

Editorial to the special issue on Applicable semiparametrics of computational statistics

Ostap Okhrin¹ · Stefan Trück²

Published online: 5 September 2015
© Springer-Verlag Berlin Heidelberg 2015

1 Introduction

Over the last 50 years, semiparametric and nonparametric modeling has become extremely popular in different areas of research. Since the first main methodological breakthrough in kernel density estimation made by [Rosenblatt \(1956\)](#) and [Parzen \(1962\)](#), the use of these techniques has increased exponentially. Nowadays, not only methodological, but also a wide range of empirical and applied work using non- and semi-parametric methods, is conducted in many quantitatively oriented sciences like risk management, biostatistics, hydrology, and finance. The main reason for the popularity of the techniques originates in their flexibility and adaptability. Many important contributions to this area of research, both from a methodological and an applied perspective, have been made by Wolfgang Karl Härdle. To highlight this, a conference entitled “Applicable Semiparametrics” was organized in October 2013 at Humboldt-

Ostap Okhrin acknowledges support from Deutsche Forschungsgemeinschaft through the CRC 649 “Economic Risk”, at Humboldt-Universität zu Berlin, Stefan Trück acknowledges support from the Centre for International Finance and Regulation (CIFR) through the Research Grant “Early-Warning Systems and Managing Systemic Risks using Real-Time Financial and Business Conditions Indicators”.

✉ Ostap Okhrin
ostap.okhrin@tu-dresden.de

✉ Stefan Trück
stefan.trueck@mq.edu.au

¹ Chair of Econometrics and Statistics esp. Transportation, Institute of Economics and Transport, Faculty of Transportation, Dresden University of Technology, Helmholtzstraße 10, 01069 Dresden, Germany

² Centre for Financial Risk, Faculty of Business and Economics, Macquarie University, Building E4A, North Ryde, NSW 2109, Australia

University Berlin. This special issue collects some of the outstanding work related to this conference.

This editorial is organized as follows: in Sect. 2, we briefly postulate main directions of applied non- and semi-parametric modeling and also provide a review of Wolfgang Karl Härdle's contributions to the area. Section 3 presents the content of this special issue.

2 Applicable semiparametrics

Nonparametric techniques refer to methods without assumptions on the underlying model. Nonparametric statistics are obtained purely from the data and do not depend on any parameters. In time series applications the model structure frequently changes over time, which might lead to different parameter values or parameters itself. Nonparametric techniques are suitable to find structural changes in the data. However, in many applications, the general relationship between the variables considered is predetermined, therefore some parametric form can still be assumed. This leads to semiparametric modeling where one part of the model is assumed to have a parametric structure, while other parts are assumed as nonparametric.

Up to recent decades, work on non- and semi-parametric modeling was mainly focused on methodological and theoretical papers, nevertheless its main purpose is driven by the data. As the field is very broad, it is impossible to provide even a brief overview of the work in this area. However, over the last decades, Wolfgang Karl Härdle has made many seminal contributions to applicable semiparametrics, both from theoretical and applied perspectives. These contributions range from robust nonparametric regression (Härdle 1984a, b; Härdle and Marron 1985a, b, 1990a, b; Härdle and Gasser 1985; Härdle and Kelly 1987; Härdle et al. 1988, 1992a, b; Härdle and Bowman 1988; Härdle and Tsybakov 1988; Härdle and Vieu 1992; Härdle and Mammen 1993), over single index models (Härdle et al. 1993, 1997; Horowitz and Härdle 1996), real steps into computational statistics (Müller et al. 1997; Klinke et al. 1997; Härdle et al. 2001), true applicable semiparametrics (Liang et al. 1999; Hall et al. 2000; Wang et al. 2004) to the development of dynamic semiparametric factor models (Fengler et al. 2007; Brüggemann et al. 2008; Park et al. 2009; Härdle and Hlavka 2009). The conducted work has not only a strong theoretical background, but also solid empirical justification, which makes the methods developed by Wolfgang Karl Härdle highly valuable. Worth mentioning is also his series of books, for example, Härdle et al. (2004) or Sperlich et al. (2006), some of which have become standard text books in academia.

The conference on *Applicable Semiparametrics* in October 2013 brought together researchers working on a huge variety of topics related to semiparametric theory as well as the application of semiparametric methods to economics, finance, and other fields. Taking into account Wolfgang Karl Härdle's influence on this field, the conference as well as this special issue have been dedicated to his 60th birthday.

3 Contents of this special issue

This special issue of Computational Statistics on *Applicable Semiparametrics* consists of twelve articles, some of them with a methodological focus, while others are clearly of more applied nature.

The issue starts with theoretical contributions into semiparametric and functional regression with changing coefficients. [Aneiros and Vieu \(2015\)](#) propose new types of partially linear models which allow for more than one functional covariate. The model class allows for both continuous and discrete type effects of functional variables. The proposed model is flexible in the nonparametric component which gives good predictive properties, the linearity allows for interpretation and the simplicity of the model allows for ease in practical applications. Using functional regression models with scalar response, in the next paper of this issue, [Goia and Vieu \(2015\)](#) develop a methodology for approximation of the unknown regression operator in a semiparametric manner. Hereby, they propose to use a single index approach that can also take into account potential structural changes.

The third paper of the issue by [Sperlich and Theler \(2015\)](#) investigates the question of how to deal with heterogeneity when studying causal effects. The authors illustrate the inadequacy of standard linear models that simply interpret the coefficients as average returns or effects and then show, why varying-coefficient models have great potential to model heterogeneity in various important problems and applications of regression. Staying in the area of time series models with random coefficients, [Liu et al. \(2015\)](#) propose a new class of nonparametric seasonal time series models under the framework of functional coefficients. Hereby, the authors develop a local linear approach for the estimation of the nonparametric trend and seasonal effect functions. The paper also includes an interesting application of the developed model using data on tourist visits to a major holiday destination. In-line with the weather-based application of the previous paper, [Cao et al. \(2015\)](#) model temperature in space and time using two different approaches: Kriging, which became a standard technique in spatial analysis, and dynamic semiparametric factor models (DSFM) which are nonparametric in space and parametric in time. The latter technique has shown a very impressive performance for modeling non-spatial data such as, for example, brain activity, implied correlation, implied volatility, or empirical pricing kernels. However, the conducted work is innovative as it conducts an extensive empirical study applying the proposed technique also to spatial data. Another application to spatial statistics is provided by [Kolbe et al. \(2015\)](#): using adaptive weights smoothing (AWS), the authors estimate a map of land values for Berlin based on a unique data set of undeveloped land prices transacted between 1996 and 2009. An iterative AWS piecewise constant regression is applied to deal with the challenging problem of transaction noise in the data.

The special issue also contains two papers dealing with the prediction of electricity prices. [Nowotarski and Weron \(2015\)](#) investigate interval forecast techniques that are based on combined individual electricity spot price forecasts and quantile regression averaging (QRA). The authors suggest that in particular QRA-based forecasting yield promising results that provide accuracy gains over predictions from individual models. [Maciejowska and Weron \(2015\)](#) apply factor models, utilizing intra-day and inter-zone relationships to predict short and long term electricity prices. They examine

restricted vector autoregressive models as well as three-factor models with common and idiosyncratic components that are estimated using principal component analysis in a semiparametric setup. The results illustrate that the use of multivariate models has the potential to significantly improve forecasts for electricity spot prices.

Another forecasting study, using support vector regression for non-linear autoregressive-moving-average (ARMA) models is proposed by [Chen et al. \(2015\)](#). The authors conduct an extensive empirical analysis, in which the performance of the developed model for predicting financial returns is examined against various benchmark models. The results suggest that the proposed model outperforms its competitors with regards to forecasting the magnitude and turning points in the considered return series. In the next paper, [Chen and Härdle \(2015\)](#) provide an in-depth examination of common factors determining systemic credit risk and the changes of credit default swap (CDS) spreads in North America and Europe. Using asymptotic principal components, they estimate common factors of CDS spreads during the pre-crisis, crisis, and post-crisis period and relate them to observable economic variables.

Semiparametric modeling is also vitally important for panel data. [Proença and Faustino \(2015\)](#) make a contribution to the class of fractional response models by introducing a nonparametric component to control for unobserved heterogeneity associated with the regressors. Their approach is based on the semi-mixed effects generalized linear model and examined in an empirical study on modeling bilateral intra-industry trade indexes between various established and emerging economies. Another application of semiparametric techniques to panel data is proposed by [Rodríguez-Poo and Soberón \(2015\)](#), using first-differences and within local linear regression estimators. Since the two considered approaches are equivalent asymptotically, the authors investigate small sample properties of the estimators in a Monte Carlo study.

Overall, all papers in this special issue are closely related to the application of semiparametric techniques in a wide range of areas. We present both theoretical work that investigates asymptotic properties as well as empirical studies ranging from spectrometry, over temperature to various financial markets. With no doubt, this shows the importance of most recent and future developments in applicable semiparametrics.

Acknowledgments We are very grateful to Yuichi Mori, the editor-in-chief of this journal, for proposing this special issue, and to all anonymous referees for their crucial help in the editorial process.

References

- Aneiros G, Vieu P (2015) Partial linear modelling with multifunctional covariates. *Comput Stat* 30. doi:[10.1007/s00180-015-0568-8](#)
- Brüggemann R, Härdle W, Mungo J, Trenkler C (2008) VAR modelling for dynamic semiparametric factors of volatility strings. *J Financ Econ* 6(3):361–381
- Cao X, Okhrin O, Odening M, Ritter M (2015) Modelling spatio-temporal variability of temperature. *Comput Stat* 30. doi:[10.1007/s00180-015-0561-2](#)
- Chen C, Härdle W (2015) Common factors in credit defaults swap markets. *Comput Stat* 30. doi:[10.1007/s00180-015-0578-6](#)
- Chen S, Jeong K, Härdle W (2015) Recurrent support vector regression for a non-linear ARMA model with applications to forecasting financial returns. *Comput Stat* 30. doi:[10.1007/s00180-014-0543-9](#)
- Fengler M, Härdle W, Mammen E (2007) A semiparametric factor model for implied volatility surface dynamics. *J Financ Econ* 5(2):189–218

- Goia A, Vieu P (2015) A partitioned single functional index model. *Comput Stat* 30. doi:[10.1007/s00180-014-0530-1](https://doi.org/10.1007/s00180-014-0530-1)
- Hall P, Härdle W, Kleinow T, Schmidt P (2000) On semiparametric bootstrap approach to hypothesis tests and confidence intervals for Hurst coefficients. *Stat Stoch Process* 3:263–276
- Härdle W (1984b) Robust regression function estimation. *J Multivar Anal* 14:169–180
- Härdle W (1984a) A law of the iterated logarithm for nonparametric regression function estimators. *Ann Stat* 12:624–635
- Härdle W, Hall P, Marron JS (1988) How far are automatically chosen regression smoothing parameters from their optimum? *J Am Stat Assoc* 83:86–101 (**with discussion**)
- Härdle W, Hall P, Marron JS (1992) Regression smoothing parameters that are not far from their optimum. *J Am Stat Assoc* 87:227–233
- Härdle W, Hart J, Marron JS, Tsybakov AB (1992) Bandwidth choice for average derivative estimation. *J Am Stat Assoc* 87:218–226
- Härdle W, Hall P, Ichimura H (1993) Optimal smoothing of single index models. *Ann Stat* 21:157–178
- Härdle W, Sperlich S, Spokoiny V (1997) Semiparametric single index versus fixed link function modelling. *Ann Stat* 25:212–243
- Härdle W, Kleinow T, Tschernig R (2001) Web quantlets for time series analysis. *Ann Inst Stat Math* 52:1–15
- Härdle W, Müller M, Sperlich S, Werwatz A (2004) *Nonparametric and semiparametric models*. Springer, Berlin
- Härdle W, Bowman A (1988) Bootstrapping in nonparametric regression: local adaptive smoothing and confidence bands. *J Am Stat Assoc* 83:102–110
- Härdle W, Gasser T (1985) On robust nonparametric estimation of derivatives of regression functions. *Scand J Stat* 12:233–240
- Härdle W, Hlavka Z (2009) Dynamics of state price densities. *J Econ* 150(1):1–15
- Härdle W, Kelly G (1987) Nonparametric kernel regression estimation—optimal choice of the bandwidth. *Statistics* 18:21–35
- Härdle W, Mammen E (1993) Comparing nonparametric versus parametric regression fits. *Ann Stat* 21:1926–1947
- Härdle W, Marron S (1985a) Bandwidth choice in nonparametric kernel regression. *Stat Decis* 2:173–177
- Härdle W, Marron S (1985b) Optimal bandwidth selection in nonparametric regression function estimation. *Ann Stat* 13:1465–1481
- Härdle W, Marron JS (1990a) Bootstrap simultaneous error bars for nonparametric regression. *Ann Stat* 19:778–796
- Härdle W, Marron JS (1990b) Semiparametric comparison of regression curves. *Ann Stat* 18:63–89
- Härdle W, Tsybakov AB (1988) Robust nonparametric regression with simultaneous scale curve estimation. *Ann Stat* 16:120–135
- Härdle W, Vieu P (1992) Kernel regression smoothing of time series. *J Time Ser Anal* 13:209–232
- Horowitz J, Härdle W (1996) Direct semiparametric estimation of single-index models with discrete covariates. *J Am Stat Assoc* 91:1632–1640
- Klinke S, Golubev Y, Härdle W, Neumann M (1997) Teaching wavelets in XploRe. *Comput Stat* 13:141–151
- Kolbe J, Schulz R, Wersing M, Werwatz A (2015) Identifying Berlin's land value map using adaptive weights smoothing. *Comput Stat* 30. doi:[10.1007/s00180-015-0559-9](https://doi.org/10.1007/s00180-015-0559-9)
- Liang H, Härdle W, Carroll R (1999) Large sample theory in a semiparametric partially linear errors-in-variables model. *Ann Stat* 27:1519–1535
- Liu X, Cai Z, Chen R (2015) Functional coefficient seasonal time series models with an application of Hawaii tourism data. *Comput Stat* 30. doi:[10.1007/s00180-015-0574-x](https://doi.org/10.1007/s00180-015-0574-x)
- Maciejowska K, Weron R (2015) Forecasting of daily electricity prices with factor models: utilizing intra-day and inter-zone relationships. *Comput Stat* 30. doi:[10.1007/s00180-014-0531-0](https://doi.org/10.1007/s00180-014-0531-0)
- Müller M, Rönz B, Härdle W (1997) Computer assisted semiparametric generalized linear models. *Comput Stat* 12:153–172
- Nowotarski J, Weron R (2015) Computing electricity spot price prediction intervals using quantile regression and forecast averaging. *Comput Stat* 30. doi:[10.1007/s00180-014-0523-0](https://doi.org/10.1007/s00180-014-0523-0)
- Park B, Mammen E, Härdle W, Borak S (2009) Dynamic semiparametric factor models. *J Am Stat Assoc* 104:284–298
- Parzen E (1962) On estimation of a probability density function and mode. *Ann Math Stat* 33(3):1065–1076

- Proença I, Faustino H (2015) Modelling bilateral intra-industry trade indexes with panel data: a semiparametric approach. *Comput Stat* 30. doi:[10.1007/s00180-015-0556-z](https://doi.org/10.1007/s00180-015-0556-z)
- Rodriguez-Poo J, Soberón A (2015) Differencing techniques in semi-parametric panel data varying coefficient models with fixed effects: a Monte Carlo study. *Comput Stat* 30. doi:[10.1007/s00180-014-0549-3](https://doi.org/10.1007/s00180-014-0549-3)
- Rosenblatt M (1956) Remarks on some nonparametric estimates of a density function. *Ann Math Stat* 27(3):832–837
- Sperlich S, Härdle W, Aydinli G (eds) (2006) *The art of semiparametrics*. Physica-Verlag, Heidelberg
- Sperlich S, Theler R (2015) Modeling heterogeneity: a praise for varying-coefficient models in causal analysis. *Comput Stat* 30. doi:[10.1007/s00180-015-0581-y](https://doi.org/10.1007/s00180-015-0581-y)
- Wang Q, Linton O, Härdle W (2004) Semiparametric regression analysis under imputation for missing response data. *J Am Stat Assoc* 99(466):334–342

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.