Project Description

Applicant: Michael Vogt, University of Bonn
Project Title: New Methods and Theory for the Comparison of Nonparametric Trend Curves

1 State of the art and preliminary work

The comparison of nonparametric curves is a classic topic in econometrics and statistics. Depending on the specific application, the curves of interest are densities, distribution functions, time trends or regression curves. The problem of testing for equality of densities has been studied in Mammen (1992), Anderson et al. (1994) and Li et al. (2009) among others. Tests for equality of distribution functions can be found for example in Kiefer (1959), Anderson (1962) and Finner and Gontscharuk (2018). Tests for equality of trend or regression curves have been developed in Härdle and Marron (1990), Hall and Hart (1990), Delgado (1993), Degras et al. (2012), Zhang et al. (2012) and Hidalgo and Lee (2014) among many others. In the proposed project, we focus on the comparison of nonparametric trend curves.

The statistical problem of comparing trends has a wide range of applications in economics, finance and other fields such as climatology and biology. In economics, one may wish is to compare trends in real gross domestic product (GDP) across different countries (cp. Grier and Tullock, 1989). Another example concerns the dynamics of long-term interest rates. To better understand these dynamics, researchers aim to compare the yields of US Treasury bills at different maturities over time (cp. Park et al., 2009). In finance, it is of interest to compare the volatility trends of different stocks (cp. Nyblom and Harvey, 2000). Finally, in climatology, researchers are interested in comparing the trending behaviour of temperature time series across different spatial locations (cp. Karoly and Wu, 2005).

Classically, time trends are modelled stochastically in econometrics; see e.g. Stock and Watson (1988). Recently, however, there has been a growing interest in econometric models with deterministic time trends; see Cai (2007), Atak et al. (2011), Robinson (2012) and Chen et al. (2012) among others. Non- and semiparametric trend modelling has attracted particular interest in a panel data context. Li et al. (2010), Atak et al. (2011), Robinson (2012) and Chen et al. (2012) considered panel models where the observed time series have a common time trend. In many applications, however, the assumption of a common time trend is quite harsh. In particular when the number of observed time series is large, it is quite natural to suppose that the time trend may differ across time series. More flexible panel settings with heterogeneous trends have been studied, for example, in Zhang et al. (2012) and Hidalgo and Lee (2014).

In what follows, we consider a general panel framework with heterogeneous trends which is useful for a number of economic and financial applications: Suppose we observe a panel of n time series $\mathcal{Z}_i = \{(Y_{it}, X_{it}) : 1 \le t \le T\}$ for $1 \le i \le n$, where Y_{it} are real-valued random variables and $X_{it} = (X_{it,1}, \ldots, X_{it,d})^\top$ are d-dimensional random vectors. Each time series $\mathcal{Z}_i$ is modelled by the equation

$$Y_{it} = m_i\Big(\frac{t}{T}\Big) + \beta_i^\top X_{it} + \alpha_i + \varepsilon_{it} \tag{1}$$

for $1 \le t \le T$, where $m_i : [0,1] \to \mathbb{R}$ is a nonparametric (deterministic) trend function, X_{it} is a vector of regressors or controls and β_i is the corresponding parameter vector. Moreover, α_i are so-called fixed effect error terms and ε_{it} are standard regression errors with $\mathbb{E}[\varepsilon_{it}|X_{it}] = 0$ for all t. Model (1) nests a number of panel models which have recently been considered in the literature. Special cases of model (1) with a nonparametric trend specification are for example considered in Atak et al. (2011), Zhang et al. (2012) and Hidalgo and Lee (2014). Versions of model (1) with a parametric trend are studied in Vogelsang and Franses (2005), Sun (2011) and Xu (2012) among others.

As usual in nonparametric regression, the trend functions m_i in model (1) depend on rescaled time t/T rather than on real time t; cp. Robinson (1989), Dahlhaus (1997) and Vogt and Linton (2014) for the use and some discussion of the rescaled time argument. The functions m_i are only identified up to an additive constant in model (1): One can reformulate the model as $Y_{it} = [m_i(t/T) + c_i] + \beta_i^\top X_{it} + [\alpha_i - c_i] + \varepsilon_{it}$, that is, one can freely shift additive constants c_i between the trend $m_i(t/T)$ and the error component α_i. In order to obtain identification, one may impose different normalization constraints on the trends m_i. One possibility is to normalize them such that $\int_0^1 m_i(u)du = 0$ for all i. In what follows, we take for granted that the trends m_i satisfy this constraint.

Within the general framework of model (1), we can formulate a number of interesting statistical questions concerning the set of trend functions $\{m_i : 1 \le i \le n\}$.

(a) Testing for equality of nonparametric trend curves

In many application contexts, an important question is whether the time trends m_i in model (1) are all the same. Put differently, the question is whether the observed time series have a common trend. This question can formally be addressed by a statistical test of the null hypothesis

$$H_0 : \text{There exists a function } m : [0,1] \to \mathbb{R} \text{ such that } m_i = m \text{ for all } 1 \le i \le n.$$

A closely related question is whether all time trends have the same parametric form.

To formulate the corresponding null hypothesis, let $m(\theta,\cdot):[0,1]\to\mathbb{R}$ be a function which is known up to the finite-dimensional parameter $\theta\in\Theta$, where Θ denotes the parameter space. The null hypothesis of interest now reads as follows:

$$H_{0,\text{para}}:\ \text{There exists } \theta\in\Theta \text{ such that } m_i(\cdot)=m(\theta,\cdot) \text{ for all } 1\le i\le n.$$

If $m(\theta,w)=a+bw$ with $\theta=(a,b)$, for example, then H_0 is the hypothesis that all trends m_i are linear with the same intercept a and slope b. A somewhat simpler but yet important hypothesis is given by

$$H_{0,\text{const}}: m_i\equiv 0 \text{ for all } 1\le i\le n.$$

Under this hypothesis, there is no time trend at all in the observed time series. Put differently, all the time trends m_i are constant. (Note that under the normalization constraint $\int_0^1 m_i(w)dw=0$, m_i must be equal to zero if it is a constant function.) A major goal of our project is to develop new tests for the hypotheses H_0, $H_{0,\text{para}}$ and $H_{0,\text{const}}$ in model (1). In order to keep the exposition as clear as possible, we focus attention on the hypothesis H_0 in what follows. Tests of $H_{0,\text{para}}$, $H_{0,\text{const}}$ and related hypotheses have for example been studied in Lyubchich and Gel (2016) and Chen and Wu (2018).

In recent years, a number of different approaches have been developed to test the hypothesis H_0. Degras et al. (2012) consider the problem of testing H_0 within the model framework

$$Y_{it}=m_i\Big(\frac{t}{T}\Big)+\alpha_i+\varepsilon_{it}\qquad(1\le t\le T,\,1\le i\le n),\tag{2}$$

where $\mathbb{E}[\varepsilon_{it}]=0$ for all i and t and the terms α_i are assumed to be deterministic. Obviously, (2) is a special case of (1) which does not include additional regressors. Degras et al. (2012) construct an L_2-type statistic to test H_0. The statistic is based on the difference between estimators of the trend with and without imposing H_0. Let $\hat{m}_{i,h}$ be the estimator of m_i and $\hat{m}_h$ the estimator of the common trend m under H_0, where h denotes the bandwidth parameter. With these estimators, the authors define the statistic

$$\Delta_{n,T}=\sum_{i=1}^n\int_0^1\big(\hat{m}_{i,h}(u)-\hat{m}_h(u)\big)^2du,\tag{3}$$

which measures the L_2-distance between $\hat{m}_{i,h}$ and $\hat{m}_h$. In the theoretical part of their paper, they derive the limit distribution of $\Delta_{n,T}$. Chen and Wu (2018) develop theory for test statistics closely related to those from Degras et al. (2012), but under more general conditions on the error terms.

Zhang et al. (2012) investigate the problem of testing the hypothesis H_0 in a slightly restricted version of model (1), where $\beta_i=\beta$ for all i. The regression coefficients β_i

are thus assumed to be homogeneous in their setting. They construct a residual-based test statistic as follows: First, they obtain profile least squares estimators $\hat{\beta}$ and $\hat{m}_h(t/T)$ of the parameter vector β and the common trend m under H_0, where h denotes the bandwidth. With these estimators, they compute the residuals $\hat{u}_{it} = Y_{it} - \hat{\beta}^T X_{it} - \hat{m}_h(t/T)$. These residuals are shown to have the form $\hat{u}_{it} = \Delta_i(t/T) + \eta_{it}$, where Δ_i is a deterministic function with the property that $\Delta_i \equiv 0$ under H_0 and η_{it} denotes the error term. Testing H_0 is thus equivalent to testing the hypothesis $H_0' : \Delta_i \equiv 0$ for all $1 \leq i \leq n$. The authors construct a test statistic for the hypothesis H_0' on the basis of nonparametric kernel estimators of the functions Δ_i and derive its limit distribution.

The tests of Zhang et al. (2012), Degras et al. (2012) and Chen and Wu (2018) are based on nonparametric estimators of the trend functions m_i that depend on one or several bandwidth parameters. Unfortunately, it is far from clear how to choose these bandwidths in an appropriate way. This is a general problem concerning essentially all tests based on nonparametric curve estimators. There are of course many theoretical results on optimal bandwidth choice for estimation purposes. However, the optimal bandwidth for curve estimation is usually not optimal for testing. Optimal bandwidth choice for tests is indeed an open problem, and only little theory for simple cases is available (cp. Gao and Gijbels, 2008). Since tests based on nonparametric curve estimators are commonly quite sensitive to the choice of bandwidth and theory for optimal bandwidth selection is not available, it appears preferable to work with bandwidth-free tests.

A classical way to obtain a bandwidth-free test of the hypothesis H_0 is to use CUSUM-type statistics which are based on partial sum processes. This approach is taken in Hidalgo and Lee (2014). A more modern approach to obtain a bandwidth-free test is to employ multiscale methods. These methods avoid the need to choose a bandwidth by considering a large collection of bandwidths simultaneously. More specifically, the basic idea is as follows: Let S_h be a test statistic for the null hypothesis of interest, which depends on the bandwidth h. Rather than considering only a single statistic S_h for a specific bandwidth h, a multiscale approach simultaneously considers a whole family of statistics $\{S_h : h \in \mathcal{H}\}$, where $\mathcal{H}$ is a set of bandwidth values. The multiscale test then proceeds as follows: For each bandwidth or scale h, one checks whether $S_h > q_h(\alpha)$, where $q_h(\alpha)$ is a bandwidth-dependent critical value (for given significance level α). The multiscale test rejects if $S_h > q_h(\alpha)$ for at least one scale h. The main theoretical difficulty in this approach is of course to derive appropriate critical values $q_h(\alpha)$. Specifically, the critical values $q_h(\alpha)$ need to be determined such that the multiscale test has the correct (asymptotic) level, that is, such that $\mathbb{P}(S_h > q_h(\alpha)$ for some $h \in \mathcal{H}) = (1 - \alpha) + o(1)$.

Multiscale methods have been developed for a variety of different test problems in recent years. Chaudhuri and Marron (1999, 2000) introduced the so-called SiZer method which has been extended in various directions; see for example Hannig and Marron (2006) and Rondonotti et al. (2007). Horowitz and Spokoiny (2001) proposed a multiscale test for the parametric form of a regression function. Dümbgen and Spokoiny (2001) constructed a multiscale approach which works with additively corrected supremum statistics. This general approach has been very influential in recent years and has been further developed in numerous ways; see for example Dümbgen (2002), Rohde (2008) and Proksch et al. (2018) for multiscale methods in the regression context and Dümbgen and Walther (2008), Rufibach and Walther (2010), Schmidt-Hieber et al. (2013) and Eckle et al. (2017) for methods in the context of density estimation. Importantly, all of these studies are restricted to the case of independent data. It turns out that it is highly non-trivial to extend the multiscale approach of Dümbgen and Spokoiny (2001) to the case of dependent data. A first step into this direction has recently been made in Khismatullina and Vogt (2018). They developed multiscale methods to test for local increases/decreases of the nonparametric trend function m in the univariate time series model $Y_t = m(t/T) + \varepsilon_t$.

To the best of our knowledge, multiscale tests of the hypotheses H_0, $H_{0,\text{para}}$ and $H_{0,\text{const}}$ in model (1) are not available in the literature. The only exception is Park et al. (2009) who developed SiZer methods for the comparison of nonparametric trend curves in a strongly simplified version of model (1). Their analysis, however, is mainly methodological and not fully backed up by theory. Indeed, theory has only been derived for the special case $n = 2$, that is, for the case that only two time series are observed.

(b) Clustering of nonparametric trend curves

Consider the situation that the null hypothesis $H_0 : m_1 = \ldots = m_n$ is violated in the general panel data model (1). Even though some of the trend functions m_i are different in this case, there may still be groups of time series with the same time trend. Formally, a group stucture can be defined as follows within the framework of model (1): There exist sets or groups of time series $G_1, \ldots, G_{K_0}$ with $\{1, \ldots, n\} = \dot{\bigcup}_{k=1}^{K_0} G_k$ such that for each $1 \le k \le K_0$,

$$m_i = m_j \quad \text{for all } i, j \in G_k. \tag{4}$$

According to (4), the time series of a given group G_k all have the same time trend. In many applications, it is very natural to suppose that there is such a group structure in the data. An interesting statistical problem which we aim to investigate in our project is how to estimate the unknown groups $G_1, \ldots, G_{K_0}$ and their unknown number K_0 from the data.

Several approaches to this problem have been proposed in the context of models closely related to (1). Degras et al. (2012) used a repeated testing procedure based on L_2-type test statistics of the form (3) in order to estimate the unknown group structure in model (2). Zhang (2013) developed a clustering method within the same model framework which makes use of an extended Bayesian information criterion. Vogt and Linton (2017) constructed a thresholding method to estimate the unknown group structure in the panel model $Y_{it} = m_i(X_{it}) + u_{it}$, where X_{it} are random regressors and u_{it} are general error terms that may include fixed effects. Their approach can also be adapted to the case of fixed regressors $X_{it} = t/T$. As an alternative to a group structure, factor-type structures have been imposed on the trend and regression functions in panel models. Such factor-type structures are studied in Kneip et al. (2012), Boneva et al. (2015) and Boneva et al. (2016) among others.

The problem of estimating the unknown groups $G_1, \ldots, G_{K_0}$ and their unknown number K_0 in model (1) has close connections to functional data clustering. There, the aim is to cluster smooth random curves that are functions of (rescaled) time and that are observed with or without noise. A number of different clustering approaches have been proposed in the context of functional data models; see for example Abraham et al. (2003), Tarpey and Kinateder (2003) and Tarpey (2007) for procedures based on k-means clustering, James and Sugar (2003) and Chiou and Li (2007) for model-based clustering approaches and Jacques and Preda (2014) for a recent survey.

The problem of finding the unknown group structure in model (1) is also closely related to a developing literature in econometrics which aims to identify unknown group structures in parametric panel regression models. In its simplest form, the panel regression model under consideration is given by the equation $Y_{it} = \beta_i^\top X_{it} + u_{it}$ for $1 \le t \le T$ and $1 \le i \le n$, where the coefficient vectors β_i are allowed to vary across individuals i and the error terms u_{it} may include fixed effects. Similar to the trend functions in model (1), the coefficients β_i are assumed to belong to a number of groups: there are K_0 groups $G_1, \ldots, G_{K_0}$ such that $\beta_i = \beta_j$ for all $i, j \in G_k$ and all $1 \le k \le K_0$. The problem of estimating the unknown groups and their unknown number has been studied in different versions of this modelling framework; cp. Su et al. (2016), Su and Ju (2018) and Wang et al. (2018) among others. Bonhomme and Manresa (2015) considered a related model where the group structure is not imposed on the regression coefficients but rather on some unobserved time-varying fixed effect components of the panel model.

Virtually all the proposed procedures to cluster nonparametric curves in panel and functional data models related to (1) depend on a number of bandwidth or smoothing parameters required to estimate the nonparametric functions m_i. In general, nonparametric curve estimators are strongly affected by the chosen bandwidth parameters. A clustering procedure which is based on such estimators can be expected to be strongly influenced by the choice of bandwidths as well. Moreover, as in the context of statistical testing, there is no theory available on how to pick the bandwidths optimally for the clustering problem. Hence, as in the context of testing, it is desirable to construct a clustering procedure which is free of bandwidth or smoothing parameters that need to be selected.

One way to obtain a clustering method which does not require to select any bandwidth parameter is to use multiscale methods. This approach has recently been taken in Vogt and Linton (2018). They develop a clustering approach in the context of the panel model $Y_{it} = m_i(X_{it}) + u_{it}$, where X_{it} are random regressors and u_{it} are general error terms that may include fixed effects. Imposing the same group structure as in (4) on their model, they construct estimators of the unknown groups and their unknown number as follows: In a first step, they develop bandwidth-free multiscale statistics $\hat{d}_{ij}$ which measure the distance between pairs of functions m_i and m_j. To construct them, they make use of the multiscale testing methods described in part (a) of this section. In a second step, the statistics $\hat{d}_{ij}$ are employed as dissimilarity measures in a hierarchical clustering algorithm.

1.1 Project-related publications

1.1.1 Articles published by outlets with scientific quality assurance, book publications, and works accepted for publication but not yet published

Boneva, L. and Linton, O. and Vogt, M. (2015). A semiparametric model for heterogeneous panel data with fixed effects. *Journal of Econometrics*, **188** 327-345.

Boneva, L. and Linton, O. and Vogt, M. (2016). The effect of fragmentation in trading on market quality in the UK equity market. *Journal of Applied Econometrics*, **31** 192-213.

Vogt, M. and Linton, O. (2017). Classification of non-parametric regression functions in longitudinal data models. *Journal of the Royal Statistical Society: Series B*, **79** 5-27.

1.1.2 Other publications

Khismatullina, M. and Vogt, M. (2018). Multiscale inference and long-run variance estimation in nonparametric regression with time series errors. *arXiv, ??.*

Vogt, M. and Linton, O. (2018). Multiscale clustering of nonparametric regression curves. *arXiv, ??.*

2 Objectives and work programme

2.1 Anticipated total duration of the project

2 years from October 1, 2019 to September 30, 2021.

2.2 Objectives

The main purpose of the project is to develop new methods and theory for the comparison and clustering of nonparametric trend curves. In particular, we aim to develop test and clustering methods which are free of classical bandwidth parameters and thus avoid the notoriously difficult problem of optimal bandwidth selection. To do so, we will build on the techniques from statistical multiscale testing reviewed in Section 1. The methodological and theoretical analysis of the project will be complemented by simulations and empirical applications. First of all, we will carry out a detailed simulation study to examine the finite sample performance of the proposed test and clustering methods. Moreover, we intend to demonstrate their usefulness by two empirical data examples which are described in more detail in the work programme in Section 2.3.

As a modelling framework, we will consider the general panel setting (1) introduced in Section 1. We briefly summarize the model setting once again for convenience: Suppose we observe a panel of n time series $\mathcal{Z}_i = \{(Y_{it}, X_{it}) : 1 \le t \le T\}$ for $1 \le i \le n$, where Y_{it} are real-valued random variables and X_{it} are d-dimensional random vectors. Each time series $\mathcal{Z}_i$ is modelled by the equation

$$Y_{it} = m_i\Big(\frac{t}{T}\Big) + \beta_i^\top X_{it} + \alpha_i + \varepsilon_{it} \tag{5}$$

for $1 \le t \le T$, where m_i is a nonparametric time trend curve, X_{it} is a vector of regressor or control variables, α_i are unobserved fixed effects and ε_{it} are idiosyncratic error terms with $\mathbb{E}[\varepsilon_{it}|X_{it}] = 0$. For each i, $\mathcal{P}_i = \{(X_{it}, \varepsilon_{it}) : 1 \le t \le T\}$ is assumed to be a general time series process which fulfills some weak dependence conditions (e.g. conditions formulated in terms of strong mixing coefficients or in terms of the physical dependence measure introduced by Wu (2005)). We will not only allow for time series dependence in the data, but also for cross-sectional dependence. To derive our theoretical results, we will assume that the time series length T tends to infinity. The number of time series n, in contrast, may either be bounded or diverging.

(a) Contributions to statistical multiscale testing

The first main challenge of the project is to develop novel multiscale tests for the comparison of the trend curves m_i in model (5). More specifically, we aim to develop multiscale tests for the hypotheses H_0, $H_{0,\text{para}}$ and $H_{0,\text{const}}$ defined at the beginning of Section 1(a). The statistical strategy to construct such tests is laid out in the work programme in Section 2.3.

Compared to existing test procedures, the proposed multiscale tests have the following main advantages: (i) By construction, they do not depend on a specific bandwidth parameter h but take into account multiple scales or bandwidths h simultaneously. They thus do not require a bandwidth choice. (ii) They are much more informative than non-multiscale tests. To make this point more precise, consider the hypothesis $H_0 : m_1 = \ldots = m_n$. Practitioners are not only interested in whether H_0 is violated. They would also like to know which violations occur, in particular, which time trends are different and during which time periods they differ from each other. As explained in more detail in Section 2.3, our multiscale test is designed to convey this additional information. In particular, it does not only allow to test whether the overall null hypothesis H_0 is violated. It also allows to make rigorous statistical confidence statements about which time series have a different trend and over which time periods these trends differ.

To the best of our knowledge, the only multiscale method available to test H_0 in (a strongly simplified version of) model (5) is the SiZer method of Park et al. (2009). However, as already discussed in Section 1, their analysis is mainly methodological and not backed up by a general theory. In particular, theory is only available for the special case of $n = 2$ time series. Moreover, the theoretical results are only valid under very severe restrictions on the set of bandwidths $\mathcal{H}$ that is taken into account by the multiscale method. In particular, the bandwidths h in the set $\mathcal{H}$ are assumed to be bounded away from zero. Put differently, they are not allowed to converge to zero as the sample size grows, which is obviously a very severe limitation. In contrast to this, we attempt to derive a complete asymptotic theory for our multiscale tests which is valid under general conditions.

(b) Contributions to curve clustering

The second main objective of the project is to develop a multiscale clustering method in the framework of model (5) which allows to estimate the unknown groups $G_1, \ldots, G_{K_0}$ and their unknown number K_0 defined in (4). A brief summary of how to construct such a method is provided in the work programme in Section 2.3. To the best of our knowledge, the only multiscale clustering method available in the literature is due to Vogt and Linton (2018). There, however, only basic consistency results were derived for the estimators of the unknown group structure. We aim to go beyond this basic theory and establish more advanced theory for our estimators of the unknown groups $G_1, \ldots, G_{K_0}$ and their unknown number K_0. A more detailed discussion of the theoretical statements we aim to establish is given in the work programme in Section 2.3.

2.3 Work programme incl. proposed research methods

In what follows, we provide a detailed account of how we intend to address the challenges formulated in Section 2.2.

(a) Methods for statistical multiscale testing

We first describe our strategy to construct multiscale tests of the hypotheses H_0, $H_{0,\text{para}}$ and $H_{0,\text{const}}$ in model (5). To keep the exposition focused, we restrict attention to the hypothesis $H_0 : m_1 = \ldots = m_n$. For any interval $[u-h, u+h] \subseteq [0,1]$, consider the hypothesis

$$H_0^{[i,j]}(u,h) : m_i(w) = m_j(w) \text{ for all } w \in [u-h, u+h].$$

Obviously, H_0 can be reformulated as

$$\begin{aligned} H_0 : &\text{ The hypothesis } H_0^{[i,j]}(u,h) \text{ holds true for all intervals } [u-h, u+h] \subseteq [0,1] \\ &\text{ and for all } 1 \leq i < j \leq n. \end{aligned}$$

We now set up a multiscale method which simultaneously tests the hypothesis $H_0^{[i,j]}(u,h)$ for all possible points (u,h) and all pairs (i,j) with $i < j$.[1] Our strategy to derive such a method can be outlined as follows:

Step 1: Construction of the test statistic.

(i) Construct nonparametric kernel estimators $\hat{m}_{i,h}$ of the trend functions m_i, where h denotes the bandwidth parameter.

(ii) For each given (u,h) and (i,j), construct a test statistic $\hat{S}_{ij}(u,h)$ of the hypothesis $H_0^{[i,j]}(u,h)$. A simple choice is a statistic of the form $\hat{S}_{ij}(u,h) = \sqrt{Th}(\hat{m}_{i,h}(u) - \hat{m}_{j,h}(u))/\hat{\nu}_{ij,h}(u)$, where the term $\hat{\nu}_{ij,h}(u)$ is chosen to normalize the asymptotic variance of the statistic to 1.

(iii) Aggregate the statistics $\hat{S}_{ij}(u,h)$ for all possible (u,h) and (i,j) into a multiscale statistic. In order to do so, we will use the aggregation scheme proposed by Dümbgen and Spokoiny (2001). The resulting multiscale statistic has the form

$$\hat{\Psi}_{n,T} = \max_{1 \leq i < j \leq n} \sup_{u,h} \big\{ |\hat{S}_{ij}(u,h)| - \lambda(h) \big\},$$

where $\lambda(h)$ are (appropriately chosen) additive correction terms. As one can see, the multiscale statistic $\hat{\Psi}_{n,T}$ is not obtained by simply taking the supremum of the individual statistics $\hat{S}_{ij}(u,h)$. We rather take the supremum of the additively

[1]Obviously, in practice, we cannot consider all points $u \in (0,1)$ and all $h > 0$ but have to restrict attention to a finite subset of points. We ignore this in our presentation for simplicity.

corrected statistics $|\hat{S}_{ij}(u,h)| - \lambda(h)$ as first suggested in Dümbgen and Spokoiny (2001).

Step 2: Construction of the test procedure.

(i) Suppose for a moment we could compute the $(1-\alpha)$-quantile $q^*_{n,T}(\alpha)$ of the multiscale statistic $\hat{\Psi}_{n,T}$ under the null H_0. Then our multiscale test would be carried out as follows:

(T*) Reject the overall null hypothesis H_0 if $\hat{\Psi}_{n,T} > q^*_{n,T}(\alpha)$.

By construction, the decision rule (T*) is a rigorous level-α-test, which means that $\mathbb{P}(\hat{\Psi}_{n,T} > q^*_{n,T}(\alpha)) = 1 - \alpha$ under H_0.

(ii) The quantile $q^*_{n,T}(\alpha)$ is a highly complicated quantity which is not known in practice. Hence, it cannot be used to set up the test. We thus need to come up with an (asymptotic) approximation $q_{n,T}(\alpha)$ which is computable in practice. In particular, we require $q_{n,T}(\alpha)$ to be an approximation of the quantile $q^*_{n,T}(\alpha)$ in the sense that

$$\mathbb{P}(\hat{\Psi}_{n,T} > q_{n,T}(\alpha)) = (1-\alpha) + o(1). \tag{6}$$

One of the main theoretical challenges of the project is to construct a suitable approximation $q_{n,T}(\alpha)$ and to verify that it has the property (6). Once we have succeeded in doing so, the multiscale test can be carried out as follows:

(T) Reject the overall null hypothesis H_0 if $\hat{\Psi}_{n,T} > q_{n,T}(\alpha)$.

(iii) By using the decision rule (T), we regard the constructed multiscale method as a test of the overall hypothesis H_0. Alternatively, one may view it as a simultaneous test of the family of hypotheses $H_0^{[i,j]}(u,h)$ for all points (u,h) and pairs (i,j). Looking at the method this way, one may proceed as follows:

(T$_{\text{mult}}$) For each interval $[u-h, u+h]$, reject the hypothesis $H_0^{[i,j]}(u,h)$ if the corrected test statistic $|\hat{S}_{ij}(u,h)| - \lambda(h)$ is above the critical value $q_{n,T}(\alpha)$, that is, if $|\hat{S}_{ij}(u,h)| - \lambda(h) > q_{n,T}(\alpha)$.

We conjecture that it is possible to prove the following theoretical result on the multiple testing procedure (T_{mult}) under appropriate regularity conditions:

With asymptotic probability $\geq 1-\alpha$, the hypothesis $H_0^{[i,j]}(u,h)$ is violated for all pairs (i,j) and for all intervals $[u-h,u+h]$ for which $|\hat{S}_{ij}(u,h)| - \lambda(h) > 0$.

According to this conjecture, we can make the following simultaneous confidence statement: We can claim, with (asymptotic) confidence at least $1-\alpha$, that the hypothesis $H_0^{[i,j]}(u,h)$ is violated for all pairs of time series (i,j) and for all intervals $[u-h,u+h]$ for which our test rejects. Hence, the multiscale test does not only give us information on whether the overall null hypothesis H_0 is violated. It also allows us to make rigorous statistical confidence statements about (i) which pairs of time series (i,j) have different trends and (ii) in which time regions $[u-h, u+h]$ these trends differ. This is valuable information in many applications.

In order to derive the theory for the multiscale test outlined above, we will build on the theoretical results developed in Khismatullina and Vogt (2018). Under very restrictive assumptions (such as $n = 2$ and no cross-sectional dependence at all in the data) the results should carry over although the technical details need to be worked out. However, under more realistic assumptions (such as a general number of curves n and cross-sectional dependence in the data), there is considerable work to be done and the results from Khismatullina and Vogt (2018) do not carry over in a straightforward way.

(b) Multiscale clustering methods

In order to construct a multiscale clustering algorithm in model (5), we will adapt the clustering approach of Vogt and Linton (2018). In particular, we plan to proceed as follows: (i) We use the multiscale test statistics from the first part of the project to construct distance measures between pairs of trends m_i and m_j. (ii) From these distance measures, we obtain so-called dissimilarity measures which form the basis of a hierarchical clustering algorithm.

Vogt and Linton (2018) derived some basic consistency results for their estimators of the unknown groups $G_1, \ldots, G_{K_0}$ and their unknown number K_0. Letting $\hat{K}_0$ be the estimator of K_0 and $\{\hat{G}_1, \ldots, \hat{G}_{\hat{K}_0}\}$ the estimator of the group structure $\{G_1, \ldots, G_{K_0}\}$, they in particular show that under appropriate regularity conditions,

$$\mathbb{P}\big(\hat{K}_0 = K_0\big) \to 1 \quad \text{and} \quad \mathbb{P}\Big(\{\hat{G}_1, \ldots, \hat{G}_{\hat{K}_0}\} = \{G_1, \ldots, G_{K_0}\}\Big) \to 1 \tag{7}$$

as the sample size goes to infinity. We intend to go beyond these basic consistency results and derive more advanced theory for our estimators. To be more specific, consider the problem of estimating the unknown number of groups K_0. The estimator of K_0 in Vogt and Linton (2018) depends on a tuning parameter $\pi_{n,T}$. Only a heuristic rule is available for the choice of this parameter. In contrast to this, we plan to derive rigorous theory for the choice of this tuning parameter. In particular, we intend to choose $\pi_{n,T} = q_{n,T}(\alpha)$, where $q_{n,T}(\alpha)$ is the approximate $(1-\alpha)$-quantile of the multiscale statistic defined above. With this choice, our estimator $\hat{K}_0$ of K_0 implicitly depends on the significance level α, that is, $\hat{K}_0 = \hat{K}_0(\alpha)$. For a given α, we conjecture that it is possible to prove that

$$\mathbb{P}(\hat{K}_0 = K_0) \geq (1-\alpha) + O(r_{n,T}), \tag{8}$$

where $r_{n,T}$ is the rate of the lower order terms. This statement can be interpreted as follows: For given α, $\hat{K}_0$ is equal to the true number of groups K_0 with (asymptotic) probability at least $1 - \alpha$. Hence, we can tune the clustering algorithm in such a way that the probability of misestimating the number of groups K_0 is (asymptotically) controlled. (8) can thus be understood as an asymptotic confidence statement about the estimator $\hat{K}_0$. In addition, we aim to derive an analogous confidence statement for the estimators $\hat{G}_1, \ldots, \hat{G}_{\hat{K}_0}$ of the unknown groups.

Empirical applications

Our test and clustering methods have a wide range of potential applications in economics and finance. Among other things, they can be used to compare the volatility trends of different stocks (Nyblom and Harvey, 2000), short-term risk-free interest rates (Fan and Yao, 2003; Park et al., 2009) or long-term rates across countries (Park et al., 2009). Another potential application is concerned with economic growth, which has been a key topic in macroeconomics for many decades. Economists are very much interested in the question whether gross domestic product (GDP) growth has been faster in some countries than in others. A suitable econometric framework to investigate this question is the panel data model (5). Zhang et al. (2012) used a special case of this model to analyze data from 16 OECD countries. For each of the $n = 16$ countries, quarterly time series data on gross domestic product (GDP), capital stock (K), labour input (L) and human capital (H) were available. The data were assumed to follow the model

$$\Delta \log GDP_{it} = m_i\Big(\frac{t}{T}\Big) + \beta_1 \Delta \log L_{it} + \beta_2 \Delta \log K_{it} + \beta_3 \Delta \log H_{it} + \alpha_i + \varepsilon_{it}$$

with $1 \le t \le T = 140$ and $1 \le i \le n = 16$, where m_i is the time trend of country i, β_k are unknown regression coefficients and $\Delta \log Z_{it} = \log Z_{it} - \log Z_{it-1}$ for $Z_{it} = GDP_{it}$, L_{it}, K_{it}, H_{it}. Zhang et al. (2012) tested the widely used common trends hypothesis $H_0 : m_1 = \ldots = m_n$ in this framework. Their analysis provided evidence against H_0. Specifically, their test rejected H_0 at the 10% confidence level. However, even if the common trends hypothesis is violated, there may still be groups of countries with the same time trend. It may thus be interesting to cluster the OECD countries into groups. We intend to use the multiscale methods developed in the project to produce such a clustering and, more generally, to analyze an updated version of the data sample from Zhang et al. (2012).

Another application we would like to explore deals with the analysis of temperature data, which has attracted some attention in econometrics in recent years; see for example Gao and Hawthorne (2006), Atak et al. (2011) and Davidson et al. (2016). Over the last decades, large panel data sets have become available which contain long temperature time series $\mathcal{Z}_i = \{Y_{it} : 1 \le t \le T\}$ for a huge number of different spatial locations i; see the Berkeley Earth project at `http://berkeleyearth.org` for examples of such big data sets. A simple trend model for the time series $\mathcal{Z}_i$ is given by the equation

$$Y_{it} = m_i\Big(\frac{t}{T}\Big) + \alpha_i + \varepsilon_{it}$$

with $\mathbb{E}[\varepsilon_{it}] = 0$, where m_i is the temperature trend at location i. If covariates are available, we could also work with the more general model (5). Climatologists are very much interested in analyzing the trending behaviour of temperature time series. Information on the trending behaviour is needed to better understand long-term climate variability. Among other things, they would like to know whether the time trends m_i are the same across locations or whether they can be clustered into groups. We aim to investigate these questions by the test and clustering methods developed in the project.

Summary

The work programme consists of two main work packages: Package (a) is devoted to the multiscale test methods described in part (a) of Section 2.2. Package (b) will focus on the multiscale clustering methods described in part (b) of Section 2.2. In each of the two work packages, we will proceed as follows: We will first develop the statistical methodology, then derive the theoretical properties thereof, evaluate the finite sample performance by simulations, and finally illustrate the methods by the application examples discussed above.

Work package	2019	2020	2021
(a) Multiscale tests	Oct–Dec	Jan–Dec	
(b) Multiscale clustering			Jan–Oct

The results of work package (a) will be presented at conferences in the second half of 2020. Possible venues are the World Congress of the Econometric Society (Aug 2020) and the CMStatistics Conference (Dec 2020). The results of package (b) will be presented at conferences in summer/autumn 2021. Possible venues are the European Summer Meeting of the Econometric Society and the European Meeting of Statisticians.

3 Bibliography

ABRAHAM, C., CORNILLON, P. A., MATZNER-LØBER, E. and MOLINARI, N. (2003). Unsupervised curve clustering using B-splines. *Scandinavian Journal of Statistics*, **30** 581–595.

ANDERSON, N. H., HALL, P. and TITTERINGTON, D. M. (1994). Two-sample test statistics for measuring discrepancies between two multivariate probability density functions using kernel-based density estimates. *Journal of Multivariate Analysis*, **50** 41–54.

ANDERSON, T. W. (1962). On the distribution of the two-sample Cramér-von Mises criterion. *Annals of Mathematical Statistics*, **33** 1148–1159.

ATAK, A., LINTON, O. and XIAO, Z. (2011). A semiparametric panel model for unbalanced data with application to climate change in the United Kingdom. *Journal of Econometrics*, **164** 92–115.

BONEVA, L., LINTON, O. and VOGT, M. (2015). A semiparametric model for heterogeneous panel data with fixed effects. *Journal of Econometrics*, **188** 327–345.

BONEVA, L., LINTON, O. and VOGT, M. (2016). The effect of fragmentation in trading on market quality in the UK equity market. *Journal of Applied Econometrics*, **31** 192–213.

BONHOMME, S. and MANRESA, E. (2015). Grouped patterns of heterogeneity in panel data. *Econometrica*, **83** 1147–1184.

CAI, Z. (2007). Trending time-varying coefficients time series models with serially correlated errors. *Journal of Econometrics*, **136** 163–188.

CHAUDHURI, P. and MARRON, J. S. (1999). SiZer for the exploration of structures in curves. *Journal of the American Statistical Association*, **94** 807–823.

CHAUDHURI, P. and MARRON, J. S. (2000). Scale space view of curve estimation. *Annals of Statistics*, **28** 408–428.

CHEN, J., GAO, J. and LI, D. (2012). Semiparametric trending panel data models with cross-sectional dependence. *Journal of Econometrics*, **171** 71–85.

CHEN, L. and WU, W. B. (2018). Testing for trends in high-dimensional time series. *Forthcoming in Journal of the American Statistical Association.*

CHIOU, J.-M. and LI, P.-L. (2007). Functional clustering and identifying substructures of longitudinal data. *Journal of the Royal Statistical Society: Series B*, **69** 679–699.

DAHLHAUS, R. (1997). Fitting time series models to nonstationary processes. *The Annals of Statistics*, **25** 1–37.

DAVIDSON, J., STEPHENSON, D. and TURASIE, A. (2016). Time series modeling of paleoclimate data. *Environmetrics*, **27** 55–65.

DEGRAS, D., XU, Z., ZHANG, T. and WU, W. B. (2012). Testing for parallelism among trends in multiple time series. *IEEE Transactions on Signal Processing*, **60** 1087–1097.

Delgado, M. A. (1993). Testing the equality of nonparametric regression curves. *Statistics & Probability Letters*, **17** 199–204.

Dümbgen, L. (2002). Application of local rank tests to nonparametric regression. *Journal of Nonparametric Statistics*, **14** 511–537.

Dümbgen, L. and Spokoiny, V. G. (2001). Multiscale testing of qualitative hypotheses. *Annals of Statistics*, **29** 124–152.

Dümbgen, L. and Walther, G. (2008). Multiscale inference about a density. *Annals of Statistics*, **36** 1758–1785.

Eckle, K., Bissantz, N. and Dette, H. (2017). Multiscale inference for multivariate deconvolution. *Electronic Journal of Statistics*, **11** 4179–4219.

Fan, J. and Yao, Q. (2003). *Nonlinear time series: nonparametric and parametric methods*. Springer.

Finner, H. and Gontscharuk, V. (2018). Two-sample Kolmogorov-Smirnov-type tests revisited: old and new tests in terms of local levels. *Annals of Statistics*, **46** 3014–3037.

Gao, J. and Gijbels, I. (2008). Bandwidth selection in nonparametric kernel testing. *Journal of the American Statistical Association*, **103** 1584–1594.

Gao, J. and Hawthorne, K. (2006). Semiparametric estimation and testing of the trend of temperature series. *The Econometrics Journal*, **9** 332–355.

Grier, K. B. and Tullock, G. (1989). An empirical analysis of cross-national economic growth, 1951–1980. *Journal of Monetary Economics*, **24** 259–276.

Hall, P. and Hart, J. D. (1990). Bootstrap test for difference between means in nonparametric regression. *Journal of the American Statistical Association*, **85** 1039–1049.

Hannig, J. and Marron, J. S. (2006). Advanced distribution theory for SiZer. *Journal of the American Statistical Association*, **101** 484–499.

Härdle, W. and Marron, J. S. (1990). Semiparametric comparison of regression curves. *Annals of Statistics*, **18** 63–89.

Hidalgo, J. and Lee, J. (2014). A CUSUM test for common trends in large heterogeneous panels. In *Essays in Honor of Peter C. B. Phillips*. Emerald Group Publishing Limited, 303–345.

Horowitz, J. L. and Spokoiny, V. G. (2001). An adaptive, rate-optimal test of a parametric mean-regression model against a nonparametric alternative. *Econometrica*, **69** 599–631.

Jacques, J. and Preda, C. (2014). Functional data clustering: a survey. *Advances in Data Analysis and Classification*, **8** 231–255.

James, G. M. and Sugar, C. A. (2003). Clustering for sparsely sampled functional data. *Journal of the American Statistical Association*, **98** 397–408.

Karoly, D. J. and Wu, Q. (2005). Detection of regional surface temperature trends. *Journal of Climate*, **18** 4337–4343.

Khismatullina, M. and Vogt, M. (2018). Multiscale inference and long-run variance estimation in nonparametric regression with time series errors. *Preprint*.

Kiefer, J. (1959). K-sample analogues of the Kolmogorov-Smirnov and Cramér-v. Mises tests. *Annals of Mathematical Statistics*, **30** 420–447.

Kneip, A., Sickles, R. C. and Song, W. (2012). A new panel data treatment for heterogeneity in time trends. *Econometric Theory*, **28** 590–628.

Li, D., Chen, J. and Gao, J. (2010). Nonparametric time-varying coefficient panel data models with fixed effects. *The Econometrics Journal*, **14** 387–408.

Li, Q., Maasoumi, E. and Racine, J. S. (2009). A nonparametric test for equality of distributions with mixed categorical and continuous data. *Journal of Econometrics*, **148** 186–200.

Lyubchich, V. and Gel, Y. R. (2016). A local factor nonparametric test for trend synchronism in multiple time series. *Journal of Multivariate Analysis*, **150** 91–104.

Mammen, E. (1992). *When does bootstrap work? Asymptotic results and simulations.* New York, Springer.

Nyblom, J. and Harvey, A. (2000). Tests of common stochastic trends. *Econometric Theory*, **16** 176–199.

Park, C., Vaughan, A., Hannig, J. and Kang, K.-H. (2009). SiZer analysis for the comparison of time series. *Journal of Statistical Planning and Inference*, **139** 3974–3988.

Proksch, K., Werner, F. and Munk, A. (2018). Multiscale scanning in inverse problems. *Forthcoming in Annals of Statistics.*

Ray, S. and Mallick, B. (2006). Functional clustering by Bayesian wavelet methods. *Journal of the Royal Statistical Society: Series B*, **68** 305–332.

Robinson, P. M. (1989). Nonparametric estimation of time-varying parameters. In *Statistical Analysis and Forecasting of Economic Structural Change*. Springer, 253–264.

Robinson, P. M. (2012). Nonparametric trending regression with cross-sectional dependence. *Journal of Econometrics*, **169** 4–14.

Rohde, A. (2008). Adaptive goodness-of-fit tests based on signed ranks. *Annals of Statistics*, **36** 1346–1374.

Rondonotti, V., Marron, J. S. and Park, C. (2007). SiZer for time series: a new approach to the analysis of trends. *Electronic Journal of Statistics*, **1** 268–289.

Rufibach, K. and Walther, G. (2010). The block criterion for multiscale inference about a density, with applications to other multiscale problems. *Journal of Computational and Graphical Statistics*, **19** 175–190.

Schmidt-Hieber, J., Munk, A. and Dümbgen, L. (2013). Multiscale methods for shape constraints in deconvolution: confidence statements for qualitative features. *Annals of Statistics*, **41** 1299–1328.

Stock, J. H. and Watson, M. W. (1988). Testing for common trends. *Journal of the American Statistical Association*, **83** 1097–1107.

Su, L. and Ju, G. (2018). Identifying latent grouped patterns in panel data models with interactive fixed effects. *Journal of Econometrics*, **206** 554–573.

Su, L., Shi, Z. and Phillips, P. C. (2016). Identifying latent structures in panel data. *Econometrica*, **84** 2215–2264.

Sun, Y. (2011). Robust trend inference with series variance estimator and testing-optimal smoothing parameter. *Journal of Econometrics*, **164** 345–366.

Tarpey, T. (2007). Linear transformations and the k-means clustering algorithm. *The American Statistician*, **61** 34–40.

Tarpey, T. and Kinateder, K. K. (2003). Clustering functional data. *Journal of Classification*, **20** 093–114.

Vogelsang, T. J. and Franses, P. H. (2005). Testing for common deterministic trend slopes. *Journal of Econometrics*, **126** 1–24.

Vogt, M. and Linton, O. (2014). Nonparametric estimation of a periodic sequence in the presence of a smooth trend. *Biometrika*, **101** 121–140.

Vogt, M. and Linton, O. (2017). Classification of non-parametric regression functions in longitudinal data models. *Journal of the Royal Statistical Society: Series B*, **79** 5–27.

Vogt, M. and Linton, O. (2018). Multiscale clustering of nonparametric regression curves. *Preprint*.

Wang, W., Phillips, P. C. and Su, L. (2018). Homogeneity pursuit in panel data models: theory and application. *Journal of Applied Econometrics*, **33** 797–815.

Wu, W. B. (2005). Nonlinear system theory: another look at dependence. *Proc. Natn. Acad. Sci. USA*, **102** 14150–14154.

Xu, K.-L. (2012). Robustifying multivariate trend tests to nonstationary volatility. *Journal of Econometrics*, **169** 147–154.

Zhang, T. (2013). Clustering high-dimensional time series based on parallelism. *Journal of the American Statistical Association*, **108** 577–588.

Zhang, Y., Su, L. and Phillips, P. C. (2012). Testing for common trends in semi-parametric panel data models with fixed effects. *The Econometrics Journal*, **15** 56–100.

4 Requested modules/funds

4.1 Basic Module

4.1.1 Funding for Staff

No.	Position	Funds
1.	Doctoral researcher (Bes.Gr. E13 Stufe 2 / E14 Stufe 1, 75%)	49725 € per calender year
2.	Student Assistant (10 hours per week for half a year)	2400 €

Job description for requested staff:

1. As research staff, a doctoral student is required who already possesses a thorough expertise in statistical multiscale methods. Marina Khismatullina who is a member of the Bonn Graduate School of Economics fits these requirements very well. Ms. Khismatullina and the applicant have already collaborated on a project which is concerned with statistical multiscale techniques; cp. Khismatullina and Vogt (2018). Ms. Khismatullina would thus bring in the expertise needed for the project. Moreover, with her very advanced programming skills, she will be able to develop the computational software required for simulations and empirical applications.

2. At the onset of the project, a student assistant position should be available for support with data collection, exploratory data analysis and auxiliary programming. Specifically, the data for the empirical applications need to be collected from various sources, cleaned and brought into a form which allows to process them. The prerequisites are strong analytical and programming skills.

4.1.2 Direct Project Costs

4.1.2.1 Travel Expenses

	2019	2020	2021
International conferences	0 €	1000 €	1000 €

The doctoral researcher is supposed to present the project at international conferences and workshops. Possible venues are the meetings of the Econometric Society, the European Meeting of Statisticians and the CMStatistics conference. In order to cover the travel expenses of the doctoral researcher, the above funds are requested.

4.1.2.2 Project-related publication expenses

	2019	2020	2021
Journal submission fees	0 €	200 €	200 €

5 Project requirements

5.1 Employment status information

Vogt, Michael, Professor, tenured position

5.2 First-time proposal data

Vogt, Michael

5.3 Composition of the project group

Vogt, Michael, Professor, tenured position

5.4 Cooperation with other researchers

5.4.1 Researchers with whom you have collaborated scientifically within the past three years

Holger Dette – Ruhr University Bochum, Germany
Oliver Linton – University of Cambridge, UK
María Dolores Martínez-Miranda – University of Granada, Spain
Enno Mammen – University of Heidelberg, Germany
Jens Perch Nielsen – CASS Business School, London, UK
Matthias Schmid – University of Bonn, Germany
Christopher Walsh – Technical University of Dortmund, Germany

5.5 Scientific equipment

The University of Bonn has a sufficient infrastructure in hard- and software to carry out the project. Personal computers are available and can be used within the project. Equipment like printers and copiers can be used as well.

6 Additional information

A request for funding this project has not been submitted to any other addresses. In the case that such a request will be submitted elsewhere, the DFG will be informed immediately. The DFG liaison officer Katrin Hahlen of the University of Bonn (`hahlen@verwaltung.uni-bonn.de`) has been informed about this application.