proposed by Müller and Watson (2018) can be of value to econometric and economic researchers and knowledge on the relationship between the S&P 500 Index and the S&P 500 Futures is interesting to worldwide investing market participants. It is observed that the impact of different data frequencies on the long-run covariability between the S&P 500 Index and the S&P 500 Futures does not follow a clear pattern. Furthermore, differences between sub-sample are captured by the (A, B, c, d) model considering, for example, the sharp decrease of long-run covariability between both variables after the U.S. stock market crisis in December 2018.

Contents

1 Introduction and literature

Research on long-run covariability between economic variables has a long history in finance, economics and econometrics. Predictions on the relationships between economic variables are common in economic theories, but the statistical tools to investigate the validity of these predictions are limited. Although there are limited statistical tools, long-run relationships are widely used by influential economists, financial institutions, and policymakers. There are two fundamental problems that econometrists encounter when investigating long-run economic theories. The first problem consists of a shortage of long-run sample information on economic variables, as data samples that are used to investigate long-horizon relationships do not contain enough observations. To deal with this for example, Valdovinos (2003) uses long-run data components instead of original data to demonstrate that there is a clear negative relation between inflation and economy growth. The long-run data components that are used in this investigation are extracted from a filter that isolates business-cycle fluctuations in macroeconomic time series initiated by Bakter and King (1995). The second problem is the high dependence of the inference on the long-run persistence of the data that is used. Probability distributions of data series are uniquely characterized and different from what is observed in comparable series. The paucity of statistic inference about long-run phenomena and the high use of economic relations in practice emphasize the importance of research on long-run covariability. The development of methods that are able to deal with the described problems can be of value for economic and econometric research in many fields.

Many methods to measure the long-run covariability between specific stochastic processes are developed in the last 45 years. Inference in cointegration models has grown during these decades, which is particularly interesting as cointegration is another widely used tool to measure long-run relationships between variables. Phillips (1991) shows that spectral methods are useful to regressions for certain nonstationary time series by the use of the block triangular error correction model representation of cointegrated systems. Stock and Watson (1993) propose in their analysis on the long-run U.S. money demand efficient GLS estimators of cointegration vectors for systems with higher orders of integration. As cointegrated variables have unit long-run correlations, many long run relationships fall outside the standard framework of cointegration. Inference about long-run covariability for wider ranges of long-run persistence is developed by Hounyo (2014) in his investigation on distributions of covariance estimators in integrated processes. Müller and Watson (2008) contribute to inference on general patterns of long-run persistence by constructing methods

that use low-frequency averages to isolate long-run sample information.

Müller and Watson (2018), the main source for this thesis, focuses on low-frequency data averages in order to develop methods that construct confidence intervals for long-run relationship parameters. The data transformations applied in this paper are based on specific band-pass filters that isolate business-cycle variations in macroeconomic time-series (Baxter & King, 1999). Müller and Watson (2017a) conclude that inference about variability and covariabilty of integrated processes in the long-run can be conducted from low-frequency weighted averages, as they follow large sample normal distributions. This paper elaborates on the properties of the weighted averages compared to asymptotic implications of certain time series (Müller & Watson, 2008). The parameters that are used to measure the long-run covariability parameters between variables are based on the low-frequency band spectrum regression from Engle (1974). Müller and Norets (2016) explore the validness of estimates and confidence intervals for parameters of $I(0)$, $I(1)$, near unit roots, fractionally integrated models and linear combinations of variables that are characterized by these persistence forms. The least favorable distribution and the construction of confidence intervals based on testing problems that are used to construct these confidence intervals are explained by Elliott, Müller and Watson (2015).

The long-run covariability inference that is developed by Müller and Watson (2018) is applied on two economic relationships in the first part of the paper. First, the relation between GDP and consumption is one of the "Great Ratios" of the U.S. economy (Klein & Kosobud, 1961). Cochrane (1994) uses two-variable autoregressions to conclude that there is cointegration between consumption and income, which is also concluded by Campbell (1987) in his post war analysis on relationships between conditional means and conditional variances. The conclusion of these papers is contradicted by Lettau and Luvigson (2013), using three historical market shocks to analyze the post-war dynamics of consumer spending, labor earnings, and household wealth.

The second main application of the methods proposed by Müller and Watson (2018) is the long-run relationship between short-term and long-term interest rates. Earlier analysis of this correlation concluded cointegration between short-term and long-term interest rates by modeling the interest rates as $I(1)$ processes (Campbell & Shiller, 1987). Dai and Singleton (2000) use affine term structure models to explain historical behavior of interest rates and accessory long-run persistence.

The two long-run relations that are explained above are the main applications of the methods used by Müller and Watson (2018), involving an analysis on the long-run covariability between variables that are observed quarterly and monthly. In the last section of their paper, Müller and

Watson (2018) apply their methods to other relationships between quarterly and monthly observed variables. The absence of applications to data of higher frequencies triggers curiosity for findings on the application of the low-frequency approach on weekly, daily and intraday data. To investigate the performance of the methods initiated by Müller and Watson (2018) on high-frequency data, I consider the relationship between the S&P 500 Index and the S&P 500 Futures in the extension part of this paper. The S&P 500 Index is a market-capitalization-weighted index, based on the 500 largest publicly traded companies in the U.S. The S&P 500 Futures are derivative contracts that provide buyers prices based on the S&P 500 Index's expected future value to with investments. Both series are important indicators of the worldwide investing market.

There are three sub-analyses that are used to investigate the long-run covariability between the S&P 500 Index and the S&P 500 Futures. First, the impact of different data frequencies on the covariability between the S&P 500 Index and the S&P 500 Futures is analyzed. The performance of the methods proposed by Müller and Watson is measured on the basis of covariability differences between monthly, weekly and daily S&P 500 series. For the first analysis, the U.S. stock market cycle of about 42 months is taken into account (Chonga, Li, Chen & Hinich, 2010). Second, the high volatility of the investing market has encouraged to compare the covariability between the S&P 500 Index and the S&P 500 Futures on different sub-samples. The purpose of the final sub-analysis is to investigate whether the (A, B, c, d) model, that is initiated by Müller and Watson (2018), can be used to obtain an estimation of the lead time of the S&P 500 Futures on the S&P 500 Index by the use of intraday data. The highly volatile relationship between the S&P 500 Index and the S&P 500 Futures is earlier investigated by Kwaller, Koch and Koch (1990). This paper concludes that futures volatility is higher than index volatility and both volatilities increase as the futures trading volume increases. Arshanapalli and Doukas (1994) elaborate on these conclusions by showing that the S&P 500 Index and the S&P 500 Futures do not share the same volatility process. The lead time of the S&P 500 Futures is investigated by Kwaller, Koch and Koch (1987) by a three-stage-least squares approach. Their analysis suggests that the lead time from futures to cash prices ranges from twenty to 45 minutes. In the same year, Herbst, McCormack and West (1987) use spectral analysis to conclude that the S&P 500 spot index reacts to changes in the futures contract in less than one minute. Knowledge on the covariability between the S&P 500 Index and the S&P 500 Futures and the lead time of the S&P 500 Futures can be of value to participants in the worldwide investing market. Besides this, additional knowledge on the performance of the methods that are proposed by Müller and Watson (2018), contributes to the inference on long-run covariability models.

The results of the first analysis indicate that the long-run covariability between the S&P 500 Index and the S&P 500 Futures is effected by the frequency of the data that is used, but the parameter values do not follow a clear pattern as the data frequency increases or decreases. Covariability differences between weekly and daily sub-periods are captured by the (A, B, c, d) model in the second sub-analysis and the clearly lower covariability in the fourth daily sub-period that contains the U.S. stock market crash in December 2018 is the most eye catching example for this. The use of the (A, B, c, d) model in the final sub-analysis does not result into a clear estimation of the lead time of the S&P 500 Futures on the S&P 500, as multiple relationships with different Futures lead times and data frequencies are estimated to be approximately equally correlated in the long-run.

The outline of this paper is as follows: section 2 consists of descriptions of the data and methods that are used to obtain all results in this paper. These results are presented in section 3, that can be divided into three subsections: section 3.1 contains simulation results, section 3.2 contains the replications of the findings of Müller and Watson (2018) and section 3.3 is the extension section that elaborates on Müller and Watson (2018) by application of proposed methods to the relationship between the S&P 500 Index and the S&P 500 Futures. Finally, section 4 concludes and contains a discussion of this paper.

2 Data and Methodology

This section contains a full clarification of the data and methodology that are used in order to obtain the results of this paper. First, the data of the paper is described for both the replication and the extension part of this thesis. Secondly, subsection 2.2 to 2.5 describe the methods and computations that are performed to obtain the findings of the replicaon part and the extension part. Finally, section 2.6 describes the simulation that is used to draw conclusions on the confidence intervals that are estimated by the methods. A description of the used Matlab code that used to apply described methods can be found in Appendix J.

2.1 Data

In this paper, two data sets are used. First, the replication part of this paper makes use of the same data set as Müller and Watson (2018). All variables that are used in the replication part of this paper are observed at monthly or quarterly frequencies. Second, the extension data set is used to obtain the results from the extension part of this paper. Growth rates in % of the S&P

500 Index and S&P 500 Futures prices are obtained from Investing.com and the observed data frequencies range from monthly to minutely. Appendix A contains a list of definitions, data sources, frequencies, and observation periods of all used variables.

2.2 Long-run projections

Properties of economic time series are typically analyzed in the time domain, but due to limited available data, Müller and Watson (2018) analyze the long-run economic relationships between economic variables in the frequency domain approach, i.e. spectral analysis. In most spectral analyses, Fourier transforms are used to transform data from the time domain to the frequency domain. Fourier transforms represent time series as series in cosines and sines and are useful in dealing with periodic and rhythmic economic phenomena. This makes picking out cyclical behavior in time series generally the main application of analyzing time series in the frequency domain.

In Müller and Watson (2018), the real data is transformed into projections onto low-frequency periodic functions. These transformations are constructed such that fluctuations of periods longer than eleven years can be captured. This minimum captured periods are longer than the U.S. business cycle of approximately eight years (Baxter & King, 1999). The data is transformed as follows: let $x_t, t = 1, ..., T$ denote a non-transformed data series from the data set. Let $\Psi_j(s) = \sqrt{2}\cos(js\pi)$ denote the cosine function for the j-th, $j = 1 \dots q$, cosine transform, such that $\Psi(s) = [\Psi_1(s), \Psi_2(s), ..., \Psi_q(s)]'$ becomes a matrix with dimensions $T \times q$ and t-th row given by $\Psi((t - 1/2)/T)'$. q is chosen such that periodicities longer than $\frac{2T}{q}$ are captured by the model. For the data set used in the replication part, this means that periodicities longer than 11.3 years are captured by $q = 12$ for GDP and consumption and $q = 11$ for the interest rates. The projection of the data x_T onto $\Psi((t - 1/2)/T$, for t = 1,...T, is given by

$$\begin{aligned} \overline{x}_t &= X_T' \Psi((t - 1/2)/T, \\ X_T &= T^{-1} \sum_{t=1}^{T} \Psi((t - 1/2)/T) x_t, \end{aligned} \tag{1}$$

where X_T denotes the $q \times 2$ projection coefficients matrix that represents the cosine-weighted averages of the data. Because of orthoginality of the cosine regressors Ψ_T, explained by Müller and Watson (2017b), the variability and covariability in the cosine transforms (X_T and Y_T) and long-run projections ($\overline{x}_t$ and $\overline{y}_t$) are strongly related. The equivalence between the long-run projection

coefficients and the cosine transforms shown in equation (2), indicates that the sample covariability in the cosine transforms coincides with the sample covariability in time series projections:

$$T^{-1}\sum_{t=1}^{T}\begin{pmatrix}\overline{x}_t\\ \overline{y}_t\end{pmatrix}(\overline{x}_t\overline{y}_t) = T^{-1}\begin{pmatrix}X_T'\\ Y_T'\end{pmatrix}\Psi_T'\Psi_T(X_TY_T) = \begin{pmatrix}X_T'X_T & X_T'Y_T\\ Y_T'X_T & Y_T'Y_T\end{pmatrix} \tag{2}$$

The long-run projections show only little differences with the low-pass moving averages (Müller & Watson, 2018) and due to the orthogonality of the cosine regressors Ψ_T, tight connections between the covariability in $(x,\, y)$ and in (X_T, Y_T) are observed.

2.3 Long-run covariability measurements

Müller and Watson (2018) find in their research a clear way to measure the long-run covariability between variables. Let the covariance matrix of projection coefficients X_T and Y_T be denoted by Σ_T, partitioned as $\Sigma_{XX,T}$, $\Sigma_{XY,T}$, $\Sigma_{YX,T}$, $\Sigma_{YY,T}$. Then, the average covariance matrix of the projections $\overline{x}_t$ and $\overline{y}_t$, Ω_T, is defined as follows:

$$\Omega_T = T^{-1}\sum_{t=1}^{T}E\left[\begin{pmatrix}\overline{x}_t\\ \overline{y}_t\end{pmatrix}(\overline{x}_t\overline{y}_t)\right] = \sum_{j=1}^{q}E\left[\begin{pmatrix}X_{jT}\\ Y_{jT}\end{pmatrix}\begin{pmatrix}X_{jT}\\ Y_{jT}\end{pmatrix}'\right] = \begin{pmatrix}tr(\Sigma_{XX,T}) & tr(\Sigma_{XY,T})\\ tr(\Sigma_{YX,T}) & tr(\Sigma_{YY,T})\end{pmatrix} \tag{3}$$

The long-run covariability parameters that are used to evaluate the covariability between variables in this paper are constructed in the following equation:

$$\begin{aligned}\rho_T &= \Omega_{xy,T}/\sqrt{\Omega_{xx,T}\Omega_{yy,T}}\\ \beta_T &= \Omega_{xy,T}/\Omega_{xx,T}\\ \sigma^2_{y|x,T} &= \Omega_{yy,T} - (\Omega_{xy,T})^2/\Omega_{xx,T},\end{aligned} \tag{4}$$

where Ω_T is constructed as $\begin{pmatrix}\Omega_{xx,T} & \Omega_{xy,T}\\ \Omega_{xy,T} & \Omega_{yy,T}\end{pmatrix}$. If we regress the long-run projection of one variable $\overline{y}_t$ on the long-run projection of another variable $\overline{x}_t$, β_T, $\sigma_{y|x,T}$ and $\rho_{xy,T}$ are interpreted as the linear regression coefficient, average variance of the prediction error and population R^2 respectively. Together these three parameters are accurate measurements of the long-run covariability between two selected variables x_t and y_t. A more extended derivation of these parameters is provided by Müller and Watson (2017a).

2.4 Parameterizing long-run persistence and covariability

The previous subsection of the methodology shas shown that ρ, β and $\sigma_{y|x}$ are functions of Ω, which is defined as the average covariance matrix of the long-run projections ($\overline{x}$ and $\overline{y}$). This 2×2 matrix is defined as a function of the covariance matrix of the cosine projection coefficients, Σ. Because cosine projection coefficients are smooth averages of the long-run projections, central limit theorem effects suggest that these projection coefficients are Gaussian and therefore $\sqrt{T}\begin{pmatrix}X_T\\Y_T\end{pmatrix} \Rightarrow \begin{pmatrix}X\\Y\end{pmatrix} \sim \mathcal{N}(0,\Sigma)$ with limiting Σ, if the spectral density converges close to zero for all frequencies. These large-sample results hold if z_t has a Fourier Transform of the autocorrelation function and indicate that, as $T\Sigma_T \rightarrow \Sigma$, also $(\rho_T, \beta_T, \sigma^2_{y|x,T}) \rightarrow (\rho, \beta, sigma^2_{y|x}$ and efficient inference in smaller problems amounts to large-sample efficient inference (Müller & Watson, 2018). The elements of Σ are now dependent on the shape of the spectral density close to the origin. This function is called the local-to-zero spectrum, denoted as $S_{\Delta z}$ for the first-difference of z and as $S_z(\omega) = \omega^{-2}S_{\Delta z}(\omega)$, with $z_t = (x_t, y_t)'$ and frequency ω for the level of z_t (Müller & Watson, 2015).

A process is integrated of order d, $I(d)$, if the d-th difference, $\Delta^d\xi$, is $I(0)$, which in turn is a process that is stationary and has finite and positive long-run variance. An $I(0)$ process can be written as $\delta + u_t$, where u_t is a zero-mean stationary process with positive long-run variance. If z_t is $I(0)$ with long-run covariance matrix Λ, then $\Sigma = \Lambda \otimes I_q$ and $\Omega = \Lambda$. This means that $z \sim iid\, \mathcal{N}(0,\Lambda)$ and the covariance matrix of the original data, Λ, is equal to the covariance matrix of the long-run projections, Ω. If z_t is $I(1)$ and differences Δz_t have covariance matrix Λ, let D be some diagonal matrix with dimensions $q \times q$ and elements $D_{ii} = \frac{1}{(i\pi)^2}$. Then it holds that $z \sim id\, \mathcal{N}(0, \Lambda/i^2)$ and $\Sigma = \Lambda \otimes D$, from which it follows that Ω is proportional to Λ, as Σ is a matrix multiplication of Λ and Ω is a function of Σ. This points out that there is correspondence between the covariance matrix of the long-run projections $\overline{x}_t$ and $\overline{y}_t$ and the long-run covariance matrix of the first differences of z_t, Δz_t (Müller & and Watson, 2017b). These two examples indicate the strong dependence of Ω on the local-to-zero spectrum $S_z(\omega)$ and Σ.

In order to be able to model persistence in economic time series with integration order $d \in [-0.4; 1]$, Müller and Watson (2018) allow for all local-to-zero spectra of the form

$$S_z(\omega) \propto A \begin{pmatrix} (\omega^2 + c_1^2)^{-d_1} & 0 \\ 0 & (\omega^2 + c_2^2)^{-d_2} \end{pmatrix} A' + BB' \tag{5}$$

The (A, B, c, d) model is an extension of the (B, c, d) model that is introduced in Müller and Watson (2016) in order to compute prediction sets for long-term forecasts of variables. The (A, B, c, d) model

has more parameters, but appears to be more flexible than the $I(0)$ and $I(1)$ models by allowing for more spectra and combining standard spectral shapes. For example, $A = 0$ yields the $I(0)$ model and when $B = 0,\ c = 0$ and $d_1 = d_2 = 1$, the (A, B, c, d) model yields the $I(1)$ spectrum. This flexibility enables the (A, B, c, d) model to allow for various long-run phenomena such as overdifferencing and slow mean reversion (Baillie, Bollerslev & Mikkelsen, 1996). The estimated values of ρ_T, β_T and $\sigma_{y|x,T}$ are medians of the Bayesian posterior density using the $I(d)$-model with integration order d. In spite of discussed flexibility of the (A, B, c, d) model, there are models that are not encompassed by (5). For example, if $B = \mathbf{0}$ and the off-diagonal elements of matrix $\begin{pmatrix} (\omega^2 + c_1^2)^{-d_1} & 0 \\ 0 & (\omega^2 + c_2^2)^{-d_2} \end{pmatrix}$ are not equal to zero, this model falls outside the (A, B, c, d) model. In this calculation, the regression coefficient of regressing Y onto X can be equal to the cointegration coefficient, which can be expected by no reason (Müller & Watson, 2017b).

2.5 Constructing confidence intervals for ρ, β and $\sigma_{y|x}$

In order to describe the uncertainty of the long-run covariability parameters, Müller and Watson (2018) construct confidence intervals. By definition, confidence intervals of level $1-\alpha$ cover the true parameter θ with a probability of $1-\alpha$, for all permitted values of θ. In this paper, parameter vector θ consists of the parameters of the (A, B, c, d) model that characterize the probability distribution of the projection coefficients X_T and Y_T. Let Θ denote the parameter space of $A,\ B,\ c$ and d. The confidence sets are constructed as informative as possible, which means that the intervals can not be tighter without violation of the coverage constraint that the intervals contain the true parameter with a probability of $100^*(1-\alpha)\%$. This is described by the following optimization problem:

$$\begin{aligned} &\min_{H} \int E_\theta[vol(H(X,Y))\ dW(\theta)], \\ &s.t. \quad \sup_{\theta\in\Theta} P_\theta(\gamma \in H(X,Y)) \geq 1-\alpha \end{aligned} \tag{6}$$

In equation (6), γ denotes the parameter vector of interest ρ, β or $\sigma_{y|x}$, $H(X,Y)$ denotes the confidence interval of γ with length $vol(H(X,Y))$. Equation (6) shows that the expected volume of the confidence interval H depends on the weighted values of θ, $dW(\theta)$. This follows by the dependence of the probability distribution of the long-run projection coefficients (X,Y) on θ. The expected length of the confidence interval can be expressed in terms of the power of hypothesis tests of $H_0 : \gamma = \gamma_0$ versus $H_1 : \theta \sim W$. The idea of minimizing the average expected length of some confidence interval by averaging according to some weighting distribution on permitted values of the

parameters is described by Pratt (1961). The minimizing approach on the length of confidence sets that is declared in his paper, is different from the Bayesian approach that is based on a prior observed distribution. A family of most powerful hypothesis tests leads to the optimal confidence interval. The algorithm in Appendix B computes the approximate least favorable distribution (ALFD) in order to determine this family of most powerful tests γ_0. This numerical approach is initiated by Elliott, Müller and Watson (2016) The weighting function $dW(\theta)$ follows the Bayesian prior density in the baseline model. The distribution is, as well as the Bayes prior, based on the bivariate $I(d)$ model with the following parameter restrictions: Set $B = 0,\ c_i = 0,\ d_i \sim iid\ U[-0.4, 1],\ i = 1, 2$ and $A = R(\pi U_1)G(15^{U_0}, 1)R(\pi U_2)$, where $U_j \sim iid\ U[0, 1]$, G a diagonal matrix and $R(\psi)$ a 2×2 rotation matrix of angle ψ. As the number of parameters in the (A, B, c, d) model is large and as the (A, B, c, d) model does not always guarantee correct coverage probabilities of the confidence intervals, Müller and Watson (2018) follow Müller and Norets (2016) when constructing desired confidence intervals. Müller and Norets (2016) propose a method to minimize weighted average length criteria as in equation (6), subject to the inclusion of a Bayesian credible set. This method prevents confidence intervals from being empty or unreasonably short, which might cause situations in which the constructed confidence interval does not contain the true parameter value. The difference between the confidence intervals and the Bayes credible set vary from 3% to 8% (Müller and Norets, 2016).

Finally, the parameter space Θ for parameter vector $\theta = (A, B, c, d)$ is as follows: $A, B \in \mathbb{R}$, B lower-triangular and (A, B) are only allowed if Ω is non-singular. Parameters c and d are restricted such that $c_i \geq 0$ and $-0.4 \leq d_i \leq 1,\ i = 1, 2$. This construction of Θ enables the (A, B, c, d) model to ensure confidence interval coverage for a wide range of persistence patterns if the least favorable distribution, the weighting distribution W and the Bayes prior distribution are given. The invariance/equivariance restrictions that are displayed in Appendix B are the second specification on the constructed confidence intervals. These restrictions require use of maximal invariants are used instead of the original (X,Y) and a slight modification on the form of the optimal test statisctic explained by Müller and Norets (2016).

2.6 Simulation

The unknown integration order of the used data series is the main argument for the use of the (A, B, c, d) model, that has the flexibility to account for fractional integration orders. The simulation results in section 3.1 investigate the patterns that the confidence intervals follow for dif-

ferent integration orders and three different models: the $I(0)$ model ($S_z \propto BB'$), the $I(1)$ model ($S_z(\omega) = \omega^{-2}AA'$) and the (A, B, c, d) model with $S_z(\omega)$ given by (5). For every fractional integration order $d \in [-0.5, 1.5]$ with one or no decimal places, such that $d = -0.5, -0.4, -0.3, \ldots, 1.4, 1.5$, 2000 random fractionally integrated time series are simulated according to the $ARFIMA(p, d, k)$ model. The general $ARFIMA(p, d, k)$ model, explained by Haldrup and Eduardo Vera Valdes(2017) and Baillie, Chong & Tieslau (1996), is described by integration order d and orders of the autoregressive and moving average part p and k respectively:

$$\Phi(L)(1-L)^d(y_t - \mu) = K(L)\epsilon_t,\ \epsilon_t \sim iid(0, \sigma_\epsilon^2), \tag{7}$$

where L denotes the backward-shift operator, $\Phi(L)$ and $\Theta(L)$ are given by $\Phi(L) = 1 - \phi_1 L - \ldots$-$\phi_p L^p$, $\Theta(L) = 1 + \kappa_1 L + \ldots + \kappa_q L^q$ and $(1-L)^d$ represents the fractional differencing operating operator which is in turn given by $(1-L)^d = \sum_{s=0}^{\infty} \frac{\Gamma(s+d)L^s}{\Gamma(d)\Gamma(s+1)}$. μ and Γ are defined as some constant and the Gamma Function respectively.

The data series that are used in the simulation, are simulated with the following parameter values: the moving-average part of the $ARFIMA$ model is neglected, so that $k = 0$, and the autoregressive order of the processes is equal to 1, with $\phi_1 = 0.2$. As described above, these $ARMA(1, 0)$ processes are simulated for every order of integration $d = -0.5, -0.4, -0.3, \ldots, 1.4, 1.5$.

After the simulation of the 2000 time series for every integration order d, the estimates of ρ, β and $\sigma_{y|x}$ and their empirical confidence sets and those that are estimated by the methods used by Müller and Watson (2018), are calculated following the simulation algorithm in Appendix C. This algorithm is executed three times. First, the confidence intervals of the simulated data are obtained using the $I(0)$ model. Second, the $I(1)$ model is used and third, the general (A, B, c, d) model is used. As the (A, B, c, d) is considered as the main contribution of Müller and Watson (2018), there will be focused on the results obtained using this model. The results that are obtained from both the $I(0)$ and the $I(1)$ model have not been worthless at all, as the results from these models emphasize the importance of the (A, B, c, d) model as a model that accounts for a wide range of persistence patterns. In addition, these results have contributed to my understanding on the dependence between the fractional integration order d in data series and the output of the three models used by Müller and Watson (2018).

3 Results

The results section contains all findings of this paper and consists of three different parts. Section 3.1 contains the results of the simulation that is explained in section 2.6. Section 3.2 contains replications of the findings of Müller and Watson (2018) and Section 3.3 proposes the extension results that are found by application of the replication methods on the relationship between the S&P 500 Index and S&P 500 Futures.

3.1 Simulation results

The results of the simulation are presented in Appendix D, containing all figures and tables that are discussed in the this subsection. Tables 3, 4 and 5 contain the average values, empirical confidence intervals, and average confidence intervals estimated by the methods used by Müller and Watson (2018) for parameters ρ, β and $\sigma_{y|x}$ and every integration order d between -0.5 and 1.5. Empirical confidence intervals in these tables differ across the models, as the parameter estimates depend on the used model, which is also demonstrated in the replication part of this paper. Figures 3, 4 and 5 provide visualizations of the differences between the empirical 90% confidence intervals and the average 90% confidence intervals that are estimated by the (A, B, c, d) model from Müller and Watson (2018).

Table 3 presents the simulation results for long-run covariability parameter ρ. The results that are obtained by use of the $I(0)$ and the $I(1)$ model can be found in columns 2-7. From these columns, it is noted that the confidence intervals that are estimated by the $I(0)$ and the $I(1)$ model are closest to the empirical confidence intervals when the integration order of the simulated process is equal to zero for the $I(0)$ model and equal to one for the $I(1)$ model. Differences between the empirical confidence intervals and the average estimated confidence intervals become larger when the order of integration of the simulated processes deviates from zero (when using the $I(0)$ model) or one (when using the $I(1)$ model). When the (A, B, c, d) model is used, this pattern is not observed, as the average estimated confidence intervals by the (A, B, c, d) model are close to the corresponding empirical confidence intervals for almost all orders of integration of the simulated processes. When the integration order of the simulated processes d exceeds 1.2, the performance of the (A, B, c, d) becomes worse, as the differences between the empirical and estimated confidence intervals become larger. This is also observed from figure 3. Furthermore, table 3 shows that the average confidence intervals that are estimated by the (A, B, c, d) model are wider than the empirical

confidence interval, which can be a consequence of the restriction on the width of the estimated confidence intervals to be supersets of the Bayes credible sets.

The simulation results for β in table 4 show that the average confidence intervals that are estimated by the (A, B, c, d) model differ most from the corresponding empirical confidence intervals when the integration order of the simulated data series d is between 0.5 and 1. This can also be observed from figure 4. Overall, the differences between the empirical confidence intervals and the average estimated confidence intervals of β are smaller and more constant using the (A, B, c, d) model than when the $I(0)$ model or the $I(1)$ model is used.

Finally, the simulation results of $\sigma_{y|x}$ are presented in table 5 and figure 5. As $\sigma_{y|x}$ is proportional to the values of the data, it is more difficult to draw conclusions on the relationship between the integration order of the simulated series d and the confidence intervals. First, we observe that there are higher estimates of $\sigma_{y|x}$ obtained when the (A, B, c, d) model is used than when the $I(0)$ or the $I(1)$ model is used. Second, we observe that the differences in terms of percentage between the empirical confidence intervals and the average confidence interval obtained using the (A, B, c, d) model, explode when the integration order of the simulated data series d exceeds 0.6. This is mainly caused by the estimated upper bounds that are higher than the corresponding empirical upper bounds.

3.2 Replication results

This subsection contains replications of the findings from Müller and Watson (2018). First, the figures in Müller and Watson (2018) are replicated that present applied data transformations. Second, the tables that display all long-run covariability parameter estimates and confidence intervals are replicated. This is done by application of three models: results of the $I(0)$ and $I(1)$ model are presented in table 1 and table 2 contains the results of the (A, B, c, d) model.

Data transformations

Figure 1a and 1b below show the average growth rates of GDP and consumption over six subsamples between January 1948 and December 2015, each containing 45 observations. The captured periodicities are equal to or longer than 11.3 years and therefore longer than the U.S. business cycle of 32 quarters that is introduced by Baxter and King (1999). From the figures below, it is observed that GPD and consumption in the U.S. show approximately equal patterns, as growth rates of variables are relatively high in the first decades after World War II and show sharp decreases during

the third sub-sample from 1970 to 1981. After this decrease, GDP growth increases less than the consumption growth rate, but during the sixth sub-sample, both growth rates are approximately equal again.

Figure 1c plots the long-run projections of the data onto the low-frequency cosine terms. The moving averages in this sub-figure are designed to isolate data variation with periods longer than eleven years. The long-run projections of GDP and consumption are more closely related in figure 1c than in figure 1a. This is confirmed by comparison of the scatterplot of the long-run projections and projection coefficients in figure 1d with the scatterplot in figure 1b.

If the number of sub-samples is decreased to four, figure 6b in Appendix E shows that the relationship between GDP and consumption is less close during the period from 1982 to 1998 than in the other time periods. Increase in the number of sub-samples to eight shows that the difference is largest from 1982 to 1990 (figure 7, Appendix E). The early 1980s recession is propably a reason for the relatively large difference in this sub-period. Figure 8 and figure 9 in Appendix E show plots of the long-run projections and scatterplots of the long-run projections and projection coefficients for $q = 8$ and $q = 16$ respectively. If $q = 16$, periodicities longer than 8.5 years are captured, which is approximately equal to the U.S. business cycle of 32 quarters. The increase of q results in a relationship that is closer than the relationship for $q = 12$. If $q = 8$, the model captures periodicities that are longer than 17 years and the plot in figure 8 in Appendix E again shows higher differences between GDP and consumption growth rates from 1982 to 1994.

Figure 2a below plots the short-term interest rates of three months and the long-term interest rates of ten years and the long-run projections of both interest rates. Due to a shorter data period, q is set equal to 11 in order to capture periodicities longer than 11.3 years. Figure 2a shows higher long-run projections of long-term interest rates than short-term interest rates projections, which might be related to the increase in interest rates before 1981 and the interest decrease after 1981. Figure 2b displays high correlation between the projection coefficients X_{jT} and Y_{jT} of the short- and long-term interest rates. In order to analyze the sensitivity of the relation between the interest rates, q is adjusted and figures 10 and 11 in Appendix E show that the relation between the long-run projection coefficients becomes closer if q increases.

Estimation by the $I(0)$ and $I(1)$ model

Table I and II from Müller and Watson (2018) are replicated below. All replicated values that do not correspond with the results from Müller and Watson (2018) are coloured in red. Table 1 below

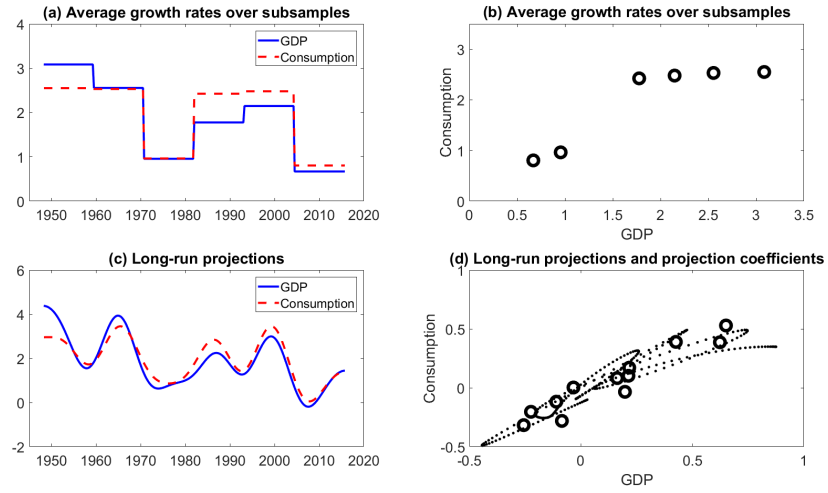


Figure 1: Panel (a) plots sample averages of consumption and GDP over np sub-periods, $np = 6$. Panel (c) shows the data projections onto the low-frequency terms with added sample means, $q = 12$. Panels (b) and the small dots in (d) are scatterplots of the variables in (a) en (c), while the large circles in (d) are of the projection coefficients (X_{jT}, Y_{jT}) from (c).

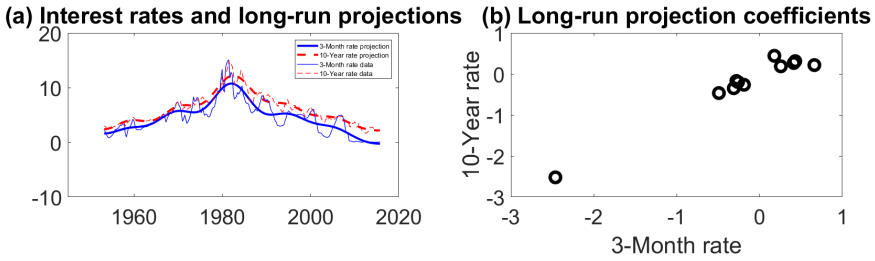


Figure 2: Panel (a) plots the projections of the short- and long-term interest rates onto the low-frequency cosine terms with $q = 11$ with added sample means. Panel (b) is a scatterplot of the projection coefficients (X_{jT}, Y_{jT}) from (a).

shows estimates and confidence intervals for long-run covariability parameters ρ_T, β_T and $\sigma_{y|x,T}$ using the $I(0)$ and $I(1)$ models with $q = 12$ (panel a) and $q = 11$ (panel b). All values in table 1 are equal to the corresponding values in the paper of Müller and Watson (2018). The estimate of the long-run correlation between GDP and consumption is equal to 0.93 in both the $I(0)$ and the $I(1)$ model and the 90% confidence intervals fall between 0.8 and 1.0 for both models. Regressing consumption onto GDP results in a higher regression coefficient when the $I(1)$ model is used than when the $I(0)$ model is used.

Short-term and long-term interest rates are more highly correlated than GDP and consumption according to the long-run correlation coefficients that follow from the $I(0)$ and $I(1)$ models. The

lower range of the 90% confidence interval of ρ exceeds 0.9 in both models, indicating a high long-run correlation between the interest rates. Contrary to table 1a, regression coefficients by use of the models are approximately equal.

Decreasing q to 8 for GDP and consumption and 7 for the interest rates captures only periodicities that are longer than 17 years. Table 6 and 7 in Appendix E present the results of the $I(0)$ and $I(1)$ models after decrease and increase of q. Decrease of q results in higher estimated long-run covariability between GDP and consumption, while the long-run covriability between the short-term and long-term interest rates remains equal. Increase of q causes lower rong-run coavriability in both investigated relationships. The width of the constructed confidence intervals does not follow a clear dependence on the value of q.

Table 1: Replication of the long-run covariability estimates and confidence intervals using the $I(0)$ and $I(1)$ models: periods longer than 11 years

		ρ	β	$\sigma_{y\|x}$
		(a) GDP and consumption		
I(0)	Estimate	0.93	0.76	0.36
	67% CI	0.87, 0.96	0.67, 0,85	0.30, 0.46
	90% CI	0.80, 0.97	0.60, 0.92	0.27, 0.55
I(1)	Estimate	0.93	0.84	0.35
	67% CI	0.88, 0.96	0.74, 0.94	0.29, 0.45
	90% CI	0.82, 0.97	0.66, 1.01	0.26, 0.54
		(b) Short- and long-term interest rates		
I(0)	Estimate	0.98	0.96	0.60
	67% CI	0.96, 0.98	0.90, 1.03	0.50, 0.79
	90% CI	0.93, 0.99	0.84, 1.08	0.44, 0.96
I(1)	Estimate	0.97	0.93	0.38
	67% CI	0.93, 0.98	0.85, 1.01	0.32, 0.50
	90% CI	0.90, 0.98	0.78, 1.07	0.28, 0.61

Periods longer than 11 years correspond to q = 12 for panel (a) and q = 11 for panel (b). The estimates are the maximum likelihood estimates found by the large-sample distribution of the cosine transforms for the $I(0)$ and $I(1)$ models. 67% and 90% confidence intervals are displayed below the estimates. Values that do not correspond with Mueller and Watson (2018), are coloured in red.

(A, B, c, d) model

Table 2 below contains the replication of table II from Müller and Watson (2018). The estimated 67% confidence intervals for ρ_T in table 2 differ slightly from the estimated 67% confidence intervals in Müller and Watson (2018), as the lower bound is equal to 0.83 in Müller and Watson (2018). As only individual comments are added to the code from Müller and Watson (2018) that replicates this table, this difference is assumed to be a consequence of rounding mistakes in the original paper. All the other values in table 2 correspond with the values in Müller and Watson (2018).

The values on the long-run covariability between GDP and consumption in table 2 are closer

to those estimated by the $I(0)$ model than those that are estimated by the $I(1)$ model. Although the estimate of the long-run correlation differs only slightly from corresponding estimates in table 1, the confidence intervals that are estimated by the (A, B, c, d) model are wider. This observation of wider confidence intervals, which is also observed for β and $\sigma_{y|x}$, is in accordance with what is concluded in the simulation section of this paper. Furthermore, only the 90% confidence interval and 90% Bayes credible set of the regression coefficient are different.

For the relationship between short-term interest rates and long-term interest rates, again wider confidence intervals are constructed by the (A, B, c, d) model than in table 1, as expected from the simulation. The estimate of the correlation, regression coefficient and conditional variance have only changed slightly to 0.96, 0.95 and 0.63 respectively. The confidence intervals for β and $\sigma_{y|x}$ are wider than corresponding Bayes credible sets, which indicates that not restricting the confidence sets to be supersets of the Bayes credible sets would result into tighter confidence interval for ρ.

Table 8 in Appendix E shows the values of the long-run covariability parameters, confidence sets and Bayes credible sets, using the (A, B, c, d) model $q = 8$ (table 8a) and $q = 7$ (table 8b). By the decrease of q, the (A, B, c, d) model captures only periodicities that are longer than 17 years. The long-run correlation coefficient increases in both relationships when q decreases, while the regression coefficient of regressing consumption onto GDP remains approximately equal. When comparing the width of the confidence intervals and Bayes credible sets, all confidence sets and intervals are wider if $q = 12$ for the interest rates, while the confidence intervals for ρ and Bayes credible sets of ρ, β and $\sigma_{y|x}$ for GDP and consumption are wider if $q = 8$.

Table 2: Replication of the long-run covariability estimates, confidence intervals and credible sets using the (A, B, c, d) model: periods longer than 11 years

		ρ	β	$\sigma_{y\|x}$
		a) GDP and consumption		
(A, B, c, d)	Estimate	0.91	0.77	0.41
	67% CI	0.84, 0.96	0.66, 0.87	0.33, 0.53
	90% CI	0.71, 0.97	0.48, 0.96	0.29, 0.66
	67% Bayes CS	0.84, 0.96	0.66, 0.87	0.33, 0.53
	90% Bayes CS	0.71, 0.97	0.58, 0.96	0.29, 0.66
		b) Short- and long-term interest rates		
(A, B, c, d)	Estimate	0.96	0.95	0.63
	67% CI	0.92, 0.98	0.87, 1.07	0.49, 0.97
	90% CI	0.89, 0.99	0.76, 1.16	0.42, 1.27
	67% Bayes CS	0.92, 0.98	0.87, 1.03	0.49, 0.82
	90% Bayes CS	0.89, 0.99	0.81, 1.09	0.42, 1.02

Periods longer than 11 years correspond to $q = 12$ for panel (a) and $q = 11$ for panel (b). The estimates are the posterior median based on the $I(d)$ model. The 67% and 90% confidence intervals and Bayes credible sets based on the posterior are displayed below the estimates. Values that do not correspond with Mueller and Watson (2018), are coloured in red.

3.3 Extension results

This subsection presents the results of the investigation on the long-run covariability between the S&P 500 Index and the S&P 500 Futures and the extension is structured in three sub-analyses.

Analysis of the impact of different data frequencies

In the analysis of the impact of different data frequencies, q is set equal to twelve for the monthly-weekly analysis and equal to ten for the weekly-daily analysis. These values of q ensure that the (A, B, c, d)-model captures periodicities longer than the S&P 500 cycle, equal to 3.5 years, and one year respectively. As expected from figures 12, 13, 14d and 15 (Appendix F), tables 9 and 10 (Appendix G) show that the S&P 500 Index and the S&P 500 Futures based on daily, weekly and monthly observed data are highly correlated. Effects of the use of weekly data between November 1997 and June 2019, instead of monthly data, consist of slight decreases of the conditional prediction error and regression coefficient of regressing the S&P 500 Futures on the S&P 500 Index. The use of daily data instead of weekly data between January 2014 and June 2019 causes a decrease of estimated long-run correlation coefficient ρ and slight increases of the regression coefficient and conditional prediction error. β and ρ remain higher than or equal to 0.9 with probability of 90% despite of described changes. As the effects of the use of daily data instead of weekly data are not in line with the effects of the use of weekly data instead of monthly data, there are no clear dependencies observed between considered long-run covariability parameters and data frequencies.

Sub-sample analysis

In order to obtain more information on the relationship between the S&P 500 Index and the S&P 500 Futures, the long-run covariability between both investing market indicators based on weekly and daily data is investigated on different sub-periods. Figures 14 (weekly data) and 16 (daily data) in Appendix F plot the S&P 500 Index and S&P 500 Futures growth rates and corresponding long-run projections and projection coefficients over different sub-samples. Tables 11 and 12 in Appendix H display the parameter values and confidence intervals that are obtained using the (A, B, c, d) model. In order to visualize the patterns that the estimates of the long-run covariability parameters follow, the parameter estimates are plotted in figure 23 in Appendix H based on a rolling window with steps equal to one week and one day. In this figure, most recent sub-sample of 282 weekly observations or 271 daily observation between November 1997 and June 2019 (weekly data) and January 2014 and

June 2019 (daily data) is displayed left, and least recent sub-samples in named time intervals are displayed right. For this analysis, q is set equal to ten, which ensures that periodicities longer than 1.09 years are captured in the weekly analysis and periodicities longer than 0.21 years are captured in the daily analysis.

The estimate of the long-run correlation parameter does not change during the four constructed sub-periods, but the tighter confidence intervals from 24 August 2008 to 12 January 2014 indicate a slightly higher correlation during this sub-period. The tight confidence intervals that are constructed for all three parameters in the third sub-period are explained by figures 23a, 23c and 23e in Appendix H, considering the relative low volatility of all three parameter estimates between October 1999 and January 2012.

In order to analyze the covariability patterns during the last weekly sub-sample from January 2014 to June 2019, the daily data during this period is divided into 5 sub-periods. The correlation coefficient remains approximately constant during the first four sub-periods, but decreases sharply during the last sub-period. The regression coefficient and conditional prediction error of regressing consumption onto GDP during the fifth sub-sample. These observations are confirmed by figures 23b, 23d and 23f in Appendix H. The change in parameter values, occurred on Christmas Eve 2018, is a consequence of what is called “The worst December since the Great Depression for stocks" (Isidore, 2018). The crash of the American stock market has had higher long-run covariability parameter volatilities from this moment as a consequence.

Approximation of the lead time of the S&P 500 Futures

To investigate whether the (A, B, c, d) model can be used to estimate the lead time of the S&P 500 Futures on the S&P 500 Index, there is made use of data frequencies that are equal to one hour, five minutes and one minute. For each investigated data frequency, the long-run covariability parameters are calculated on the relationship between the growth rates of the S&P 500 Index and the S&P 500 Futures at equal points in time. After this, the S&P 500 Futures growth rates are shifted backward n steps in the data set, and the long-run covariability parameters on the relationship between the growth rates of the S&P 500 Index and the S&P 500 Futures of n steps ahead are calculated. Data plots for this sub-analysis can be found in Appendix F, while the tables that contain all parameter values, confidence intervals and Bayes credible sets can be found in Appendix I. For this sub-analysis, q is kept equal to ten, resulting in captured periodicities of minimum length of 8.24 trading days (hourly data), 269 trading minutes (data per five minutes) and 54 trading minutes (minutely data).

The long-run parameter values and confidence intervals on the relationship between the hourly growth rates of the S&P 500 Index and the S&P 500 Futures are shown in table 13. The estimates of the long-run correlation coefficient and the regression coefficient are equal for the relationship between the S&P 500 Futures of one hour ahead and the S&P 500 Index and the relationship between the investing market indicators observed at the same time. The covariability between the S&P 500 Index and the S&P 500 Futures of two hours ahead shows lower parameter estimates and as a consequence of that, the constructed confidence intervals are wider. Results of the investigation of the long-run covariability between the growth rates of the S&P 500 Index and the S&P 500 Futures per five minutes are shown in table 14. From this table, a clear decreasing pattern of covariability is observed when the lead of the S&P 500 Futures increases. The estimates and confidence intervals of long-run correlation parameter ρ and regression coefficient β are highest when the lead time of the S&P 500 Futures is equal to zero minutes. The conditional prediction error is estimated equal to zero in every panel of table 14, as the growth rates of the S&P 500 Index and the S&P 500 Futures have become too small for the 5 minutes and minutely data frequencies.

Lastly, the relationship between the minutely growth rates of the S&P 500 Index and the S&P 500 Futures is investigated. According to the estimated parameter values, the minutely growth rates of the S&P 500 Index and S&P 500 Futures are more highly correlated than the growth rates per five minutes. Table 15 shows that correlation parameter estimates of the long-run correlation and the regression coefficient are highest when the lead time of the S&P 500 Futures is equal to zero minutes. The confidence intervals of these parameters follow the pattern that is observed before and become wider as the lead time of the S&P 500 Futures increases.

As there are multiple relationships of different data frequencies and with different lead times of the S&P 500 Futures with approximately equal long-run covariability parameter estimates and confidence intervals, no conclusion can be drawn on the lead time of the S&P 500 Futures by the use of the (A, B, c, d) model.

4 Conclusion and discussion

This paper builds on previous research by Müller and Watson (2018), who propose methods that make use of small numbers of approximately normally distributed low-frequency weighted averages to construct asymptotically efficient confidence intervals for long-run covariability parameters. The approximately normally distributed low-frequency weighted averages solve the problems of a paucity

of sample inference and of inference that critically depends on the long-run persistence of the data. As the (A, B, c, d) model is defined as a flexible parametrization of the so-called local-to-zero spectrum, confidence intervals can be constructed for a wide range of persistent processes. The simulation section of this paper shows the accurate form of the constructed confidence intervals that are obtained from the (A, B, c, d) model for parameters ρ and β. Furthermore, in the replication of figures I and II and tables I and II from Müller and Watson (2018), no important deviations are found.

The performance of the methods that are proposed by Müller and Watson (2018) on higher data frequencies is measured on the basis of the relationship between the S&P 500 Index and the S&P 500 Futures. First, it is demonstrated that monthly, weekly and daily S&P 500 Index and S&P 500 Futures growth rates lead to different long-run covariability parameter values, but these values do not follow a clear pattern as the data frequency increases. From the next sub-sample analysis, it is derived that the U.S. stock market crash in December 2018 has reduced the long-run covariability between the S&P 500 Index and the S&P 500 Futures. It is concluded that the differences between sub-periods are captured and covariability crashes can be traced by the long-run covariability estimates obtained from the (A, B, c, d) model, and the rolling-window figures that are plotted from these estimates. Finally, the use of the (A, B, c, d) model in the final sub-analysis does not result into a clear estimation of the lead time of the S&P 500 Futures on the S&P 500 Index, as multiple relationships with different Futures lead times and data frequencies are estimated to be approximately equally correlated in the long-run.

In future research, the performance of the methods used by Müller and Watson (2018) on high frequency data can be further examined by consideration of other relationships that can be observed at high frequencies. Consideration of relationships that are less highly correlated, such as exchange rates or cryptocurrencies, might result into different or more observable differences in parameter estimates and confidence intervals than observed in this extension. Besides that, the approximation of the lead time of the S&P 500 Futures on the S&P 500 Index might be possible for researchers that have more time to perform this analysis and have access to larger data sets of intraday data on the S&P 500 Index and the S&P 500 Futures. When thinking of extensions on the models that are proposed by Müller and Watson (2018), development of a multivariate extension of the bivariate (A, B, c, d) would be a particularly interesting and challenging addition to the existing inference.

Bibliography

Arshanapalli, B. and Doukas, J. (1994). Common volatility in S&P 500 stock index and S&P 500 index f Prices during October 1987, *The Journal of Futures Markets, 14*, 915-925.

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

Baillie, R. T., Chung, C. F., Tieslau, M. A. (1996). Analysing inflation by the fractionally integrated ARFIMA–GARCH model. *Journal of applied econometrics, 11(1)*, 23-40.

Baxter, M., and King, R.G. (1999). Measuring business cycles, approximate band-pass filters for economic time series," *Review of economics and statistics, 81(4)*, 575-593

Campbell, J.Y. (1987). Does savings anticipate declining labor income? An alternative test of the permanent income hypothesis,, *Econometrica, 55*, 1249-1273.

Campbell, J.Y. (1991). The response of consumption to income - A cross-country investigation, *European Economic Review, 35*, 723-767.

Campbell, J.Y. and Shiller, R.J. (1987). Cointegration and tests of present Value models, *Journal of Political Economy, 95*, 1062-1088.

Chong, T.T., Li, Z., Chen, H. and M.J.Hinich (2010). An investigation of duration dependence in the American stock market cycle, Journal of Applied Statistics, 37, 1407-1416.

Chong, T. T. and Lui, G.C. (1999). Estimating the fractionally integrated process in the presence of measurement errors, *Economics Letters, 63*, 285-294.

Cochrane, J. (1994). Permanent and transitory components of GNP and stock prices, *Quarterly Journal of Economics, 109*, 241-266.

Dai, Q. and Singleton, K.J. (2002). Specification analysis of affine term structure models, *The Journal of Finance, 55(5)*, 1943-1978.

Elliott, G., Mueller, U.K. and Watson, M.W. (2015). Nearly optimal tests when a nuisance parameter is present under the null hypothesis, *Econometrica, 83*, 771-811. [788,789]

Engle, R. F. (1974). Band spectrum regression, *International Economic Review, 15*, [776,781]

Granger, C.W.J. and Newbold, P. (1973). Spurious regressions in econometrics", *Journal of Econometrics, 2*, 111-120

Haldrup, N. and Eduardo Vera Valdes, J. (2017). Long memory, fractional integration, and cross-sectional aggregation", *Jounal of Econometrics, 199*, 1-11

Hakkio, C. (1991). Cointegration, how short is the long run?, *Journal of International Money and Finance, 10*, 571-581.

Harvey, A.C. (1978). Linear Regression in the frequency domain, *International Economic Review, 2*, 507-512

Herbst, A.F., McCormack, J.P. and West, E.N. (1987). Investigation of a Lead-Lag relationship between spot stock indices and their futures contracts, *The Journal of Futures Markets, 7*, 373-381.

Hounyo, U. (2017). Bootstrapping integrated covariance matrix estimators in noisy jump–diffusion models with non-synchronous trading. *Journal of econometrics, 197(1)*, 130-152.
Isidore, C. (2018). 2018 was the worst for stocks in 10 years. Retrieved from https://edition.cnn.com/2018/12/31/investing/dow-stock-market-today/index.html

Johansen, S. (1988). Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control, 12*, 231-254.

Kawaller, I.G., Koch P.D. and Koch, T.W. (1987). The temporal price relationship between S&P 500 Futures and the S&P 500 Index, *The Journal of Finance, 42*, 1309-1329.

_____ (1990). Intraday relationships between volatility in S&P 500 futures prices and volatility in the S&P 500 index, *Journal of Banking & Finance, 14*, 373-397.

Klein, L.R., and Kosobud, R.F. (1961). Some econometrics of growth, Great Ratios of economics, *Quarterly Journal of Economics, 75*, 173-198.

Lettau, M. and Ludvigson, S.C. (2013). Shocks and crashes, *NBER Macroeconomics Annual, 28(1)*, 293-354.

Mueller, U.K. and Norets, A. (2016). Credibility of confidence sets in nonstandard econometric problems, *Econometrica, 84*, 2183–2213

Mueller, U.K. and Watson, M.W. (2008). Testing models of low-frequency variability, *Econometrica, 84*, 2183-2213.

_____ (2016). Measuring uncertainty about long-run predictions, *Review of Economic Studies, 83*, 1711-1740. [785,7920].

_____ (2017a). Low-frequency econometrics, *Advances in Economics, ekeventh World Congress of the Econometric Society*, ed. by Honoré, B. and Samuelson, L. Cambridge, Cambridge University Press. [782,283,802].

_____ (2017b). Long-run covariability. Working paper, 2017.

_____ (2018). Long-run covariability, *Econometrica Supplemental Material, 86*, https,//doi.org/10.3982/ECTA15047. [777]

_____ (2018). Long-run covariability. Econometrica, 86(3), 775-804.

Onatski, A. and Wang, C. (2018). Spurious factor analysis, *Faculty of Economics, University of Cambridge*.

Phillips, P. (1986).“Understanding spurious regressions in econometrics, *Journal of Econometrics, 33*, 311-340

_____ (1991). Optimal inference in cointegrated systems, *Econometrica, 59*, 283-306. Stock, J.H.,

and Watson, M.W. (1993). A simple estimator of cointegrating vectors in higher order integrated systems, *Econometrica, 61*, 783-820.

Valdovinos, C. G. F. (2003). Inflation and economic growth in the long run. Economics Letters, 80(2), 167-173.

Appendix A. Data sources

Replication data

For the replication part of the paper, data sources and definitions are listed per variable below. The first list contains all variables that are monthly observe and the second list contains all variables that are observed on a monthly frequency.

Monthly observed variables

- 10-Year Treasury Constant Maturity Rate (GS10): obtained from FRED. Observed from 1953-04-01 to 2016-11-01.
- 3-Month Treasury Bill: Secondary Market Rate (TB3MS): obtained from FRED. Observed from 1947-01-01 to 2016-11-01.

Quarterly observed variables

- Gross Domestic Product (GDP): obtained from FRED and deflated by the price index for personal consumer expenditures (PCECTPI, obtained from FRED). Observed from 1947-01-01 to 2016-07-01.
- Personal Consumption Expenditures: Durable Goods (PCDG): obtained from FRED and deflated by the price index for personal consumer expenditures (PCECTPI, obtained from FRED). Observed from 1947-01-01 to 2016-07-01.
- Personal Consumption Expenditures: Chain-type Price Index (PCECTPI): obtained from FRED. Observed from 1947-01-01 to 2016-07-01.
- Personal Consumption Expenditures: Services (PCESV): obtained from FRED and deflated by the price index for personal consumer expenditures (PCECTPI, obtained from FRED). Observed from 1947-01-01 to 2016-07-01.
- Personal Consumption Expenditures: Nondurable Goods (PCND): obtained from FRED and deflated by the price index for personal consumer expenditures (PCECTPI, obtained from FRED). Observed from 1947-01-01 to 2016-07-01.

- Private Nonresidential Fixed Investment (PNFI): obtained from FRED and deflated by the price index for personal consumer expenditures (PCECTPI, obtained from FRED). Observed from 1947-01-01 to 2016-07-01.
- Private Residential Fixed Investment (PRFI): obtained from FRED and deflated by the price index for personal consumer expenditures (PCECTPI, obtained from FRED). Observed from 1947-01-01 to 2016-07-01.

Extension data

Data on the &P 500 Index (SPX) and the S&P 500 Futures (SPXF) are obtained from the investing database of Investing.com (at https://nl.investing.com/indices/us-spx-500 and https://nl.investing.com/indices/us-spx-500-futures respectively). Data frequencies and time ranges that are observed are listed below:

- Observed at a monthly frequency from November 1997 to June 2019
- Observed at a weekly frequency from 02/11/1997 to 09/06/2019
- Observed at a daily frequency from 21/01/2014 to 10/06/2019
- Observed at a hourly frequency from 24/04/2019 4 pm to 18/06/2019 4 pm (UCT +2:00)
- Observed at a 5 minutely frequency from 12/06/2019 7:05 pm to 17/06/2019 21:55 pm (UCT +2:00)
- Observed at a minutely frequency from 17/06/2019 4:30 pm to 17/06/2019 20:59 pm (UTC +2:00)

All data on the S&P 500 Index and the S&P 500 Futures is only observed on trading days from 3:30 pm to 10 pm (UTC +2:00).

Appendix B. Additional formulas on methodology

Least favorable distribution algorithm

The basics of the algorithm to find the least favorable distribution Λ^*:

1. Let $\Theta_c = \{\theta_1, \ldots, \theta_m\}$ be candidate set for support of Λ.
2. Compute ALFD weights and **cv** such that size of test is equal to $\alpha - \epsilon$ on Θ_c.
3. Check size control on Θ
   - If size is controlled, we are done (compare length to bound generated from $\mathbf{cv}'$ that induces Λ-weighted size that is equal to α.
   - If size violated, add violating θ to Θ_c and go to step 2.
4. The algorithm executes when 500 consecutive Broyden–Fletcher–Goldfarb–Shanno searches do not violate the size of Θ.

Invariance and equivariance restrictions on the parameter confidence intervals

Let H^ρ, H^β and H^σ denote confidence intervals for ρ, β and $\sigma_{y|x}$, then the confidence intervals are restricted as follows (Müller and Watson, 2018):

$$\rho \in H^\rho(X, Y) \Leftrightarrow \rho \in H^\rho(b_x X, b_y Y) for b_x b_y > 0 \tag{8}$$

$$\beta \in H^\beta(X, Y) \Leftrightarrow \frac{b_y \beta + b_{yx}}{b_x} \in H^\beta(b_x X, b_y Y) \text{ for } b_x, b_y \neq 0 \text{ and all values of } b_{yx} \tag{9}$$

$$\sigma_{y|x} \in H^\sigma(X, Y) \Leftrightarrow |b_y| \sigma_{y|x} \in H^\sigma(b_x X, b_y Y + b_{yx} X) \text{ for } b_x, b_y \neq 0 \text{ and all values of } b_{yx}. \tag{10}$$

Appendix C. Simulation algorithm

Simulation algorithm:

- for every $d = -0.5 : 0.1 : 1.5$ $(d = -0.5, -0.4, -0.3, \ldots, 1.4, 1.5)$

  1. Simulate 2000 $I(d)$ time series with 168 observations of the form:

$$(1 - 0.2y_{t-1})\Big(\sum_{s=0}^{\infty} \frac{\Gamma(s+d)L^s}{\Gamma(d)\Gamma(s+1)}\Big)(y_t - \mu) = \epsilon_t, \tag{11}$$

     with L the backward-shift operator, $\epsilon_t \sim iid(0, \sigma_\epsilon^2$ Γ denoting the Gamma function and μ some constant.

  2. Assign 1000 random pairs of time series integrated of order d $(x, y)_i$, $i = 1, \ldots, 1000$

  3. For every parameter $\gamma = \rho$, $\gamma = \beta$ and $\gamma = \sigma_{y|x}$

     (a) Construct following 5 vectors with 1000 elements: $\gamma_{\mathbf{estimate}}$, $\gamma_{\mathbf{LB},\mathbf{67}\%}$, $\gamma_{\mathbf{UB},\mathbf{67}\%}$, $\gamma_{\mathbf{LB},\mathbf{90}\%}$ and $\gamma_{\mathbf{UB},\mathbf{90}\%}$.

     (b) For $i = 1, \ldots, 1000$

        i. Algorithm returns γ_i, $\gamma_{LB^{MW}_{67\%,i}}$, $\gamma_{UB^{MW}_{67\%,i}}$, $\gamma_{LB^{MW}_{90\%,i}}$ and $\gamma_{UB^{MW}_{90\%,i}}$, where γ_i denotes the estimated long-run covariability parameter for data pair i and $\gamma_{LB^{MW}_{67\%,i}}$, $\gamma_{UB^{MW}_{67\%,i}}$, $\gamma_{LB^{MW}_{90\%,i}}$ and $\gamma_{UB^{MW}_{90\%,i}}$ denote the lower (LB) and upper (UB) bounds of the confidence intervals for data pair i of the level that is added as subscript, estimated by the methods of Mueller and Watson (2018).

        ii. Save γ_i, $\gamma_{LB^{MW}_{67\%,i}}$, $\gamma_{UB^{MW}_{67\%,i}}$, $\gamma_{LB^{MW}_{90\%,i}}$ and $\gamma_{UB^{MW}_{90\%,i}}$ as $i-th$ elements of corresponding 1000×1 vectors $\gamma_{\mathbf{estimate}}$, $\gamma_{\mathbf{LB},\mathbf{67}\%}$, $\gamma_{\mathbf{UB},\mathbf{67}\%}$, $\gamma_{\mathbf{LB},\mathbf{90}\%}$ and $\gamma_{\mathbf{UB},\mathbf{90}\%}$.

        iii. End

     (c) Sort $\gamma_{\mathbf{estimate}}$ such that $\gamma_{\mathbf{estimate}}(1) \leq \gamma_{\mathbf{estimate}}(2) \leq \ldots \leq \gamma_{\mathbf{estimate}}(1000)$

     (d) Compute parameters:

        - $\hat{\gamma} = mean(\gamma_{\mathbf{estimate}})$:
          Estimated average long-run parameter of 1000 simulated pairs of series of integration order d
        - $\widehat{\gamma_{LB^{MW}_{67\%}}} = mean(\gamma_{\mathbf{LB}\mathbf{67}\%})$:
          Average of 1000 67% lower bounds that are estimated by a model from Mueller and Watson (2018) between simulated $I(d)$ processes $I(d)$

- $\widehat{\gamma_{UB^{MW}_{67\%}}} = mean(\gamma_{\mathbf{UB},\mathbf{67}\%})$:
  Average of 1000 67% upper bounds that are estimated by a model from Mueller and Watson (2018) between simulated $I(d)$ processes $I(d)$
- $\widehat{\gamma_{LB^{MW}_{90\%}}} = mean(\gamma_{\mathbf{LB},\mathbf{90}\%})$:
  Average of 1000 90% lower bounds that are estimated by a model from Mueller and Watson (2018) between simulated $I(d)$ processes $I(d)$
- $\widehat{\gamma_{UB^{MW}_{90\%}}} = mean(\gamma_{\mathbf{UB},\mathbf{90}\%})$:
  Average of 1000 90% upper bounds that are estimated by a model from Mueller and Watson (2018) between simulated $I(d)$ processes $I(d)$
- $\gamma_{LB^{Em}_{67\%}} = \gamma_{\mathbf{estimate}}(166)$:
  Empirical 67% lower bound of parameter after computation of 1000 parameter values from 1000 pairs of simulated $I(d)$ processes
- $\gamma_{UB^{Em}_{67\%}} = \gamma_{\mathbf{estimate}}(835)$:
  Empirical 67% upper bound of parameter after computation of 1000 parameter values from 1000 pairs of simulated $I(d)$ processes
- $\gamma_{LB^{Em}_{90\%}} = \gamma_{\mathbf{estimate}}(51)$:
  Empirical 90% lower bound of parameter after computation of 1000 parameter values from 1000 pairs of simulated $I(d)$ processes
- $\gamma_{UB^{Em}_{90\%}} = \gamma_{\mathbf{estimate}}(950)$:
  Empirical 90% upper bound of parameter after computation of 1000 parameter values from 1000 pairs of simulated $I(d)$ processes
- End

(e) End

4. End

- End

This algorithm is executed for the $I(0)$, $I(1)$ and the (A, B, c, d) algorithms that are provided by Müller and Watson (2018).

Appendix D. Simulation results

Table 3: Simulated estimates and empirical and Mueller & Watson confidence intervals of ρ

Model to estimate ρ	I(0)			I(1)			(A,B,c,d)		
Simulated process	$\hat{\rho}$	67% CI	90% CI	$\hat{\rho}$	67% CI	90% CI	$\hat{\rho}$	67% CI	90% CI
I(-0.5)	0.01	-0.32, 0.34	-0.55, 0.56	0.01	-0.41, 0.42	-0.65, 0.66	0.00	-0.23 , 0.23	-0.39, 0.43
		(-0.25, 0.26)	(-0.41, 0.43)		(-0.24, 0.25)	(-0.39, 0.40)		(-0.25, 0.25)	(-0.42, 0.42)
I(-0.4)	-0.01	-0.33, 0.31	-0.51, 0.50	-0.01	-0.44, 0.40	-0.66, 0.65	0.03	-0.13, 0.26	-0.35, 0.43
		(-0.27, 0.25)	(-0.43, 0.41)		(-0.25, 0.23)	(-0.41, 0.39)		(-0.23, 0.28)	(-0.40, 0.45)
I(-0.3)	0.00	-0.32,0.31	-0.25, 0.26	0.01	-0.41, 0.42	-0.63, 0.63	-0.01	-0.25, 0.16	-0.42, 0.41
		(-0.25, 0.26)	(-0.49, 0.50)		(-0.24, 0.25)	(-0.40, 0.41)		(-0.26, 0.24)	(-0.43, 0.42)
I(-0.2)	-0.01	-0.32, 0.30	-0.49, 0.46	-0.01	-0.41, 0.39	-0.63, 0.63	0.03	-0.20, 0.26	-0.38, 0.46
		(-0.27, 0.25)	(-0.44, 0.42)		(-0.25, 0.23)	(-0.41, 0.39)		(-0.23, 0.28)	(0.41, 0.46)
I(-0.1)	0.00	-0.29, 0.30	-0.49, 0.49	0.01	-0.38, 0.37	-0.60, 0.61	0.00	-0.20, 0.20	-0.38, 0.38
		(-0.26, 0.26)	(-0.42, 0.43)		(-0.24, 0.25)	(-0.40, 0.41)		(-0.25, 0.26)	(-0.44, 0.45)
I(0)	0.00	-0.29, 0.27	-0.49, 0.49	-0.01	-0.40, 0.35	-0.61, 0.58	0.01	-0.21, 0.23	-0.44, 0.45
		(-0.26, 0.26)	(-0.43, 0.43)		(-0.26, 0.24)	(-0.42, 0.40)		(-0.24, 0.28)	(-0.44, 0.47)
I(0.1)	0.00	-0.30, 0.30	-0.47, 0.49	-0.01	-0.36, 0.39	-0.59, 0.59	-0.01	-0.21, 0.21	-0.41, 0.41
		(-0.26, 0.26)	(-0.43, 0.43)		(-0.25, 0.24)	(-0.41, -0.41)		(-0.28, 0.25)	(-0.48, 0.46)
I(0.2)	-0.02	-0.32, 0.29	-0.28, 0.24	-0.03	-0.38, 0.34	-0.60, 0.54	-0.02	-0.27, 0.20	-0.42, 0.41
		(-0.28, 0.24)	(-0.51, 0.48)		(-0.28, 0.22)	(-0.44, 0.39)		(-0.28, 0.25)	(-0.50, 0.47)
I(0.3)	-0.01	-0.32, 0.31	-0.52, 0.47	-0.01	-0.35, 0.34	-0.56, 0.57	-0.02	-0.27, 0.21	-0.44, 0.38
		(-0.26, 0.25)	(-0.43, 0.42)		(-0.26, 0.24)	(-0.42, 0.41)		(-0.30, 0.26)	(-0.52, 0.49)
I(0.4)	0.01	-0.34, 0.34	-0.54, 0.55	0.01	-0.35, 0.34	-0.55, 0.55	0.00	-0.25, 0.23	-0.43, 0.45
		(-0.25, 0.26)	(-0.41, 0.42)		(-0.25, 0.26)	(-0.41, 0.42)		(-0.29, 0.29)	(-0.52, 0.52)
I(0.5)	0.00	-0.35, 0.37	-0.58, 0.58	0.00	-0.34, 0.35	-0.53, 0.53	-0.01	-0.28, 0.23	-0.46, 0.43
		(-0.25, 0.25)	(-0.41, 0.41)		(-0.25, 0.26)	(-0.42, 0.42)		(-0.31, 0.29)	(-0.54, 0.53)
I(0.6)	0.01	-0.40, 0.42	-0.60, 0.60	0.00	-0.34, 0.31	-0.52, 0.51	0.00	-0.23, 0.27	-0.45, 0.44
		(-0.24, 0.25)	(-0.40, 0.41)		(-0.26, 0.25)	(-0.42, 0.42)		(-0.31, 0.31)	(-0.55, 0.55)
I(0.7)	-0.01	-0.41, 0.41	-0.62, 0.62	0.02	-0.29, 0.33	-0.48, 0.52	0.01	-0.27, 0.28	-0.46, 0.42
		(-0.25, 0.23)	(-0.41, 0.39)		(-0.24, 0.27)	(-0.41, 0.44)		(-0.32, 0.33)	(-0.56, 0.56)
I(0.8)	-0.01	-0.50, 0.47	-0.73, 0.69	-0.01	-0.31, 0.29	-0.50, 0.51	-0.01	-0.28, 0.27	-0.46, 0.46
		(-0.24, 0.47)	(-0.40, 0.37)		(-0.26, 0.25)	(-0.43, 0.42)		(-0.35, 0.32)	(-0.57, 0.55)
I(0.9)	-0.02	-0.56, 0.51	-0.77, 0.76	0.00	-0.29, 0.30	-0.48, 0.44	0.00	-0.30, 0.30	-0.48, 0.47
		(-0.24, 0.21)	(-0.38, 0.36)		(-0.26, 0.26)	(-0.42, 0.43)		(-0.34, 0.34)	(-0.55, 0.55)
I(1)	-0.03	-0.63, 0.59	-0.82, 0.80	-0.02	-0.31, 0.27	-0.51, 0.48	-0.01	-0.32, 0.32	-0.48, 0.47
		(0.24, 0.18)	(-0.37, 0.32)		(-0.28, 0.24)	(-0.44, 0.41)		(-0.34, 0.33)	(-0.54, 0.53)
I(1.1)	0.03	-0.66, 0.70	-0.86, 0.87	0.01	-0.29, 0.32	0.45, 0.50	0.00	-0.32, 0.32	-0.49, 0.49
		(-0.17, 0.22)	(-0.30, 0.35)		(-0.25, 0.27)	(-0.42, 0.44)		(-0.33, 0.35)	(-0.51, 0.52)
I(1.2)	-0.03	-0.74, 0.71	-0.90, 0.88	-0.03	-0.33, 0.29	-0.53, 0.50	0.01	-0.33, 0.34	-0.48, 0.49
		(-0.21, 0.16)	(-0.33, 0.28)		(-0.28, 0.23)	(-0.44, 0.40)		(-0.33, 0.35)	(-0.50, 0.52)
I(1.3)	-0.02	-0.77, 0.77	-0.91, 0.92	0.00	-0.34, 0.34	-0.51, 0.55	-0.01	-0.45, 0.39	-0.64, 0.64
		(-0.19, 0.16)	(-0.31, 0.28)		(-0.25, 0.26)	(-0.42, 0.42)		(-0.35, 0.33)	(-0.50, 0.48)
I(1.4)	-0.01	-0.82, 0.82	-0.94, 0.94	0.01	-0.40, 0.41	-0.60, 0.59	0.00	-0.47, 0.47	-0.64, 0.64
		(-0.16, 0.15)	(-0.27, 0.26)		(-0.24, 0.25)	(-0.40, 0.41)		(-0.34, 0.34)	(-0.47, 0.47)
I(1.5)	0.02	-0.86, 0.87	-0.95, 0.96	0.00	-0.46, 0.47	-0.65, 0.67	0.01	-0.47, 0.47	-0.66, 0.66
		(NaN, NaN)	(NaN, NaN)		(-0.23, 0.24)	(-0.39, 0.40)		(-0.32, 0.34)	(-0.44, 0.46)

Obtained estimates and confidence intervals for ρ from different simulated (fractionally) integrated processes and different models. The model in the upper row indicates the model that is used to estimate the values, the left column indicates the integration order of the simulated processes. The empirical confidence intervals are shown next to the estimates, the average Müller & Watson confidence intervals are given underneath corresponding empirical CI's between brackets. Simulated data series contained 168 observations and periods longer than 14 observations correspond to $q = 12$.

Table 4: Simulated estimates and confidence intervals of β

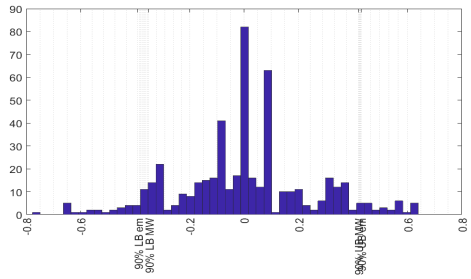
Model to estimate β	I(0)			I(1)			(A,B,c,d)		
Simulated process	$\hat{\beta}$	67% CI	90% CI	$\hat{\beta}$	67% CI	90% CI	$\hat{\beta}$	67% CI	90% CI
I(-0.5)	0.00	-0.32, 0.33	-0.59, 0.61	0.00	-0.43, 0.43	-0.74, 0.81	-0.01	-0.31, 0.28	-0.59, 0.51
		(-0.30, 0.31)	(-0.59, 0.61)		(-0.30, 0.31)	(-0.54, 0.55)		(-0.33, 0.31)	(-0.58, 0.57)
I(-0.4)	-0.01	-0.32, 0.30	-0.56, 0.57	-0.01	-0.43, 0.41	-0.78, 0.77	0.03	-0.25, 0.30	-0.45, 0.59
		(-0.31, 0.30)	(-0.54, 0.53)		(-0.31, 0.29)	(-0.54, 0.52)		(-0.30, 0.34)	(-0.55, 0.60)
I(-0.3)	0.01	-0.30, 0.33	-0.52, 0.55	0.01	-0.39, 0.43	-0.70, 0.79	-0.01	-0.31, 0.27	-0.55, 0.50
		(-0.30, 0.31)	(-0.54, 0.55)		(-0.29, 0.32)	(-0.53, 0.56)		(-0.34, 0.31)	(-0.60, 0.57)
I(-0.2)	-0.02	-0.31, 0.29	-0.59, 0.50	-0.03	-0.42, 0.37	-0.73, 0.67	0.02	-0.27, 0.33	-0.52, 0.51
		(-0.33, 0.39)	(-0.56, 0.52)		(-0.33, 0.28)	(-0.56, 0.51)		(-0.31, 0.34)	(-0.58, 0.62)
I(-0.1)	0.00	-0.28, 0.29	-0.52, 0.51	-0.01	-0.38, 0.39	-0.67, 0.70	0.00	-0.26, 0.27	-0.51, 0.48
		(-0.30, 0.31)	(-0.53, 0.54)		(-0.30, 0.31)	(-0.53, 0.55)		(-0.33, 0.33)	(-0.60, 0.61)
I(0)	0.00	-0.29, 0.26	-0.55, 0.50	-0.01	-0.40, 0.36	-0.68, 0.67	0.02	-0.30, 0.32	-0.52, 0.59
		(-0.31, 0.30)	(-0.54, 0.53)		(-0.32, 0.29)	(-0.55, 0.52)		(-0.32, 0.37)	(-0.63, 0.69)
I(0.1)	0.00	-0.29, 0.29	-0.53, 0.50	-0.01	-0.38. 0.36	-0.69, 0.69	-0.02	-0.32, 0.28	-0.57, 0.49
		(-0.31, 0.31)	(-0.54, 0.54)		(-0.31, 0.29)	(-0.55, 0.52)		(-0.37, 0.32)	(-0.69, 0.64)
I(0.2)	-0.02	-0.33, 0.28	-0.53, 0.53	-0.03	-0.39, 0.34	-0.66, 0.61	-0.02	-0.32, 0.27	-0.53, 0.47
		(-0.32, 0.29)	(-0.56, 0.53)		(-0.34, 0.28)	(-0.57, 0.52)		(-0.37, 0.32)	(-0.72. 0.68)
I(0.3)	-0.01	-0.32, 0.31	-0.58, 0.53	0.00	-0.36, 0.35	-0.64, 0.65	-0.02	-0.31, 0.26	-0.53, 0.48
		(-0.32, 0.30)	(-0.56, 0.54)		(-0.31, 0.31)	(-0.55, 0.54)		(-0.38, 0.33)	(-0.76, 0.71)
I(0.4)	0.01	-0.33, 0.36	-0.58, 0.59	0.01	-0.33, 0.34	-0.59, 0.61	-0.01	-0.30, 0.32	-0.55, 0.55
		(-0.29, 0.31)	(-0.52, 0.54)		(-0.30, 0.31)	(-0.53, 0.54)		(-0.39, 0.37)	(-0.80, 0.79)
I(0.5)	-0.01	-0.37, 0.37	-0.69, 0.61	0.00	-0.32, 0.34	-0.60, 0.58	-0.02	-0.36, 0.30	-0.54, 0.52
		(-0.32, 0.30)	(-0.56, 0.54)		(-0.31, 0.32)	(-0.60, 0.58)		(-0.41, 0.38)	(-0.84, 0.82)
I(0.6)	0.01	-0.41, 0.42	-0.74, 0.75	0.00	-0.33, 0.,32	-0.56, 0.61	0.01	-0.29, 0.32	-0.60, 0.60
		(-0.30, 0.33)	(-0.54, 0.57)		(-0.31, 0.31)	(-0.54, 0.54)		(-0.39, 0.41)	(-0.83, 0.85)
I(0.7)	-0.01	-0.42, 0.43	-0.79, 0.77	0.02	-0.29, 0.32	-0.50, 0.58	0.00	-0.31, 0.35	-0.60, 0.53
		(-0.32, 0.31)	(-0.56, 0.55)		(-0.29, 0.33)	(-0.53, 0.57)		(-0.43, 0.44)	(-0.89, 0.90)
I(0.8)	-0.01	-0.48, 0.47	-0.90, 0.91	-0.01	-0.31, 0.29	-0.56, 0.52	-0.01	-0.38, 0.31	-0.63, 0.60
		(-0.32, 0.29)	(-0.55, 0.53)		(-0.31, 0.29)	(-0.54, 0.52)		(-0.48, 0.45)	(-0.92, 0.89)
I(0.9)	-0.02	-0.52, 0.51	-1.00, 0.99	0.01	-0.29, 0.31	-0.51, 0.51	0.00	-0.37, 0.35	-0.59, 0.67
		(-0.32, 0.28)	(-0.54, 0.51)		(-0.30, 0.32)	(-0.53, 0.55)		(-0.46, 0.49)	(0.89, 0.92)
I(1)	-0.06	-0.67, 0.55	-1.21, 1.07	-0.02	-0.32, 0.28	-0.53, 0.48	-0.02	-0.40, 0.36	-0.75, 0.63
		(-0.36, 0.25)	(-0.59, 0.48)		(-0.33, 0.29)	(-0.56, 0.52)		(-0.50, 0.47)	(-0.91, 0.87)
I(1.1)	0.02	-0.71, 0.73	-1.26, 1.30	0.01	-0.28, 0.32	-0.54, 0.51	-0.01	-0.42, 0.39	-0.73, 0.67
		(-0.28, 0.32)	(-0.51, 0.55)		(-0.28, 0.32)	(-0.53, 0.55)		(-0.50, 0.49)	(-0.88, -/87)
I(1.2)	-0.03	-0.79, 0.75	-1.51, 1.43	-0.02	-0.33, 0.30	-0.60, 0.54	0.01	-0.39, 0.40	-0.66, 0.75
		(-0.34, 0.29)	(-0.58, 0.52)		(-0.34, 0.29)	(-0.57, 0.52)		(-0.47, 0.50)	(-0.83, 0.85)
I(1.3)	-0.05	-0.93, 0.83	-1.79, 1.81	0.00	-0.34, 0.34	-0.16, 0.63	-0.03	-0.46, 0.41	-0.85, 0.78
		(-0.38, 0.28)	(-0.63, 0.53)		(-0.32, 0.31)	(-0.56, 0.55)		(-0.50, 0.45)	(-0.85, 0.79)
I(1.4)	-0.01	-0.87, 0.88	-1.91, 1.87	0.01	-0.40, 0.41	-0.70, 0.78	-0.02	-0.53, 0.54	-1.04, 0.82
		(-0.31, 0.29)	(-0.54, 0.52)		(-0.30, 0.31)	(-0.53, 0.55)		(-0.51, 0.45)	(-0.85, 0.79)
I(1.5)	0.07	-0.88, 1.04	-1.78, 1.98	0.01	-0.42, 0.49	-0.78, 0.83	0.01	-0.53, 0.56	-0.98, 1.06
		(-0.25, 0.37)	(-0.48, 0.60)		(-0.29, 0.31)	(-0.52, 0.55)		(-0.43, 0.46)	(-0.76, 0.79)

Obtained estimates and confidence intervals for β from different simulated (fractionally) integrated processes and different models. The model in the upper row indicates the model that is used to estimate the values, the left column indicates the integration order of the simulated processes. The empirical confidence intervals are shown next to the estimates, the average Müller & Watson confidence intervals are given underneath corresponding empirical CI's between brackets. Simulated data series contained 168 observations and periods longer than 14 observations correspond to $q = 12$.

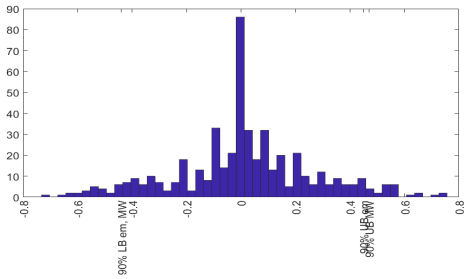
Table 5: Simulated estimates and confidence intervals of $\sigma_{y|x}$

Model to estimate $\sigma_{y\|x}$	I(0)			I(1)			(A,B,c,d)		
Simulated process	$\hat{\sigma}_{y\|x}$	67% CI	90% CI	$\hat{\sigma}_{y\|x}$	67% CI	90% CI	$\hat{\sigma}_{y\|x}$	67% CI	90% CI
I(-0.5)	0.00	0.00, 0.00 (0.00, 0.00)	0.00, 0.00 (0.00, 0.00)	0.00	0.00, 0.00 (0.00, 0.00)	0.00, 0.00 (0.00, 0.00)	0.12	0.09, 0.15 (0.10, 0.15)	0.07, 0.17 (0.09, 0.19)
I(-0.4)	0.00	0.00, 0.00 (0.00, 0.00)	0.00, 0.00 (0.00, 0.00)	0.00	0.00, 0.00 (0.00, 0.00	0.00, 0.01 (0.00, 0.01)	0.15	0.12, 0.18 (0.12, 0.19)	0.10, 0.21 (0.11, 0.23)
I(-0.3)	0.00	0.00, 0.00 (0.00, 0.00)	0.00, 0.00 (0.00. 0.01)	0.00	0.00, 0.01 (0.00, 0.01)	0.00, 0.01 (0.00, 0.01)	0.18	0.14, 0.23 (0.15, 0.24)	0.12, 0.25 (0.13, 0.29)
I(-0.2)	0.00	0.00. 0.01 (0.00. 0.01)	0.00. 0.01 (0.00. 0.01)	0.01	0.00. 0.01 (0.00. 0.01)	0.00. 0.01 (0.00. 0.01)	0.22	0.18, 0.27 (0.18, 0.29)	0.15, 0.30 (0.16, 0.36)
I(-0.1)	0.01	0.00. 0.01 (0.00. 0.01)	0.00. 0.01 (0.00. 0.01)	0.01	0.00. 0.01 (0.01, 0.01)	0.00, 0.02 (0.00, 0.02)	0.28	0.22, 0.34 (0.23, 0.37)	0.19, 0.38 (0.20, 0.46)
I(0)	0.01	0.01, 0.01 (0.01, 0.02)	0.00, 0.02 (0.00, 0.02)	0.01	0.01, 0.02 (0.01, 0.02)	0.00, 0.02 (0.01, 0.03)	0.36	0.28, 0.43 (0.29, 0.46)	0.23, 0.48 (0.25, 0.59)
I(0.1)	0.01	0.01, 0.02 (0.01, 0.02)	0.01, 0.03 (0.01, 0.03)	0.02	0.01, 0.02 (0.01, 0.03)	0.01, 0.03 (0.01, 0.04)	0.45	0.35, 0.54 (0.36, 0.58)	0.30, 0.60 (0.31, 0.75)
I(0.2)	0.02	0.01, 0.03 (0.02, 0.04)	0.01, 0.04 (0.01, 0.05)	0.02	0.01, 0.03 (0.02, 0.04)	0.01, 0.05 (0.01, 0.05)	0.56	0.45, 0.67 (0.45, 0.74)	0.37, 0.77 (0.39, 0.97)
I(0.3)	0.04	0.02, 0.05 (0.03, 0.06)	0.02, 0.07 (0.02, 0.09)	0.03	0.02, 0.05 (0.02, 0.06)	0.01, 0.07 (0.02, 0.08)	0.70	0.55, 0.85 (0.56, 0.94)	0.46, 0.97 (0.49, 1.26)
I(0.4)	0.06	0.03, 0.09 (0.04, 0.10)	0.02, 0.12 (0.03, 0.15)	0.05	0.03, 0.07 (0.03, 0.08)	0.02, 0.10 (0.03, 0.12)	0.92	0.71, 1.12 (0.74, 1.25)	0.59, 1.27 (0.64, 1.72)
I(0.5)	0.11	0.06, 0.15 (0.08, 0.18)	0.04, 0.23 (0.06, 0.26)	0.07	0.04, 0.10 (0.05, 0.12)	0.03, 0.13 (0.04, 0.17)	1.17	0.92, 1.44 (0.93, 1.63)	0.74, 1.65 (0.81, 2.24)
I(0.6)	0.19	0.09, 0.27 (0.13, 0.31)	0.06, 0.40 (0.10, 0.45)	0.11	0.06, 0.15 (0.08, 0.018)	0.04, 0.20 (0.06, 0.25)	1.49	1.13, 1.87 (1.18, 2.11)	0.94, 2.10 (1.03, 2.91)
I(0.7)	0.35	0.16, 0.52 (0.25, 0.59)	0.10, 0.84 (0.20, 0.84)	0.16	0.09, 0.23 (0.12, 0.27)	0.06, 0.31 (0.09, 0.39)	2.02	1.55, 2.50 (1.59, 2.97)	1.18, 2.87 (1.39, 4.06)
I(0.8)	0.65	0.26, 0.99 (0.46, 1.09)	0.17, 1.79 (0.36, 1.56)	0.24	0.14, 0.35 (0.17, 0.41)	0.09, 0.46 (0.14, 0.58)	2.74	2.03, 3.42 (2.14, 4.14)	1.65, 4.08 (1.87, 5.53)
I(0.9)	1.24	0.45, 1.95 (0.88, 2.08)	0.29, 3.06 (0.69, 2.98)	0.39	0.23, 0.56 (0.28, 0.66)	0.16, 0.70 (0.22, 0.94)	3.68	2.70, 4.57 (2.84, 5.76)	2.20, 5.35 (2.49, 7.43)
I(1)	2.47	0.80, 3.71 (1.76, 4.15)	0.50, 7.34 (1.38, 5.93)	0.62	0.37, 0.87 (0.44, 1.04)	0.25, 1.11 (0.35, 1.49)	4.79	3.58, 6.12 (3.67, 7.60)	2.82, 6.86 (3.23, 9.58)
I(1.1)	4.96	1.44, 7.73 (3.54, 8.34)	0.87, 15.04 (2.77, 11.93)	1.04	0.62, 1.48 (0.75, 1.76)	0.43, 1.94 (0.58, 2.51)	6.59	4.69, 8.44 (4.91, 10.69)	3.82, 9.85 (4.35, 15.66)
I(1.2)	10.70	2.52, 18.25 (7.64, 17.99)	1.32, 36.00 (5.98, 25.72)	1.81	1.03, 2.56 (1.29, 3.04)	0.68, 3.32 (1.01, 4.34)	8.93	6.34, 11.36 (6.57, 14.72)	4.88, 13.57 (5.86, 19.91)
I(1.3)	23.83	4.84, 38.02 (17.03, 40.09)	2.42, 80.41 (13.33, 57.31)	3.18	2.69, 4.63 (2.27, 5.35)	1.19, 6.38 (1.78, 7.65)	11.99	8.08, 15.99 (8.69, 19.95)	6.49, 19.42 (7.80, 31.42)
I(1.4)	48.17	8.30, 77.03 (34.41, 81.02)	3.98, 181.53 (26.93, 115.83)	5.75	2.82, 8.51 (4.11, 9.67)	1.95, 13.03 (3.22, 13.83)	16.99	10.95, 23.51 (11.97, 28.67)	8.34, 28.17 (10.91, 40.75)
I(1.5)	103.91	14.16, 157.11 (74.23, 274.76)	6.50, 401.98 (58.09, 249.84)	10.64	4.92, 16.21 (7.60, 17.89)	3.18, 25.15 (.95, 25.57)	23.46	15.23, 31.60 (16.03, 42.09)	11.32, 40.96 (14.77, 80.08)

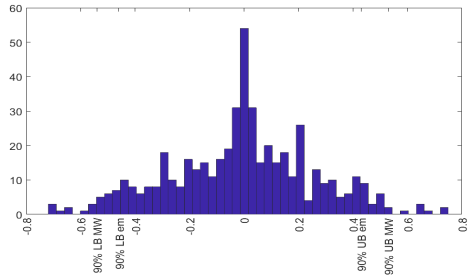
Obtained estimates and confidence intervals for $\sigma_{y|x}$ from different simulated (fractionally) integrated processes and different models. The model in the upper row indicates the model that is used to estimate the values, the left column indicates the integration order of the simulated processes. The empirical confidence intervals are shown next to the estimates, the average Müller & Watson confidence intervals are given underneath corresponding empirical CI's between brackets. Simulated data series contained 168 observations and periods longer than 14 observations correspond to $q = 12$.



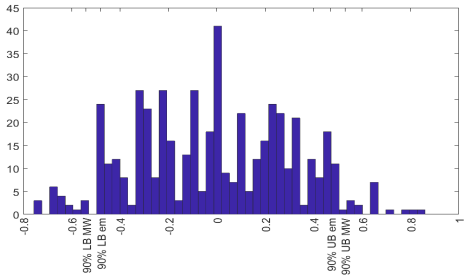
(a) Histograms and bounds of 90% confidence intervals of ρ for simulated processes of integration order $d = -0.5$.



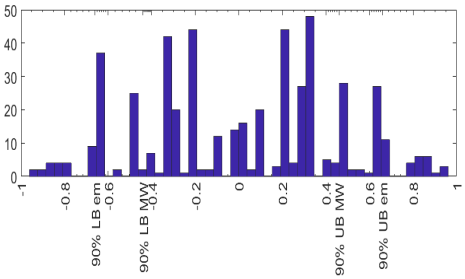
(b) Histograms and bounds of 90% confidence intervals of ρ for simulated processes of integration order $d = 0$.



(c) Histograms and bounds of 90% confidence intervals of ρ for simulated processes of integration order $d = 0.5$.

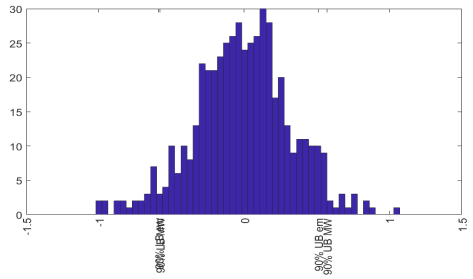


(d) Histograms and bounds of 90% confidence intervals of ρ for simulated processes of integration order $d = 1$.

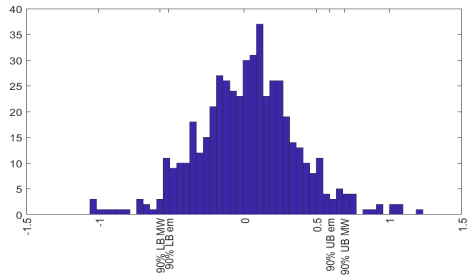


(e) Histograms and bounds of 90% confidence intervals of ρ for simulated processes of integration order $d = 1.5$.

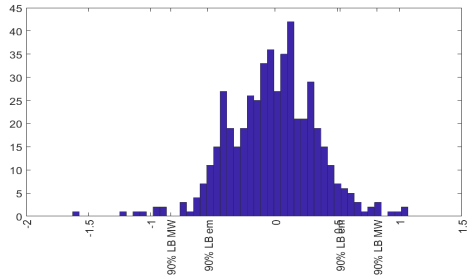
Figure 3: Histogram of ρ for simulated processes of different integration orders d. 90% lower (LB) and upper (UB) bounds of empirical confidence intervals (em) and average confidence intervals that are estimated by use of the (A, B, c, d) model from Müller and Watson (2018) (MW), are indicated at the x-axis.



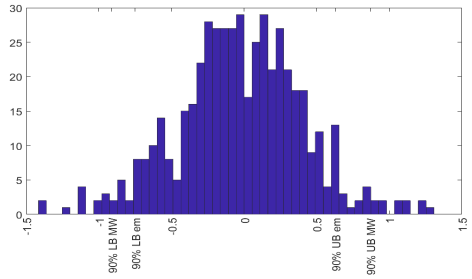
(a) Histograms and bounds of 90% confidence intervals of β for simulated processes of integration order $d = -0.5$.



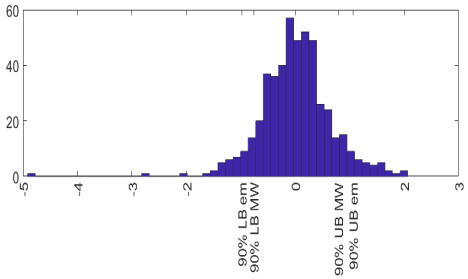
(b) Histograms and bounds of 90% confidence intervals of β for simulated processes of integration order $d = 0$.



(c) Histograms and bounds of 90% confidence intervals of β for simulated processes of integration order $d = 0.5$.

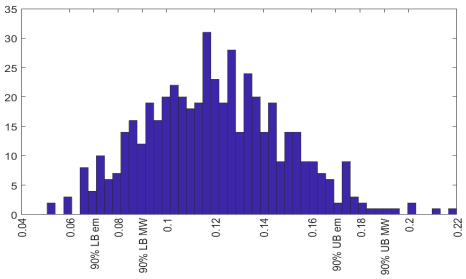


(d) Histograms and bounds of 90% confidence intervals of β for simulated processes of integration order $d = 1$.

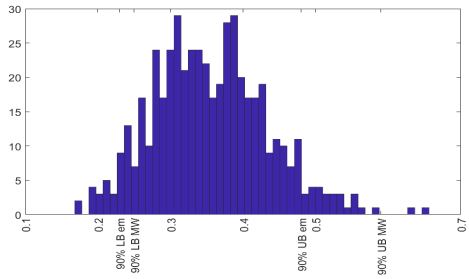


(e) Histograms and bounds of 90% confidence intervals of β for simulated processes of integration order $d = 1.5$.

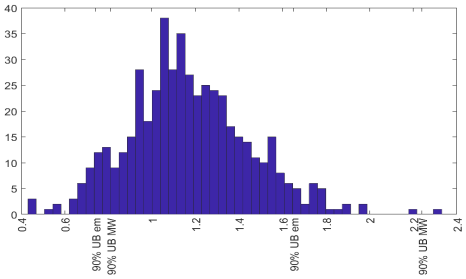
Figure 4: Histogram of β for simulated processes of different integration orders d. 90% lower (LB) and upper (UB) bounds of empirical confidence intervals (em) and average confidence intervals that are estimated by use of the (A, B, c, d) model from Müller and Watson (2018) (MW), are indicated at the x-axis.



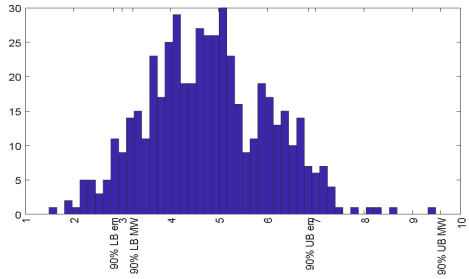
(a) Histograms and bounds of 90% confidence intervals of $\sigma_{y|x}$ for simulated processes of integration order $d = -0.5$.



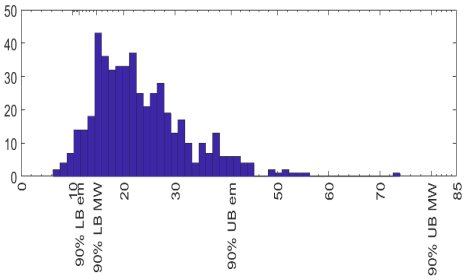
(b) Histograms and bounds of 90% confidence intervals of $\sigma_{y|x}$ for simulated processes of integration order $d = 0$.



(c) Histograms and bounds of 90% confidence intervals of $\sigma_{y|x}$ for simulated processes of integration order $d = 0.5$.



(d) Histograms and bounds of 90% confidence intervals of $\sigma_{y|x}$ for simulated processes of integration order $d = 1$.



(e) Histograms and bounds of 90% confidence intervals of $\sigma_{y|x}$ for simulated processes of integration order $d = 1.5$.

Figure 5: Histogram of $\sigma_{y|x}$ for simulated processes of different integration orders d. 90% lower (LB) and upper (UB) bounds of empirical confidence intervals (em) and average confidence intervals that are estimated by use of the (A, B, c, d) model from Müller and Watson (2018) (MW), are indicated at the x-axis.

Appendix E. Sensitivity analysis of replication results

Data plots

GDP and consumption

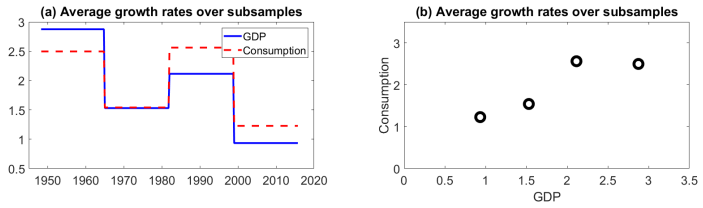


Figure 6: Panel (a) plots sample averages of consumption and GDP over 4 periods. Panel (b) is a scatter plot of the averages in (a).

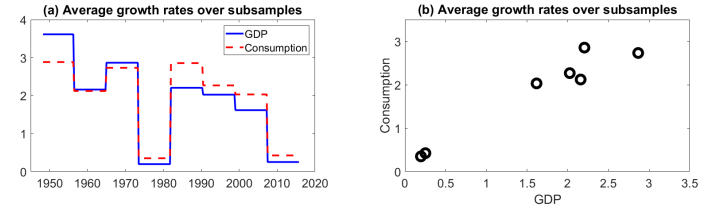


Figure 7: Panel (a) plots sample averages of consumption and GDP over 8 periods. Panel (b) is a scatter plot of the averages in (a).

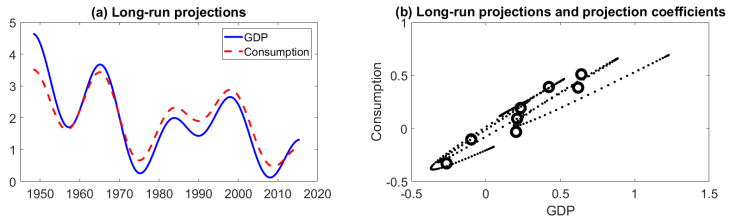


Figure 8: Panel (a) plots the projections of consumption and GDP onto the low-frequency consine terms with added sample means, $q = 8$. Panel (b) is a scatterplot of the data projections (small dots) and the projection coefficients (X_{jT}, Y_{jT}) (large dots).

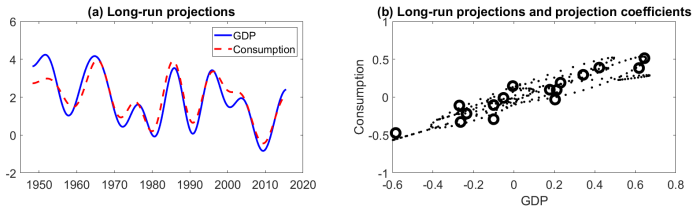


Figure 9: Panel (a) plots the projections of consumption and GDP onto the low-frequency consine terms with added sample means, $q = 16$. Panel (b) is a scatterplot of the data projections (small dots) and the projection coefficients (X_{jT}, Y_{jT}) (large dots).

Short-term and long-term interest rates

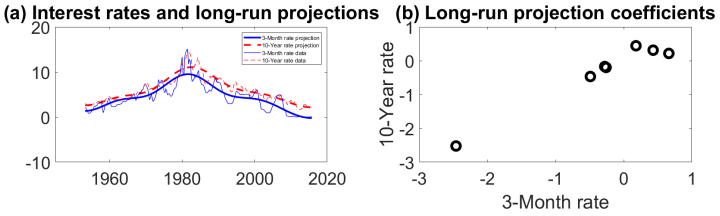


Figure 10: Panel (a) plots the projections of the short- and long-term interest rates onto the low-frequency cosine terms with added sample means, $q = 7$. Panel (b) is a scatterplot of the projection coefficients (X_{jT}, Y_{jT}) from (a).

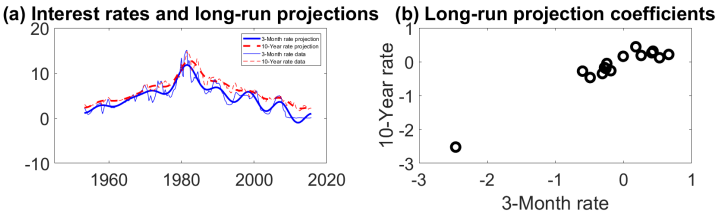


Figure 11: Panel (a) plots the projections of the short- and long-term interest rates onto the low-frequency cosine terms with added sample means, $q = 15$. Panel (b) is a scatterplot of the projection coefficients (X_{jT}, Y_{jT}) from (a).

Long-run covariability estimates

Table 6: Replication of table 1: periods longer than 17 years

		ρ	β	$\sigma_{y\|x}$
		(a) GDP and consumption		
I(0)	Estimate	0.95	0.75	0.28
	67% CI	0.89, 0.97	0.65, 0,84	0.23, 0.39
	90% CI	0.82, 0.98	0.58, 0.91	0.19, 0.50
I(1)	Estimate	0.99	0.81	0.18
	67% CI	0.97, 0.99	0.76, 0.86	0.15, 0.26
	90% CI	0.95, 0.99	0.72, 0.90	0.13, 0.32
		(b) Short- and long-term interest rates		
I(0)	Estimate	0.98	0.97	0.59
	67% CI	0.94, 0.99	0.88, 1.06	0.48, 0.86
	90% CI	0.90, 0.99	0.81, 1.13	0.41, 1.13
I(1)	Estimate	0.97	0.95	0.40
	67% CI	0.93, 0.98	0.85, 1.04	0.32, 0.58
	90% CI	0.88, 0.99	0.77, 1.13	0.27, 0.76

Periods longer than 17 years correspond to q = 8 for panel (a) and q = 7 for panel (b). The estimates are the maximum likelihood estimates found by the large-sample distribution of the cosine transforms for the $I(0)$ and $I(1)$ models. 67% and 90% confidence intervals are displayed below the estimates.

Table 7: Replication of table 1: periods longer than 8 years

		ρ	β	$\sigma_{y\|x}$
		(a) GDP and consumption		
I(0)	Estimate	0.94	0.76	0.40
	67% CI	0.89, 0.96	0.69, 0,84	0.35, 0.50
	90% CI	0.85, 0.97	0.63, 0.89	0.31, 0.58
I(1)	Estimate	0.93	0.77	0.43
	67% CI	0.89, 0.96	0.70, 0.85	0.37, 0.54
	90% CI	0.84, 0.97	0.64, 0.91	0.34, 0.62
		(b) Short- and long-term interest rates		
I(0)	Estimate	0.96	0.91	0.81
	67% CI	0.93, 0.97	0.84, 0.99	0.69, 1.01
	90% CI	0.90, 0.98	0.78, 1.04	0.62, 1.18
I(1)	Estimate	0.84	0.55	0.68
	67% CI	0.74, 0.90	0.46, 0.65	0.58, 0.85
	90% CI	0.64, 0.92	0.39, 0.72	0.52, 0.99

Periods longer than 8 years correspond to q = 16 for panel (a) and q = 15 for panel (b). The estimates are the maximum likelihood estimates found by the large-sample distribution of the cosine transforms for the $I(0)$ and $I(1)$ models. 67% and 90% confidence intervals are displayed below the estimates.

Table 8: Replication of table 2: periods longer than 17 years

		ρ	β	$\sigma_{y\|x}$
		a) GDP and consumption		
(A, B, c, d)	Estimate	0.95	0.76	0.30
	67% CI	0.85, 0.97	0.60, 0.85	0.23, 0.47
	90% CI	0.75, 0.98	0.48, 1.01	0.20, 0.65
	67% Bayes CS	0.89, 0.97	0.67, 0.85	0.23, 0.40
	90% Bayes CS	0.83, 0.98	0.59, 0.92	0.20, 0.51
		b) Short- and long-term interest rates		
(A, B, c, d)	Estimate	0.96	0.96	0.60
	67% CI	0.92, 0.98	0.84, 1.07	0.45, 1.00
	90% CI	0.85, 0.99	0.74, 1.18	0.38, 1.36
	67% Bayes CS	0.92, 0.98	0.87, 1.05	0.45, 0.84
	90% Bayes CS	0.87, 0.99	0.79, 1.13	0.38, 1.11

Periods longer than 17 years correspond to $q = 8$ for panel (a) and $q = 7$ for panel (b). The estimates are the posterior median based on the $I(d)$ model. The 67% and 90% confidence intervals and Bayes credible sets based on the posterior are displayed below the estimates.

Appendix F. Extension data plots

Plots on monthly growth rates

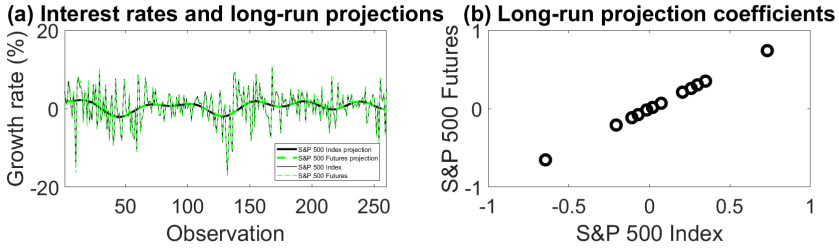


Figure 12: Monthly growth rates (in %) of S&P 500 Index and S&P 500 Futures from November 1997 to June 2019

Plots on weekly growth rates

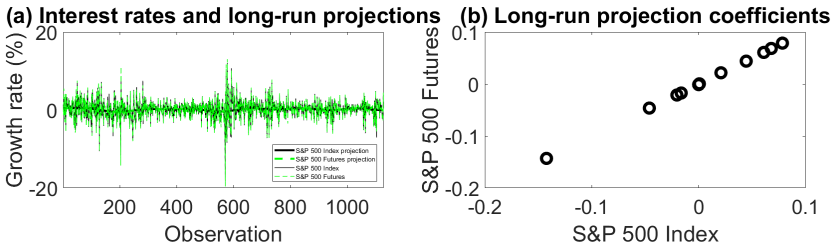
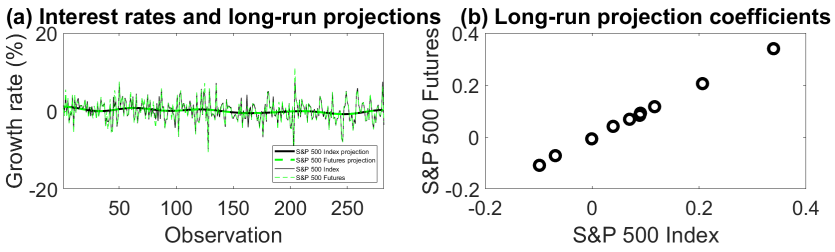
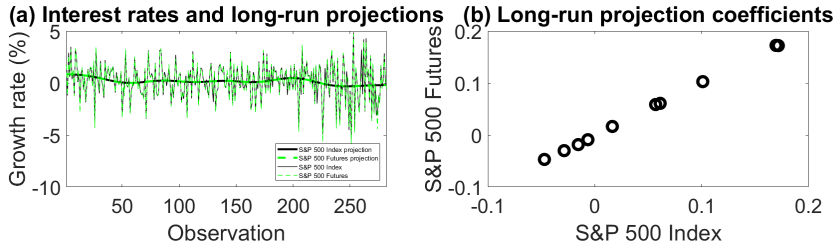


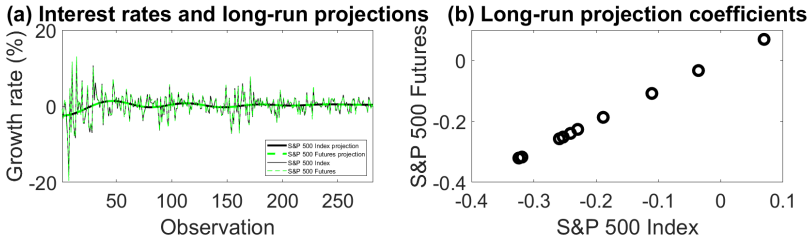
Figure 13: Weekly growth rates (in %) of S&P 500 Index and S&P 500 Futures from November 1997 to June 2019



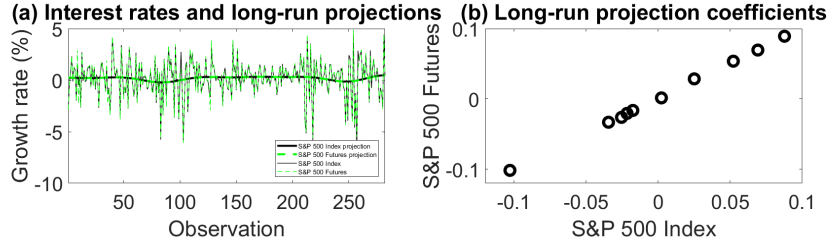
(a) Weekly market S&P 500 Index and Futures growth rates (%) from 02/11/1997 to 23/03/2003



(b) Weekly market S&P 500 Index and Futures growth rates (%) from 24/03/2003 to 17/08/2008



(c) Weekly market S&P 500 Index and Futures growth rates (%) from 24/08/2008 to 12/01/2014



(d) Weekly S&P 500 Index and Futures growth rates (%) from 19/01/2014 to 09/06/2019

Figure 14: Weekly growth rates (in %) of S&P 500 Index and S&P 500 Futures observed during four sub-samples

Plots on daily growth rates

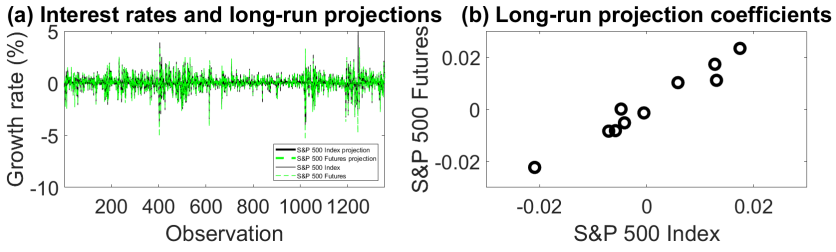
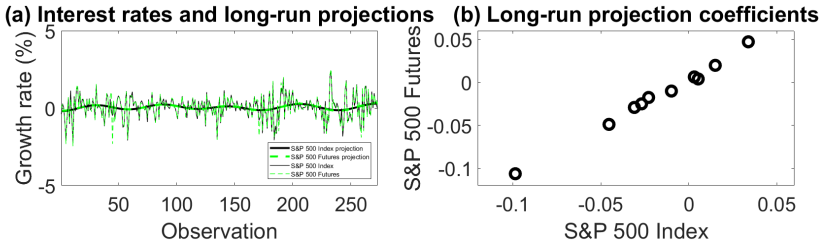
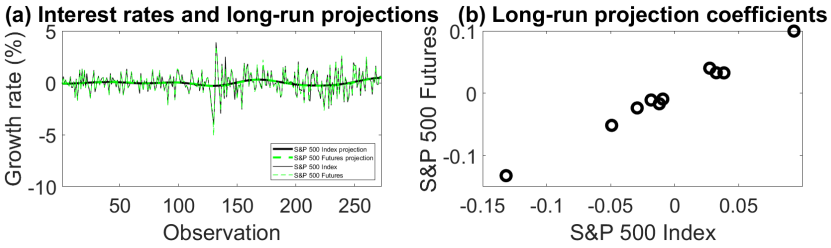


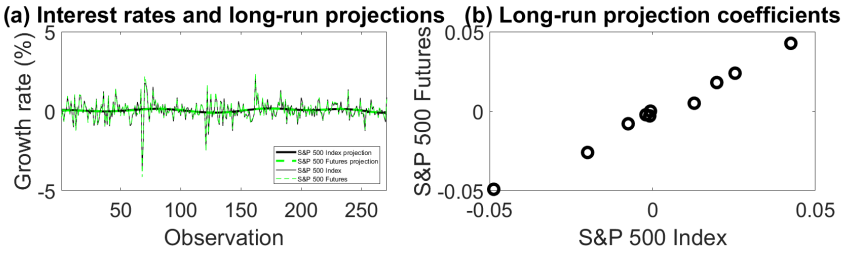
Figure 15: Daily growth rates (in %) of S&P 500 Index and S&P 500 Futures from 21/01/2104 to 09/06/2019



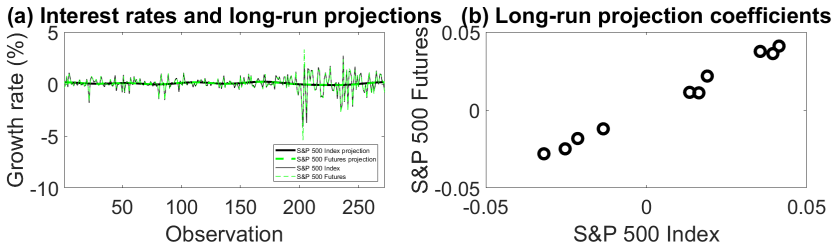
(a) Daily market S&P 500 Index and Futures growth rates (%) from 21/01/2014 to 18/02/2015



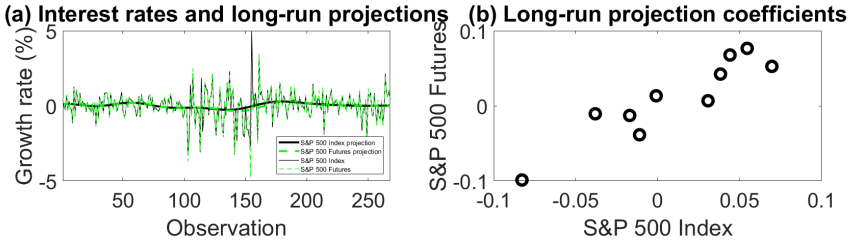
(b) Daily market S&P 500 Index and Futures growth rates (%) from 19/02/2015 to 18/03/2016



(c) Daily market S&P 500 Index and Futures growth rates (%) from 21/03/2016 to 17/04/2017



(d) Daily S&P 500 Index and Futures growth rates (%) from 18/04/2017 to 15/05/2018



(e) Daily S&P 500 Index and Futures growth rates (%) from 16/05/2018 to 10/06/2019

Figure 16: Daily growth rates (in %) of S&P 500 Index and S&P 500 Futures observed during five sub-samples

Plots on hourly growth rates

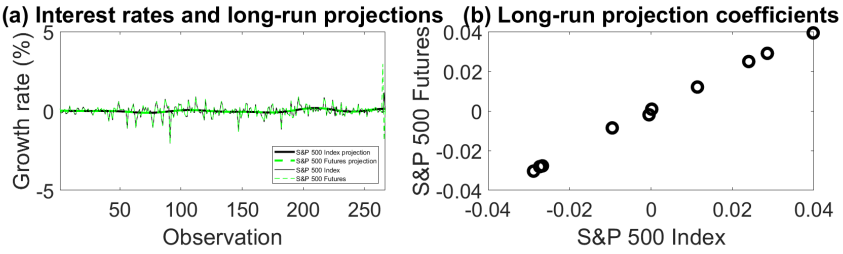
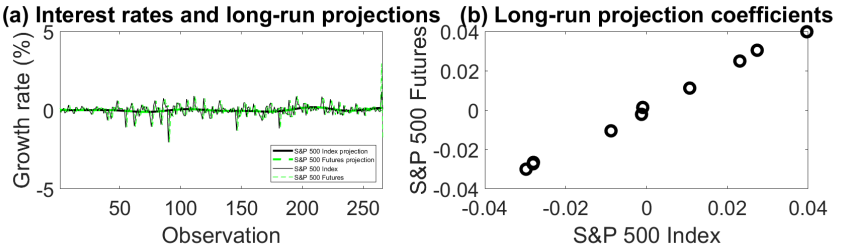
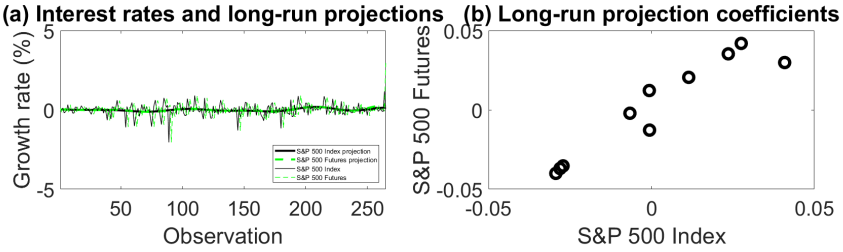


Figure 17: Hourly growth rates (in %) of S&P 500 Index and S&P 500 Futures from 24/04/2019 4pm to 18/06/2019 4pm (UTC +2:00)



(a) Hourly growth rates (in %) of S&P 500 Index and one hour ahead S&P 500 Futures



(b) Hourly growth rates (in %) of S&P 500 Index and two hours ahead S&P 500 Futures

Figure 18: Hourly growth rates (in %) of S&P 500 Index and n hours ahead S&P 500 Futures from 24/04/2019 4pm to 18/06/2019 4pm (UTC +2:00), $n = 1, 2$

Plots on growth rates per five minutes

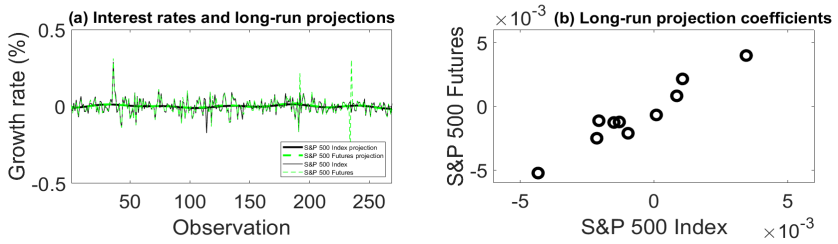
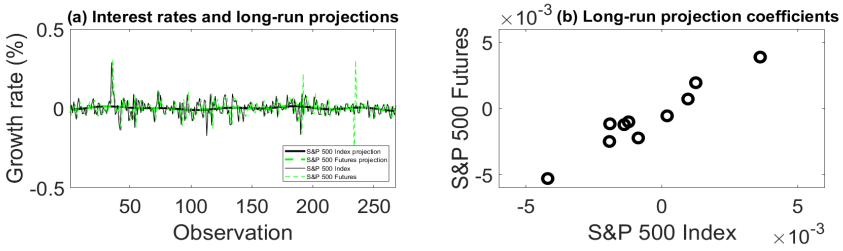
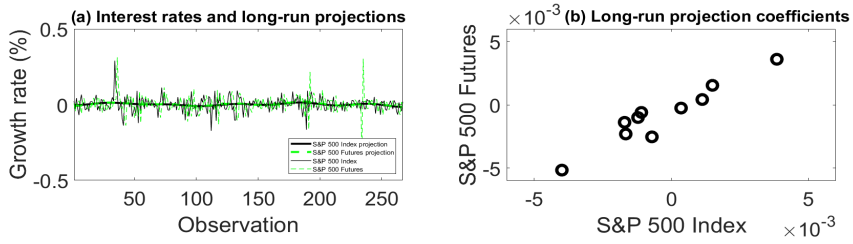


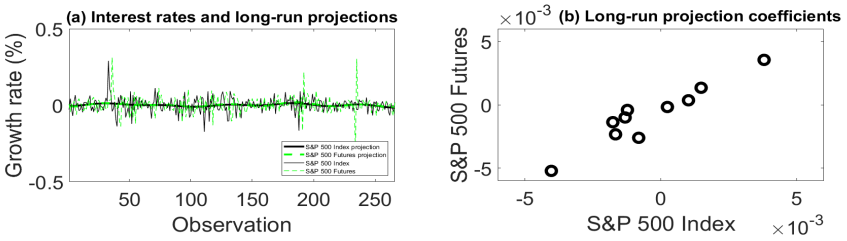
Figure 19: Growth rates (in %) of S&P 500 Futures and S&P 500 Index per 5 minutes from 12/06/2019 07:05 pm to 19/06/2019 9:55 pm (UTC +2:00)



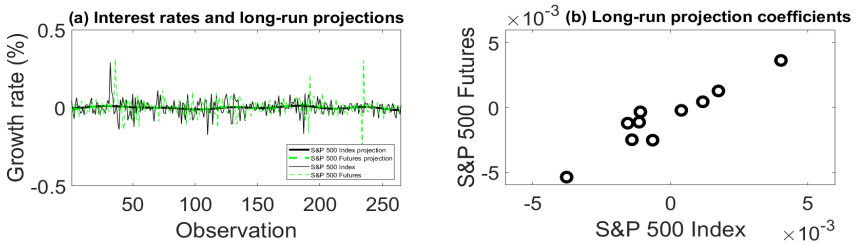
(a) Growth rates (in %) of S&P 500 Index and five minutes ahead S&P 500 Futures



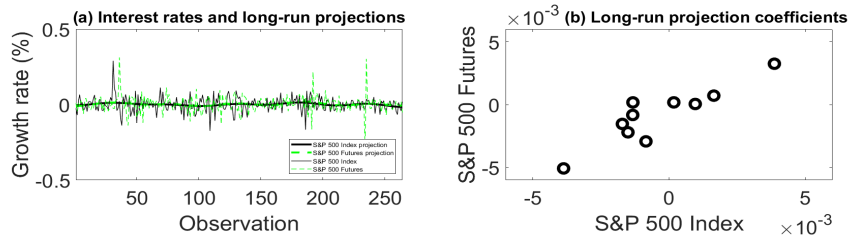
(b) Growth rates (in %) of S&P 500 Index and ten minutes ahead S&P 500 Futures



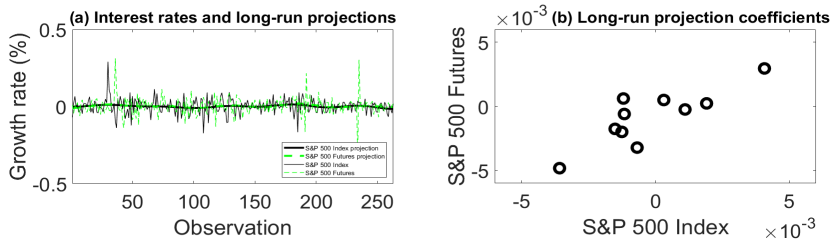
(c) Growth rates (in %) of S&P 500 Index and fifteen minutes ahead S&P 500 Futures



(d) Growth rates (in %) of S&P 500 Index and twenty minutes ahead S&P 500 Futures



(e) Growth rates (in %) of S&P 500 Index and 25 minutes ahead S&P 500 Futures



(f) Growth rates (in %) of S&P 500 Index and 30 minutes ahead S&P 500 Futures

Figure 20: Growth rates (in %) of S&P 500 Index and $n * 5$ minutes ahead S&P 500 Futures per 5 minutes from 12/06/2019 7:05 pm to 17/06/2019 21:55 pm (UTC +2:00), $n = 1, 2, 3, 4, 5, 6$

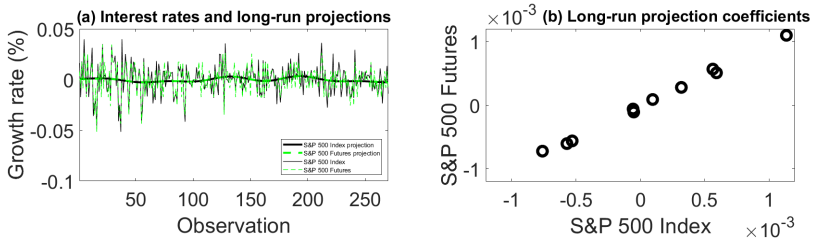
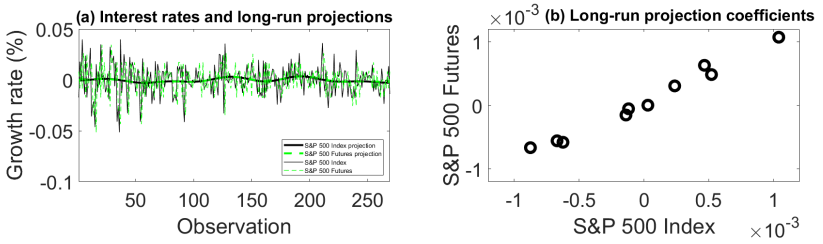
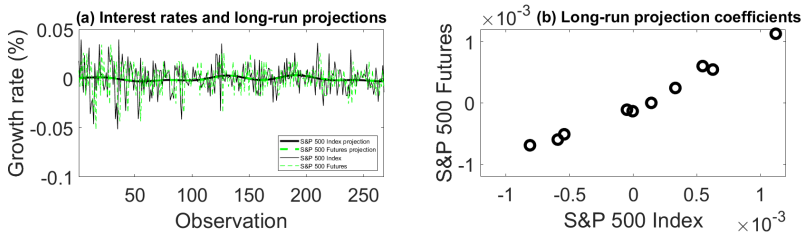


Figure 21: Minutely growth rates (in %) of S&P 500 Futures and S&P 500 Index on 17/06/2019 from 4:30 pm to 20:59 pm (UTC +2:00)

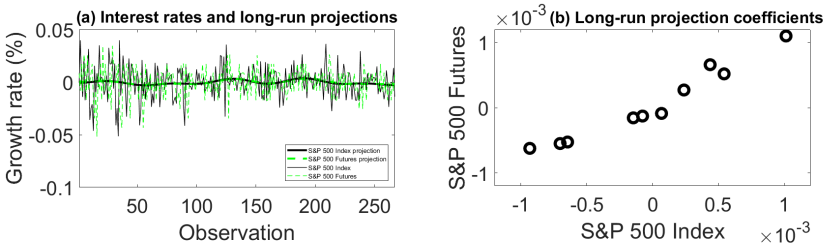
Plots on minutely growth rates



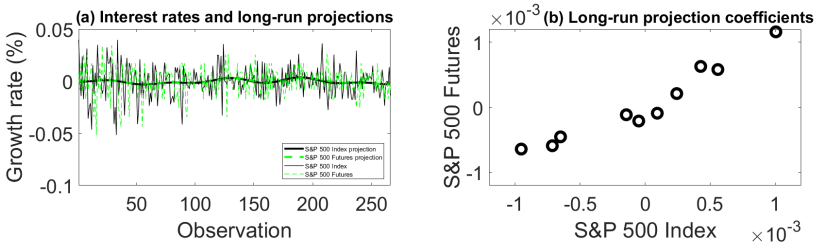
(a) Minutely growth rates (in %) of S&P 500 Index and one minute ahead S&P 500 Futures



(b) Minutely growth rates (in %) of S&P 500 Index and two minutes ahead S&P 500 Futures



(c) Minutely growth rates (in %) of S&P 500 Index and three minutes ahead S&P 500 Futures



(d) Minutely growth rates (in %) of S&P 500 Index and four minutes ahead S&P 500 Futures

Figure 22: Minutely growth rates (in %) of S&P 500 Index and n minutes ahead S&P 500 Futures on 17/06/2019 from 4:30 pm to 21:55 pm (UTC +2:00), $n = 1, 2, 3, 4$

Appendix G. Analysis of the impact of different data frequencies

Table 9: (A, B, c, d)-model parameters on monthly and weekly data: periods longer than 3.61 years

	ρ	β	$\sigma_{y\|x}$
	a) Monthly data		
Estimate	0.99	1.01	0.01
67% CI	0.99, 0.99	1.01, 1.01	0.01, 0.02
90% CI	0.98, 0.99	1.00, 1.02	0.01, 0.02
67% Bayes CS	0.99, 0.99	1.01, 1.01	0.01, 0.02
90% Bayes CS	0.98, 0.99	1.00, 1.02	0.01, 0.02
	b) Weekly data		
Estimate	0.99	1.00	0.00
67% CI	0.99, 0.99	1.00, 1.00	0.00, 0.00
90% CI	0.98, 0.99	1.00, 1.01	0.00, 0.00
67% Bayes CS	0.99, 0.99	1.00, 1.00	0.00, 0.00
90% Bayes CS	0.98, 0.99	1.00, 1.01	0.00, 0.00

Periods longer than 3.61 years correspond to $q = 12$. Growth rates are observed from November 1997 to June 2019. The estimates are the posterior median based on the $I(d)$ model. The 67% and 90% confidence intervals and Bayes credible sets based on the posterior are displayed below the estimates.

Table 10: (A, B, c, d)-model parameters on weekly and daily data: periods longer than 1.09 years

	ρ	β	$\sigma_{y\|x}$
	a) Weekly data		
Estimate	0.99	1.01	0.01
67% CI	0.99, 0.99	1.00, 1.01	0.00, 0.01
90% CI	0.98, 0.99	0.99, 1.02	0.00, 0.01
67% Bayes CS	0.99, 0.99	1.00, 1.01	0.00, 0.01
90% Bayes CS	0.98, 0.99	0.99, 1.02	0.00, 0.01
	b) Daily data		
Estimate	0.97	1.16	0.01
67% CI	0.94, 0.98	1.07, 1.25	0.01, 0.01
90% CI	0.90, 0.99	0.90, 1.32	0.01, 0.02
67% Bayes CS	0.94, 0.98	1.07, 1.25	0.01, 0.01
90% Bayes CS	0.91, 0.99	0.99, 1.32	0.01, 0.02

Periods longer than 1.09 years correspond to $q = 10$. Growth rates are observed from January 2014 to June 2019.The estimates are the posterior median based on the $I(d)$ model. The 67% and 90% confidence intervals and Bayes credible sets based on the posterior are displayed below the estimates.

Appendix H. Sub-sample analysis

Table 11: (A, B, c, d)-model parameters on weekly data: periods longer than 1.09 years

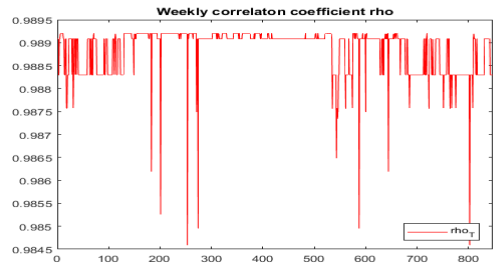
	ρ	β	$\sigma_{y\|x}$
a) Sub-sample from 02/11/1997 to 23/03/2003			
Estimate	0.99	1.00	0.02
67% CI	0.98, 0.99	0.99, 1.01	0.01, 0.02
90% CI	0.98, 0.99	0.98, 1.02	0.01, 0.03
67% Bayes CS	0.98, 0.99	0.99, 1.01	0.01, 0.02
90% Bayes CS	0.98, 0.99	0.98, 1.02	0.01, 0.03
b) Sub-sample from 30/03/2003 to 17/08/2008			
Estimate	0.99	1.01	0.01
67% CI	0.98, 0.99	1.00, 1.02	0.00, 0.01
90% CI	0.98, 0.99	1.00, 1.02	0.00, 0.01
67% Bayes CS	0.98, 0.99	1.00, 1.02	0.00, 0.01
90% Bayes CS	0.98, 0.99	1.00, 1.02	0.00, 0.01
c) Sub-sample from 24/08/2008 to 12/01/2014			
Estimate	0.99	0.99	0.01
67% CI	0.99, 0.99	0.99, 0.99	0.00, 0.01
90% CI	0.99, 0.99	0.99, 0.99	0.00, 0.01
67% Bayes CS	0.99, 0.99	0.99, 0.99	0.00, 0.01
90% Bayes CS	0.99, 0.99	0.99, 0.99	0.00, 0.01
d) Sub-sample from 19/01/2014 to 09/06/2019			
Estimate	0.99	1.01	0.01
67% CI	0.99, 0.99	1.00, 1.01	0.00, 0.01
90% CI	0.98, 0.99	0.99, 1.02	0.00, 0.01
67% Bayes CS	0.99, 0.99	1.00, 1.01	0.00, 0.01
90% Bayes CS	0.98, 0.99	0.99, 1.02	0.00, 0.01

Periods longer than 1.09 years correspond to $q = 10$. Growth rates are observed for three disjoint sub-periods between 02/11/1997 and 12/01/2014. The estimates are the posterior median based on the $I(d)$ model. The 67% and 90% confidence intervals and Bayes credible sets based on the posterior are displayed below the estimates.

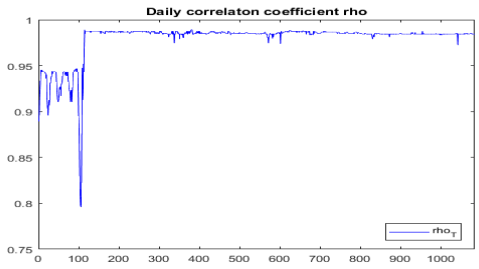
Table 12: (A, B, c, d)-model parameters on daily data: periods longer than 0.21 years

	ρ	β	$\sigma_{y\|x}$
a) Sub-sample from 20/01/2014 to 18/02/2015			
Estimate	0.98	1.07	0.01
67% CI	0.98, 0.99	1.01, 1.14	0.01, 0.02
90% CI	0.97, 0.99	1.00, 1.15	0.01, 0.02
67% Bayes CS	0.98, 0.99	1.03, 1.12	0.01, 0.02
90% Bayes CS	0.98, 0.99	1.00, 1.15	0.01, 0.02
b) Sub-sample from 19/02/2015 to 18/03/2016			
Estimate	0.98	1.02	0.02
67% CI	0.97, 0.99	0.98, 1.06	0.02, 0.03
90% CI	0.96, 0.99	0.95, 1.09	0.01, 0.04
67% Bayes CS	0.97, 0.99	0.98, 1.06	0.02, 0.03
90% Bayes CS	0.96, 0.99	0.95, 1.09	0.01, 0.04
c) Sub-sample from 21/03/2016 to 17/04/2017			
Estimate	0.98	0.99	0.01
67% CI	0.97, 0.99	0.95, 1.04	0.01, 0.01
90% CI	0.96, 0.99	0.92, 1.07	0.01, 0.02
67% Bayes CS	0.97, 0.99	0.95, 1.04	0.01, 0.01
90% Bayes CS	0.96, 0.99	0.92, 1.07	0.01, 0.02
d) Sub-sample from 18/04/2017 to 15/05/2018			
Estimate	0.99	0.95	0.01
67% CI	0.97, 0.99	0.92, 0.99	0.01, 0.01
90% CI	0.96, 0.99	0.89, 1.01	0.01, 0.02
67% Bayes CS	0.97, 0.99	0.92, 0.99	0.01, 0.01
90% Bayes CS	0.96, 0.99	0.89, 1.01	0.01, 0.02
e) Sub-sample from 16/05/2018 to 10/06/2019			
Estimate	0.89	1.04	0.07
67% CI	0.80, 0.95	87, 1.20	0.06, 0.10
90% CI	0.68, 0.98	0.75, 1.34	0.05, 0.13
67% Bayes CS	0.80, 0.95	0.87, 1.20	0.06, 0.10
90% Bayes CS	0.68, 0.98	0.75, 1.34	0.05, 0.13

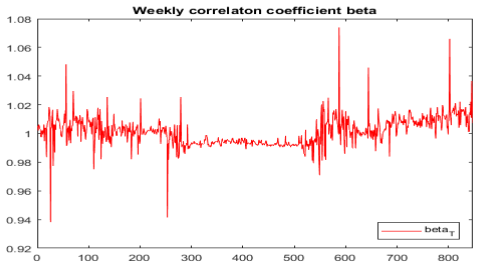
Periods longer than 0.21 years correspond to $q = 10$. Growth rates are observed for four disjoint sub-periods between 20/01/2014 and 10/06/2019. The estimates are the posterior median based on the $I(d)$ model. The 67% and 90% confidence intervals and Bayes credible sets based on the posterior are displayed below the estimates.



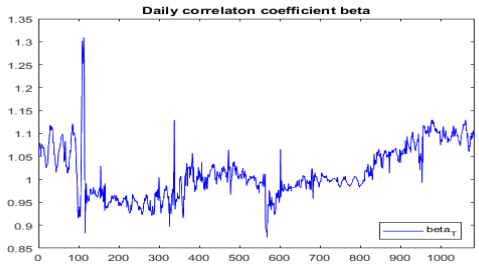
(a) ρ_T for 282 pairs of weekly observations on a rolling window with steps equal to one week between November 1997 and June 2019



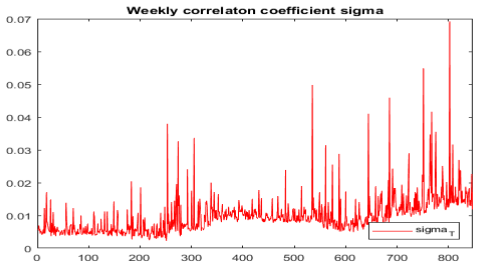
(b) ρ_T for 271 pairs of daily observations on a rolling window with steps equal to one day between January 2014 and June 2019



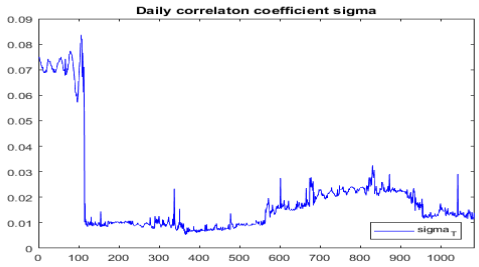
(c) β_T for 282 pairs of weekly observations on a rolling window with steps equal to one week between November 1997 and June 2019



(d) β_T for 271 pairs of daily observations on a rolling window with steps equal to one day between January 2014 and June 2019



(e) $\sigma_{y|x,T}$ for 282 pairs of weekly observations on a rolling window with steps equal to one week between November 1997 and June 2019



(f) $\sigma_{y|x,T}$ for 271 pairs of daily observations on a rolling window with steps equal to one day between January 2014 and June 2019

Figure 23: Long-run covariability parameters on rolling windows between November 2997 and June 2019 (weekly observations) and January 2014 and June 2019 (daily observations).

Appendix I. Estimation of the S&P 500 Futures lead time

Table 13: (A, B, c, d)-model parameters on hourly S&P 500 Index and n hours ahead S&P 500 Futures, $n = 0, 1, 2$: periods longer than 8.25 trading days

	ρ	β	$\sigma_{y\|x}$
a) S&P 500 Index and 0 hours ahead S&P 500 Futures			
Estimate	0.99	1.02	0.00
67% CI	0.99, 0.99	1.01, 1.03	0.00, 0.00
90% CI	0.98, 0.99	0.99, 1.05	0.00, 0.01
67% Bayes CS	0.99, 0.99	1.01, 1.03	0.00, 0.00
90% Bayes CS	0.98, 0.99	1.00, 1.04	0.00, 0.01
b) S&P 500 Index and one hour ahead S&P 500 Futures			
Estimate	0.99	1.02	0.01
67% CI	0.98, 0.99	0.99, 1.04	0.00, 0.01
90% CI	0.97, 0.99	0.98, 1.06	0.00, 0.01
67% Bayes CS	0.98, 0.99	0.99, 1.04	0.00, 0.01
90% Bayes CS	0.97, 0.99	0.98, 1.06	0.00, 0.01
c) S&P 500 Index and two hours ahead S&P 500 Futures			
Estimate	0.94	1.15	0.03
67% CI	0.88, 0.97	1.02, 1.30	0.02, 0.04
90% CI	0.81, 0.98	0.92, 1.41	0.02, 0.06
67% Bayes CS	0.88, 0.97	1.02, 1.30	0.02, 0.04
90% Bayes CS	0.81, 0.98	0.92, 1.41	0.02, 0.06

Periods longer than 8.25 trading days correspond to $q = 10$. Growth rates are observed from 24/02/2019 4pm to 18/06/2019 4pm (UTC +2:00).The estimates are the posterior median based on the $I(d)$ model. The 67% and 90% confidence intervals and Bayes credible sets based on the posterior are displayed below the estimates.

Table 14: (A, B, c, d)-model parameters on S&P 500 Index and $n * 5$ minutes ahead S&P 500 Futures per 5 minutes, $n = 0, 1, 2, 3, 4, 5, 6$: periods longer than 269 trading minutes

	ρ	β	$\sigma_{y\|x}$
a) S&P 500 Index and 0 minutes ahead S&P 500 Futures			
Estimate	0.87	0.93	0.00
67% CI	0.72, 0.93	0.75, 1.10	0.00, 0.00
90% CI	0.57, 0.95	0.62, 1.24	0.00, 0.01
67% Bayes CS	0.72, 0.93	0.75, 1.10	0.00, 0.00
90% Bayes CS	0.57, 0.95	0.62, 1.24	0.00, 0.01
b) S&P 500 Index and 5 minutes ahead S&P 500 Futures			
Estimate	0.85	0.92	0.00
67% CI	0.70, 0.93	0.74, 1.08	0.00, 0.00
90% CI	0.57, 0.95	0.60, 1.23	0.00, 0.01
67% Bayes CS	0.70, 0.93	0.74, 1.08	0.00, 0.00
90% Bayes CS	0.57, 0.95	0.60, 1.23	0.00, 0.01
c) S&P 500 Index and 10 minutes ahead S&P 500 Futures			
Estimate	0.82	0.85	0.00
67% CI	0.64, 0.90	0.66, 1.03	0.00, 0.01
90% CI	0.48, 0.94	0.52, 1.18	0.00, 0.01
67% Bayes CS	0.64, 0.90	0.66, 1.03	0.00, 0.01
90% Bayes CS	0.48, 0.94	0.52, 1.18	0.00, 0.01
d) S&P 500 Index and 15 minutes ahead S&P 500 Futures			
Estimate	0.82	0.84	0.00
67% CI	0.64, 0.90	0.65, 1.02	0.00, 0.01
90% CI	0.47, 0.94	0.51, 1.17	0.00, 0.01
67% Bayes CS	0.64, 0.90	0.65, 1.02	0.00, 0.01
90% Bayes CS	0.47, 0.94	0.51, 1.17	0.00, 0.01
e) S&P 500 Index and 20 minutes ahead S&P 500 Futures			
Estimate	0.78	0.84	0.00
67% CI	0.60, 0.89	0.64, 1.04	0.00, 0.01
90% CI	0.41, 0.93	0.48, 1.21	0.00, 0.01
67% Bayes CS	0.60, 0.89	0.64, 1.04	0.00, 0.01
90% Bayes CS	0.41, 0.93	0.48, 1.21	0.00, 0.01
f) S&P 500 Index and 25 minutes ahead S&P 500 Futures			
Estimate	0.72	0.76	0.00
67% CI	0.56, 0.85	0.55, 0.95	0.00, 0.01
90% CI	0.36, 0.91	0.39, 1.12	0.00, 0.01
67% Bayes CS	0.56, 0.85	0.55, 0.95	0.00, 0.01
90% Bayes CS	0.36, 0.91	0.39, 1.12	0.00, 0.01
g) S&P 500 Index and 30 minutes ahead S&P 500 Futures			
Estimate	0.64	0.66	0.00
67% CI	0.39, 0.79	0.43, 0.89	0.00, 0.01
90% CI	0.16, 0.86	0.24, 1.08	0.00, 0.01
67% Bayes CS	0.39, 0.79	0.43, 0.89	0.00, 0.01
90% Bayes CS	0.16, 0.86	0.24, 1.08	0.00, 0.01

Periods longer than 269 trading minutes correspond to $q = 10$. Growth rates are observed from 12/06/2019 7:05pm to 17/06/2019 9:55pm (UTC +2:00).The estimates are the posterior median based on the $I(d)$ model. The 67% and 90% confidence intervals and Bayes credible sets based on the posterior are displayed below the estimates.

Table 15: (A, B, c, d)-model parameters on minutely S&P 500 Index and n minutes ahead S&P 500 Futures, $n = 0, 1, 2, 3, 4$: periods longer than 54 trading minutes

	ρ	β	$\sigma_{y\|x}$
a) S&P 500 Index and 0 minutes ahead S&P 500 Futures			
Estimate	0.99	0.97	0.00
67% CI	0.98, 0.99	0.95, 1.00	0.00, 0.00
90% CI	0.97, 0.99	0.91, 1.03	0.00, 0.00
67% Bayes CS	0.98, 0.99	0.95, 1.00	0.00, 0.00
90% Bayes CS	0.97, 0.99	0.91, 1.02	0.00, 0.00
b) S&P 500 Index and one minute ahead S&P 500 Futures			
Estimate	0.98	0.96	0.00
67% CI	0.97, 0.99	0.90, 1.00	0.00, 0.00
90% CI	0.95, 0.99	0.86, 1.04	0.00, 0.00
67% Bayes CS	0.97, 0.99	0.90, 1.00	0.00, 0.00
90% Bayes CS	0.95, 0.99	0.86, 1.04	0.00, 0.00
c) S&P 500 Index and two minutes ahead S&P 500 Futures			
Estimate	0.98	0.95	0.00
67% CI	0.97, 0.99	0.90, 1.00	0.00, 0.00
90% CI	0.95, 0.99	0.86, 1.04	0.00, 0.00
67% Bayes CS	0.97, 0.99	0.90, 1.00	0.00, 0.00
90% Bayes CS	0.95, 0.99	0.86, 1.04	0.00, 0.00
d) S&P 500 Index and three minutes ahead S&P 500 Futures			
Estimate	0.97	0.93	0.00
67% CI	0.93, 0.98	0.84, 1.00	0.00, 0.00
90% CI	0.88, 0.99	0.78, 1.07	0.00, 0.00
67% Bayes CS	0.93, 0.98	0.84, 1.00	0.00, 0.00
90% Bayes CS	0.88, 0.99	0.78, 1.07	0.00, 0.00
e) S&P 500 Index and four minutes ahead S&P 500 Futures			
Estimate	0.95	0.92	0.00
67% CI	0.90, 0.97	0.83, 1.02	0.00, 0.00
90% CI	0.84, 0.98	0.75, 1.10	0.00, 0.00
67% Bayes CS	0.90, 0.97	0.83, 1.02	0.00, 0.00
90% Bayes CS	0.84, 0.98	0.75, 1.10	0.00, 0.00

Periods longer than 54 trading minutes correspond to $q = 10$. Growth rates are observed from 17/06/2019 4:30 pm to 17/06/2019 20:59pm (UTC +2:00).The estimates are the posterior median based on the $I(d)$ model. The 67% and 90% confidence intervals and Bayes credible sets based on the posterior are displayed below the estimates.

Appendix J. Code description

The replication zip-file consists of three main folders:

- The folder "'data" contains all raw data that is used to find obtained results. "final data.xlsx" contains the data on the extension part. The other files and particularly "lr_correlations.xlsx" contain the data on the replication of the results found by Müller and Watson (2018).
- The folder "matlab" contains all code that is used to obtain the results of the replication and extension part of this paper.
- The folder "simulation" contains all code that is used to obtain the simulation results.

The code in the folder "matlab" is structured as follows:

- The folder "m_utilities" consists of matlab files that are useful to the application of the methods that are used by Müller and Watson. For example, in this folder codes can be found that remove empty observations from the data set and perform the data transformation (psi_compute.mat, xp_compute.mat). All code in this folder is entirely provided by Müller and Watson (2018).
- The folder "initial_analysis" consists of code files for which output is included in the replication material of Müller and Watson (2018). By this reason, the code that is provided by Müller and Watson (2018) is not adjusted and only individual comments are added to the code.

- Whether the files are used in order to obtain the results from the replication part or the extension part is made clear in the titles of the matlab files. The files in the sub-folder "matlab_procs" and lrcov.mat are used for both the replication and the extension analysis to compute the long-run covariability parameter estimates and confidence intervals. All code files that are used to obtain the replication results and large parts of the code files that are used to obtain the extension results are provided by Müller and Watson (2018) and only individual comments are added to these files. All code files and explanations of what the code files do are listed below:
  - Figure_1_replication: this file plots average growth rates, long-run projections and projection coefficients.
  - Figure_1a_replication: this file plots average growth rates over subsamples.
  - Figure_2_extension: this file plots long-run projections and projection coefficients of extension data.
  - Figure_2_replication: this file plots long-run projections and projection coefficients of replication data.
  - lr_correlations_data_calendar_replication: this file sets up the data and calendars for the replication analysis.
  - lrcov: this function computes long-run covariability statistics for two variables.
  - Table_1_a_replication: this file constructs the table containing by other functions derived parameter values and confidence intervals for

GDP and consumption using the $I(0)$ and $I(1)$ model.

- Table_1_b_replication: this file constructs the table containing by other functions derived parameter values and confidence intervals for long-term and short-term interest rates using the $I(0)$ and $I(1)$ model.
- Table_2_a_replication: this file constructs the table containing by other functions derived parameter values and confidence intervals for GDP and consumption using the (A, B, c, d) model.
- Table_2_b_replication: this file constructs the table containing by other functions derived parameter values and confidence intervals for short-term and long-term interest rates using the (A, B, c, d) model.
- Table_2_extension: this file constructs the table containing by other functions derived parameter values and confidence intervals using the (A, B, c, d) model.
- Table_2_rollingwindows_total: this file has as output a vector of all parameter values that is plotted in figure 23.

The following files can be found in folder "matlab_procs":

- b_compute: this file constructs results for beta.
- beta_compute_NaN: this file computes beta for the $I(0)$ and the $I(1)$ model.
- beta_compute_varcorr: this file computes the estimate of beta.
- beta_grid: this file computes Grid of values for beta.
- compute_avg_like_h0: this file computes lambda-weighted average density under $H_0 : \gamma = \gamma_0$.

- compute_avg_like_h1: this file computes the average density under $H_1 : \theta \sim W$.
- compute_like_h1: computes the posterior under the uniform prior distribution.
- corr_compute: this file constructs results for rho.
- cosine_transforms: this file computes cosine transforms and standardized cosine transforms
- df_pct: computes the percentiles of each density.
- getdens_2xscale: this file derives the density of ρ under H_0.
- getdens_2xscale_sym: this file derives the symmetric density of ρ under H_0.
- getdens_tri_beta0: this file derives the density of β under H_0.
- getdens_tri_beta1: this file derives the density of β under H_1.
- getdens_tri_inv: this file derives the density of the maximal invariant.
- getdens_tri_stddev0: this file derives the density of $\sigma_{y|x}$ under H_1.
- getdens_tri_stddev1: this file derives the density of $\sigma_{y|x}$ under H_1.
- getni: this file computes the $(q-1)$-th power of the normal distribution with mean β.
- getnima: this file evaluates normal integral with powers equal to -1 on both coefficients, non-symmetrized.
- getnimd: this file evaluates normal integral with powers equal to q - 2 on a and q-1 on beta.

  - getnime: this file evaluates normal integral with powers equal to q-1 on both coefficients, symmetrized.
  - I0_varcorr: this file computes the long-run covariability parameter valuesusing the $I(0)$ model.
  - I1_varcorr: this file computes the long-run covariability parameter valuesusing the $I(1)$ model.
  - lower_tri_invariant: this file ensures that the input, matrix X, is invariant.
  - sigma_grid: this file computes Grid of values for sigma.
  - stddev_compute: this file constructs results for sigma.

All files in the sub-folders "matlab_procs" (listed above) and "m_utilities" are copied to the final main folder "simulation". Other files that are needed to obtain the simulation results are listed below:

- lrcov_Simulation: this function computes long-run covariability statistics for two simulated data series.
- histfit: this file computes parameter estimate, empirical confidence intervals and average estimated confidence intervals, and plots the histograms for the long-run covariability parameters.
- Simulation: This file defines the parameters of the $ARFIMA$ model that is used to simulate the data.
- Simulation_frac_integrated: This file simulates fractionally integrated series according to the given $ARFIMA$ parameters.

- Table_1_Simulation: this file computes 1000 times the long-run covariability parameters between different simulated series using the $I(0)$ and the $I(1)$ model.
- Table_2_Simulation: this file computes 1000 times the long-run covariability parameters between different simulated series using the (A, B, c, d) model.