
Applications of modern regression techniques in empirical economics

DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES DER FAKULTÄT FÜR
AGRARWISSENSCHAFTEN
GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN



VORGELEGT VON

Alexander März

GEBOREN IN GIENGEN AN DER BRENZ

GÖTTINGEN
MAI, 2016

D 7

1. Referent: Prof. Dr. Oliver Mußhoff
 2. Korreferent: Prof. Dr. Stephan v. Cramon-Taubadel
 3. Korreferent: Prof. Dr. Martin Odening (extern)
- Tag der mündlichen Prüfung: 14.07.2016

SIT DOWN BEFORE A FACT AS A LITTLE CHILD, BE PREPARED TO GIVE UP
EVERY PRECONCEIVED NOTION. FOLLOW HUMBLY WHEREVER AND TO
WHATEVER ABYSSES NATURE LEADS, OR YOU SHALL LEARN NOTHING.

(Thomas H. Huxley 1860)

Contents

I	INTRODUCTION	1
II	LINEAR REGRESSION AND BEYOND	11
III	ANALYSING FARMLAND RENTAL RATES USING BAYESIAN GEOADDITIVE QUANTILE REGRESSION	12
IV	INTERGENERATIONAL SOCIAL MOBILITY IN THE UNITED STATES: A MULTIVARIATE ANALYSIS US- ING DISTRIBUTIONAL REGRESSION	13
V	EXPERIMENTELLE UNTERSUCHUNG DER OPTI- MALITÄT VON INVESTITIONSENTSCHEIDUNGEN	14
VI	SUMMARY AND CONCLUDING REMARKS	15
	List of publications	19
	Erklärung über den geleisteten Eigenanteil an der Arbeit	20
	Eidesstattliche Erklärungen	22

Introduction

'Essentially, all models are wrong, but some are useful.' (Box 1979)

With the advent of big data and the associated increase of information in a wide range of different fields, such as administrative, business or financial domains, analysing data has become increasingly challenging. Consequently, there is a growing need for flexible statistical methods that allow for a realistic modelling of real data problems. Within the class of statistical models, regression analysis has become one of the principal statistical techniques for investigating the relationship between a response (dependent) variable and the explanatory variables (covariates) of interest. While it seems reasonable that empirical applications keep pace with recent advances in statistical methodology, it comes as a surprise that more traditional methods, such as the ordinary least squares regression model, are still among the most commonly applied in day to day research (Krueger and Lewis-Beck 2008).

Aside from their many desirable properties and the provision of economic insights, it can only be speculated as to why modern regression tools are not used more often in empirical applications. One reason might be that, similarly to statisticians, many researchers in applied economics are 'mean lovers' (Friedman 2002). Therefore, the argument put forward by Friedman (2002) might also explain the preferences of applied economists for using regression models that relate the mean of the response variable to explanatory variables within a regression setting, while ignoring other properties of the data, such as the variance or shape. While standard

linear regression models have their justification for cases in which the conditional mean provides a parsimonious summary of the data and where this is useful for the research question at hand, the researcher runs the risk of drawing incorrect scientific conclusions if he assumes that the process under study behaves according to an oversimplistic model. In order to avoid such inconclusive results, this thesis presents several applications of modern regression techniques in empirical economics that allow relaxing some of the restrictive assumptions of more traditional models. All applications are motivated from an economic point of view and address several issues that are currently being discussed in the economics literature.

To motivate the methodology that is used for the empirical analyses in subsequent chapters, Chapter II of this thesis briefly introduces the reader to Generalised Linear Models (GLM), Generalised Additive Models (GAM), Generalised Additive Models for Location, Scale and Shape (GAMLSS) and Quantile Regression. In the first analysis presented in Chapter III titled '*Analysing farmland rental rates using Bayesian geoadditve Quantile Regression*', we make use of Quantile Regression which allows us to analyze points other than the central tendency, e.g., low, medium and high quantiles, as well as to uncover potential heterogeneities of the estimated effects across the conditional response distribution. In the analysis of Chapter IV titled '*Intergenerational social mobility in the United States: a multivariate analysis using distributional regression*', we contribute to a recent strand of literature in social science and economics that highlights the importance of analyzing all parameter of a parametric response distribution within a regression setting. Instead of treating them as nuisance, we allow each parameter of a multivariate response distribution to change as functions of covariates and provide additional insight into the data and its underlying mechanisms, which would go unnoticed if more traditional regression models were used. In order to relate the investment decisions of farmers to an optimal investment point in time, the analysis presented in Chapter V titled '*Experimentelle Untersuchung der Optimalität von Investitionsentscheidungen*' employs a model that takes the multinomial structure of the data into account. With respect to the predictor specifications, the models used for the analyses offer a maximum degree of flexibility since they allow us to include linear terms, non-linear terms, random effect terms and spatial effect terms. The inclusion of the latter term is especially useful for our analyses as it

enables us to explicitly take the spatial dimension of the data into account and to provide additional insight into its spatial structure. Furthermore, by modelling the relationship between the response and the covariates semi-parametrically, we address the issue of functional form misspecification in all of our analyses that may occur when linear regression models are used. The final Chapter VI summarizes the main findings of the analyses and presents concluding remarks.

The remainder of this chapter briefly motivates and summarizes the contributions of each analysis to the literature.

Analyzing farmland rental rates

Land is an indispensable factor for agricultural production. Therefore, the question as to what determines agricultural land prices is an important research topic in agricultural economics. Given that many farmers gain access to farmland through rental markets and that payments for rental rates constitute an important factor in farmers' expenses, studies that analyze the price formation of farmland rental rates have evolved as an equally important field of research and have increased substantially over the past years.

A systematic review of the relevant literature shows that hedonic pricing theory according to Rosen (1974) provides the theoretical underpinnings for many empirical studies. In order to analyse the relationship between farmland rental rates and its determinants and to account for spatial dependencies in the data, spatial lag and spatial error models have evolved as standard tools in hedonic pricing studies. Even though these models have become the most widely used technique, there are a number of issues attached to them. Nickerson and Zhang (2014) highlight the problem of functional form misspecification as one crucial issue when estimating hedonic models. Since there is oftentimes no existing economic theory that guides researchers in terms of correctly specifying the functional form of the relationship between farmland and its determinants, the most common choice made in empirical applications is to assume a linear relationship. However, Kostov (2009) finds that misspecifications with respect to the functional form may be responsible for spurious spatial dependencies when hedonic pricing models are used. Consequently, functional form misspecifications can adversely influence statistical

inference, as they may affect both the magnitude and the significance of the estimated effects (Nickerson and Zhang 2014). Besides its relevance for statistical inference, Margarian (2008) considers the topic as being equally important from an economic point of view. In particular, Margarian (2008) emphasises that more attention should be dedicated to the analysis of the functional form in applied regression analysis rather than only on quantifying the estimated effects resulting from an assumed linear relationship. Hence, these considerations suggest that the assumption of linearity in parametric hedonic pricing models should be relaxed in order to acknowledge the possible non-linearity of some of the covariate effects in the modelling process. In the analysis presented in Chapter III, we model the relationship between farmland rental rates and its determinants non-linearly by making use of semi-parametric regression models, as this model class allows for a broader class of functional relationships than parametric linear models.

Even though the analysis of farmland is an active field of research in the agricultural economics literature, it comes as a modest surprise that comparatively few studies go beyond the modelling of conditional means. Consequently, little is known about farmland and the way it is influenced by economic variables for points other than the central tendency. In order to fill this gap, we use Quantile Regression that allows us to separately identify the determinants of rental rates for each quantile and to analyse the relationship between covariates and the response at different points in the conditional distribution. Hence, Quantile Regression is of special importance for hedonic pricing studies, as interest usually lies in analysing expensive rental rates rather than average rental rates only. In addition to allowing us to go beyond the modelling of mean rental rates, Quantile Regression enables us to address another topic that has been widely neglected in the empirical literature on farmland: the issue of heterogeneity (Bekkerman et al. 2013). Although heterogeneity has many different causes, we consider expectations of economic agents with respect to future conditions as one possible source of such heterogeneity. Since rental rates are usually mid- to long-term, rental rate agreements are negotiated at different points in time. Consequently, the different sets of expectations with respect to the future profit potential of the farmland under which the agreements were originally formed lead to heterogeneity. In a sample of heterogeneous economic agents, we expect the formation of future expectations to differ across

farmers so that it is difficult, or even infeasible, to directly and explicitly control for heterogeneity in the estimated effects in conventional hedonic pricing models. Therefore, mean hedonic pricing models may not provide an accurate estimation of the covariate effects, as these models do not allow for identifying heterogeneity resulting from factors that are not directly observable. In our analysis, we argue that heterogeneity with respect to future expectations of economic agents can be uncovered by looking at different points in the conditional distribution and that Quantile Regression provides a means of empirically uncovering the distinct and unobservable sets of factors that affect the formation of farmland rental rates.

Intergenerational social mobility in the United States

Social mobility and the lack thereof is a topic that has equally attracted the interest of academics and policy makers alike. Over the last few decades, increasing income inequality has led to a renewed interest in understanding intergenerational social mobility, which is defined as the change in social position or hierarchy that occurs during a person's lifetime. Social mobility is relevant because it implies equality of opportunity, where everyone is given equal chances of improving his or her socio-economic position through individual effort, hard work and talent, irrespective of his or her birth circumstances. Not only is it important for economic prosperity, but it is also a political issue. Indeed, it is likely to affect social justice; a society in which people have the perception they can personally pull the strings that lead to an improvement of their economic and social situation has the potential to foster social cohesion, in contrast to a society with a rigid social hierarchy, in which individual opportunities are decided at birth (Aldridge 2001). The United States, considered the land of opportunity, is particularly suited for the analysis due to the common belief that everyone has a fair shot at making it from a humble beginning to great wealth. Besides the widespread interest in the analysis of social mobility, there are a number of open issues that are of particular interest and that require further analysis.

The first issue that we address in our analysis presented in Chapter IV is concerned with the fact that two commonly used measures of social mobility, absolute and relative mobility, are usually analyzed independently from one another. While

absolute mobility captures the idea of growing prosperity for all, where everyone participates in the proceeds of economic growth, relative mobility is concerned with equality of opportunity, with a fluid and meritocratic society providing everyone with equal chances of climbing the income ladder, independent of his or her family background (Bengali and Daly 2013). Even though the literature agrees that the state of social mobility within a society can only be comprehensively assessed through a combined analysis of measures of absolute and relative mobility, existing studies of social mobility have focused on either absolute or relative mobility. However, since many Americans think of the American Dream in both absolute and relative terms, it seems more adequate to analyze the chances of Americans to move beyond their parents in terms of higher incomes, while changing their relative position in the income distribution at the same time (Haskins et al. 2008). Moreover, since both types of social mobility matter and both measures can reinforce the other, it is essential from a policy-maker point of view to better understand the interdependencies between the two in order to balance the need to promote both types of mobility. Therefore, we relax the assumption of independence between measures of absolute and relative mobility and study their interdependence by relating the coefficient of correlation between the two measures to covariates within a regression setting. This allows us to identify the factors that govern the strength of their association and to analyze the pathways through which absolute and relative mobility are linked.

With respect to our second contribution, a review of the relevant literature reveals that empirical models in economic and sociological research analyzing social mobility and its determinants have generally been developed within a conditional mean regression setting. However, characteristics of social mobility other than its central tendency are also of interest. Consider the decision of people to remain in or to move to a specific district as an example. Their decision may not be driven only by the overall level of social mobility within a particular area, but also by the chances to participate in, or, similarly, by the risk of being excluded from upward social mobility. For instance, people might prefer living in areas where the fluctuation around the mean level of social mobility is comparatively low, so that the risk of being excluded from the opportunity of climbing the income ladder is low. Therefore, instead of treating the variation of social mobility as a nuisance, we

view the analysis of the conditional variance within a heteroscedastic regression setting as a potentially important source of additional information in the data. While covariate effects in conditional mean regressions account for deviations of the group sample means from the overall mean of the response (between-group differences), covariate effects in heteroscedastic regression models explain how the variability of the response around group means changes as a function of covariates within groups (within-group difference). Consequently, parallel to studying between-group differences within a conditional mean regression setting, the analysis of within-group heterogeneity modelled as conditional heteroscedasticity allows us to provide a more complete picture of social mobility.

Instead of analyzing social mobility and its determinants cross-sectionally, our third contribution allows us to take a within country perspective and to analyze social mobility and its determinants at the level of counties and commuting zones within the United States. More particularly, our research focus lies in analyzing the factors that influence measures of absolute and relative social mobility and on investigating how well these factors capture the spatial patterns of social mobility across the United States. Given that recent studies have found geography to be an important determinant when it comes to the analysis of social phenomena, the analysis of local spatial patterns is important as it allows for a better understanding of social mobility and to draw a more precise picture at a regional level.

From a methodological point of view, we add to the empirical literature in economics and social science that is centered on analyzing the entire conditional distribution. In particular, we make use of Bayesian multivariate distributional regression, which allows us to model not only conditional means and conditional variances, but also all distributional parameter of a multivariate response distribution within a unified regression setting.

Optimality of investment decisions

The net-present value (NPV) criterion is commonly used in order to evaluate the profitability of an investment. Accordingly, an investment should be carried out if the net-present value of the expected future cash-flow resulting from the investment is positive. However, the classic NPV criterion is inconclusive if an

investment is characterized by future cash flows that are uncertain, when the costs associated with the investment cannot easily be recovered if the decision-makers want to undo their investment and in cases in which the investment is not limited to a now-or-never-decision, but can rather be moved into the future. As an alternative, the real option approach (ROA) extends the classic NPV criterion by considering the uncertainty associated with the returns on investments, the irreversibility of the investment costs and the possibility of postponing the investment into the future.

Although the real options approach is superior to the classic NPV criterion from a theoretical perspective, investment decisions often deviate, in practice, from the optimal time to invest predicted by the ROA. In particular, experimental studies show that decision-makers tend to invest too early compared to the optimal point in time under the real options approach (see, e.g., [Rauchs and Willinger 1996](#), [Richards and Green 2003](#), [Yavaş and Sirmans 2005](#), [Hinrichs et al. 2008](#), [Oprea et al. 2009](#), [Sandri et al. 2010](#)). So far, however, little empirical evidence is available so that it remains unclear how the deviations between the optimal point in time to invest under the ROA and the ones observed in practice can be brought into conformity. Also, it remains unresolved whether personal characteristics of the decision-makers or experiment-related characteristics, such as the framing of the experiment or learning effects, help to explain the inconsistencies.

Considering the existing literature, the objective of the analysis presented in Chapter V is to fill some of the existing research gaps and to investigate the influence of both decision-maker specific and experiment-related variables on the timing of investment decisions. In order to analyze the investment decisions made in practice, we make use of a model that allows us to differentiate between investment decisions that are made too early or too late and to relate these decisions to the optimal point in time to invest predicted by the real options approach. In order to analyze how country-specific differences affect the quality of investment decisions, we make use of a comprehensive data set of German and Kazakh farmers. This cross-country comparison provides an interesting extension of the existing literature, as Germany is a western market economy, while Kazakhstan has long been exposed to the impact of a centrally planned economy in the former Soviet Union.

The analyses presented in the following focus on research questions that are currently being discussed in the relevant literature. The objective is to address each of the issues raised and to discuss them by means of suitable methods and models.

References

- Aldridge, S.: 2001, *Social mobility: A discussion paper*, Performance and Innovation Unit, the Cabinet Office.
- Bekkerman, A., Brester, G. W. and McDonald, T. J.: 2013, A Semiparametric Approach to Analyzing Differentiated Agricultural Products, *Journal of Agricultural and Applied Economics* **45**(1), 79–94.
- Bengali, L. and Daly, M.: 2013, U.S. economic mobility: the dream and the data, *Federal Reserve Bank of San Francisco Economic Letter* **06**, 1–5.
- Box, George E. P.: 1979, Robustness in the strategy of scientific model building, in R. Launer and G. Wilkinson (eds), *Robustness in statistics*, Academic Press, pp. 201–236.
- Friedman, Hershey H. and Friedman, Linda W. and Amoo, Taiwo: 2002, Using humor in the introductory statistics course, *Journal of Statistics Education* **10**(3), 234–248.
- Haskins, R., Isaacs, J. and Sawhill, I.: 2008, Getting ahead or losing ground: Economic mobility in America, *Washington: The Brookings Institution*.
- Hinrichs, J., Musshoff, O. and Odening, M.: 2008, Econometric Hysteresis in Hog Production, *Applied Economics* **40**(3), 333–340.
- Kostov, P.: 2009, Spatial dependence in agricultural land prices: does it exist?, *Agricultural Economics* **40**(3), 347–353.
- Krueger, J. S. and Lewis-Beck, M. S.: 2008, Is OLS dead?, *The Political Methodologist* **15**(2), 2–4.
- Margarian, A.: 2008, Sind die Pachten im Osten zu niedrig oder im Westen zu hoch?, *Arbeitsberichte aus der vTI-Agrarökonomie - Johann Heinrich von Thünen Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries* (1), 1–42.

- Nickerson, C. J. and Zhang, W.: 2014, Modeling the Determinants of Farmland Values in the United States, *in* J.M. Duke and J. Wu (eds.), *The Oxford Handbook of Land Economics*, Chapter 5, pp. 111–137. Oxford: Oxford University Press.
- Oprea, R., Friedman, D. and Anderson, S. T.: 2009, Learning to Wait: A Laboratory Investigation, *The Review of Economic Studies* **76**(3), 1103–1124.
- Rauchs, A. and Willinger, M.: 1996, Experimental evidence on the irreversibility effect, *Theory And Decision* **40**(1), 51–78.
- Richards, T. J. and Green, G. P.: 2003, Economic Hysteresis in Variety Selection, *Journal of Agricultural and Applied Economics* **35**(1), 1–14.
- Rosen, S.: 1974, Hedonic prices and implicit markets: product differentiation in pure competition, *The Journal of Political Economy* **82**(1), 34–55.
- Sandri, S., Schade, C., Mußhoff, O. and Odening, M.: 2010, Holding on for too long? An experimental study on inertia in entrepreneurs' and non-entrepreneurs' disinvestment choices, *Journal of Economic Behavior and Organization* **76**(1), 30–44.
- Yavaş, A. and Sirmans, C. F.: 2005, Real options: Experimental evidence, *The Journal of Real Estate Finance and Economics* **31**(1), 27–52.



Linear regression and beyond

Abstract: The development of models that go beyond traditional linear regression has been a topic of great interest in statistical research over the last years. As a consequence, a powerful toolbox has emerged, allowing for a realistic modelling of a variety of real data problems. Irrespective of these advances, the parametric ordinary least squares regression model (OLS) is still among the most commonly used techniques in empirical applications. To motivate the methodology that is used for the empirical analyses in subsequent chapters, this chapter briefly introduces the reader to Generalised Linear Models (GLM), Generalised Additive Models (GAM), Generalised Additive Models for Location, Scale and Shape (GAMLSS) and Quantile Regression which iteratively relax the restrictive assumptions of the OLS model and allow for a flexible modelling of the relationship between the response and the explanatory variables of interest.

Keywords: Generalised Linear Models; Generalised Additive Models; Generalised Additive Models for Location, Scale and Shape; Quantile Regression.



Analysing farmland rental rates using Bayesian geoadditive quantile regression

ALEXANDER MÄRZ¹ · NADJA KLEIN² · THOMAS KNEIB² · OLIVER MUSSHOF¹

¹DEPARTMENT OF AGRICULTURAL ECONOMICS AND RURAL DEVELOPMENT

²CHAIR OF STATISTICS

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Published in: *European Review of Agricultural Economics* (doi:10.1093/erae/jbv028)

Abstract: Empirical studies on farmland rental rates so far have predominantly concentrated on modelling conditional means using spatial autoregressive models. While these models only focus on the central tendency of the response variable, quantile regression provides more detailed insight by modelling different points of the conditional distribution as functions of covariates. Based on data from the German agricultural census, this paper contributes to the agricultural economics literature by modelling conditional quantiles of farmland rental rates semi-parametrically using Bayesian geoadditive quantile regression. Our results stress the importance of using semi-parametric regression models, as several covariates influence rental rates in an explicit non-linear way. Moreover, our analysis allows us to uncover potential heterogeneities of the estimated effects across the conditional distribution of rental rates. By explicitly modelling and visually presenting the spatial effects, we also provide additional insight into the spatial structure of German farmland rental rates.

Keywords: Bayesian geoadditive quantile regression; Component-wise boosting; Farmland rental rates; Hedonic pricing models; Heterogeneity; Spatial statistics.

JEL codes: C11, C14, Q15.

Intergenerational social mobility in the United States: a multivariate analysis using distributional regression

ALEXANDER MÄRZ¹ · NADJA KLEIN² · THOMAS KNEIB² · OLIVER MUSSHOF¹

¹DEPARTMENT OF AGRICULTURAL ECONOMICS AND RURAL DEVELOPMENT

²CHAIR OF STATISTICS

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Centre for Statistics Working Paper, Georg-August-Universität Göttingen

Abstract: While previous studies have focused primarily on analyzing the spatial process of social mobility for the conditional mean, we model all parameters of a multivariate response distribution within a regression framework and hereby extend the literature in several ways. First, we relax the assumption of independence between measures of absolute and relative mobility and study their interdependence by relating the coefficient of correlation to economic factors. We find that the correlation between the two measures is non-constant and varies considerably, both with respect to economic variables and across the U.S. Second, and following the recent trend of bringing the attention of economists and sociologists to analyzing the conditional variance, we investigate the within-district variability of social mobility within a heteroscedastic regression framework. Our results suggest a Great Gatsby Curve for the variance, so that higher rates of income inequality increase the risk of being excluded from upward social mobility. Third, our analysis enables us to assess the ability of the factors, as discussed in the literature, to capture the spatial patterns of social mobility across the U.S. The results indicate that although these factors contribute to a better understanding of its underlying mechanisms, they leave spatial heterogeneities of social mobility mostly unexplained.

Keywords: Conditional dependence; Distributional regression; Equality of opportunity; Heteroscedastic regression; Intergenerational social mobility.

JEL codes: C11, C14, H00, J00, R00, R12.

Experimentelle Untersuchung der Optimalität von Investitionsentscheidungen

GESA HOLST¹ · ALEXANDER MÄRZ¹ · OLIVER MUSSHOF¹

¹DEPARTMENT FÜR AGRARÖKONOMIE UND RURALE ENTWICKLUNG
GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Veröffentlicht in: Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung (im Druck)

Zusammenfassung: In der Literatur finden sich zahlreiche Hinweise darauf, dass Entscheider häufig nicht den optimalen Investitionszeitpunkt wählen. Bisher gibt es jedoch nur wenige Studien, die analysieren, wovon das Treffen von zu frühen bzw. zu späten Investitionsentscheidungen abhängig ist. In diesem Beitrag wird dieser Frage nachgegangen, indem experimentell getroffene Investitionsentscheidungen hinsichtlich ihrer Güte untersucht werden. Der Realloptionsansatz wird dabei genutzt, um eine Benchmark herzuleiten. Zur Analyse der Daten, findet ein multinomiales additiv gemischtes Modell Anwendung. Die Ergebnisse machen deutlich, dass die Güte der getroffenen Investitionsentscheidungen von entscheiderspezifischen Variablen abhängig ist. Darüber hinaus beeinflusst das Design des Experiments die Güte der getroffenen Investitionsentscheidungen.

Keywords: Experimentelle Ökonomik; Realloptionen; Investitionsentscheidungen; Entscheidungsverhalten.

JEL codes: C91, D03, D81, D92.

Summary and concluding remarks

'Statistical models for data are never true. The question whether a model is true is irrelevant. A more appropriate question is whether we obtain the correct scientific conclusion if we pretend that the process under study behaves according to a particular statistical model.' (Zeger 1991)

This thesis presents several applications of modern regression techniques in empirical economics that allow to relax some of the restrictive assumptions of more traditional models. By making as many as necessary, but as few as possible, assumptions with respect to the process under study, our analyses address several issues that are currently being discussed in the economics literature and provide additional insight that would not have been detected if more traditional regression models had been used.

The first analysis presented in Chapter III investigates German farmland rental rates within a spatial quantile regression setting. Contributing to the current literature that goes beyond the modelling of the conditional mean, we model conditional quantiles of farmland rental rates which allows us to separately identify the determinants of rental rates for each quantile and to uncover the driving forces behind low, medium and high rents. For expensive rents, we find that mainly those covariates that reflect local competition for farmland or field crops with high profit margins have a strong influence. Regarding the effect of an increased land use for the production of biogas on rental rates, our results show that it influences mod-

erately low, medium and moderately high rental rates only, while it does not seem to have an effect on very expensive rents. However, since we use 'Bestandspachten' as our response variable, the inelasticity of rental rates due to the long duration of the leasing contracts mentioned in the literature might apply to our results. Hence, it may take some time before the effect of biogas is actually reflected in rental rates. Modelling a dense grid of quantiles also enables us to exploit the informational content implicit in farmland rental rates and to reveal heterogeneities of the estimated effects across the conditional response distribution. Our regression results supplement the estimates from traditional models by allowing the effects to vary with conditional quantiles. Consequently, a richer description of the data generating process underlying rental rates can be obtained. Based on our analysis, we find that there are some noticeable instances of heterogeneity across quantiles. This is particularly true for the estimated effects of livestock densities, especially for cattle density both at the farm and the county level, for the share of rented agricultural land, as well as for the share of rented arable land.

We also address the problem of functional form misspecifications by modelling conditional quantiles of farmland rental rates semi-parametrically. The flexibility of this model class alleviates the problems on statistical inference associated with functional form misspecifications and allows us to present a more detailed analysis of farmland rental rates. Our results stress the importance of making use of semi-parametric regression models, as several covariates influence farmland rental rates in an explicit non-linear way. Moreover, by explicitly modelling and visually presenting the spatial effects, we provide additional insight into the spatial structure of German farmland rental rates. Our analysis shows that rental rates are considerably lower than what can be explained with covariates in the Southwest, as well as in large parts of East and Central Germany across all quantiles. Also, the covariates are not well-suited to explain expensive rental rates, especially in the South, as well as in western and northern regions of Germany.

In Chapter IV we analyze the multifaceted dimension of social mobility by modelling all parameter of a multivariate response distribution as function of covariates within a unified regression framework. The focus of the first part lies in analyzing the factors that influence measures of absolute and relative social mo-

bility in a conditional mean regression setting and on investigating how well these factors capture the spatial patterns of social mobility across the United States. By investigating the informational content of the economic variables, as discussed in the sociology and economics literature, to capture the spatial patterns of social mobility across the United States, our results suggest that even though these factors contribute to a better understanding of its underlying mechanisms, they leave spatial heterogeneities of social mobility mostly unexplained.

Being among the first to apply heteroscedastic regression to the analysis of social mobility, we also investigate the within-district variability of absolute and relative social mobility which we interpret as reflecting the insecurity attached to participating in upward social mobility. The analysis of the conditional variance is of interest because the decision of people to remain in or to move to a specific district may not be driven solely by the overall level of social mobility, but also by the chances to participate in upward social mobility. Our results suggest that the variability of absolute mobility is high in the Great Plains, but also in parts of the Southeast, whereas the variability of relative mobility is high mainly on the West Coast. In contrast to conditional mean models, the selected covariates are better suited to capture the variability within districts. However, spatial heterogeneities remain unexplained in some areas where rates of mobility vary considerably. Moreover, our results suggest that in addition to the mean regression case, there is a Great Gatsby Curve for the variance as well. Consequently, higher rates of income inequality do not only reduce the overall level of social mobility, but also seem to increase the risk attached to participate in upward social mobility.

We also relax the assumption of independence between measures of absolute and relative mobility, and contribute to the growing literature that perceives the phenomenon of social mobility as a multivariate one. By explicitly taking the dependency between both measures into account and by modeling the parameter that governs the strength of their interdependence as a function of covariates within a regression setting, we provide insight into the way in which both measures interact and show how the coefficient of correlation between absolute and relative mobility is influenced by economic variables. The spatial pattern of the coefficient of correlation indicates that measures of social mobility show the highest correlation in the Eastern and Southeastern regions of the United States, as well as around the

Rust Belt. In particular, the strength of the coefficient of correlation is strongest in those areas in which the highest and lowest levels of both absolute and relative mobility occur. Moreover, our analysis shows that the coefficient of correlation is non-constant across the domain of the selected covariates and varies considerably according to the values of economic factors.

The analysis presented in Chapter V is concerned with the question as to why decision-makers, in practice, often fail to choose the optimal point in time to invest. Based on existing studies, little empirical evidence regarding this asymmetry is yet available. Hence, it remains unclear, how the deviations between the optimal point in time to invest and the ones observed in practice can be reconciled. In order to identify the factors that influence the investment behaviour, we make use of a comprehensive data set of German and Kazakh farmers.

Our results show that personal characteristics have a significant influence on the investment decisions of farmers. It appears, for instance, that the probability of investing too early is significantly lower for male decision-makers and that, compared to German farmers, the timing of the investment decisions of Kazakh farmers is less optimal. A second conclusion that can be drawn from our results is that experiment-related characteristics also have an impact on the quality of the investment decisions. Our results indicate that the order of the framing, as well as the number of repetitions have an influence on the quality of the investment decision, while the framing itself does not. Furthermore, we find that the probability of investing too late decreases when participants are faced with a known framing. Also, farmers tend to hit the optimal point in time more often as the number of repetitions increases during the experiment.

References

Zeger, Scott L.: 1991, Statistical Reasoning in Epidemiology, *American Journal of Epidemiology* **134**(10), 1062–1066.

List of publications

Publications in refereed scientific journals

Holst, G., März, A. and Musshoff, O.: 2015, Experimentelle Untersuchung der Optimalität von Investitionsentscheidungen, *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung (ZfbF)*; *in print*.

März, A., Klein, N., Kneib, T. and Musshoff, O.: 2015, Analysing farmland rental rates using Bayesian geoaddivitive quantile regression, *To appear in European Review of Agricultural Economics*, doi: 10.1093/erae/jbv028.

URL: <http://erae.oxfordjournals.org/content/early/2015/11/03/erae.jbv028.abstract>

Submitted Papers & Preprints

März, A., Klein, N., Kneib, T. and Musshoff, O.: 2016, Intergenerational social mobility in the United States: a multivariate analysis using distributional regression. *Submitted*.

Other Publications

März, A., Klein, N., Kneib, T. and Musshoff, O.: 2014, Analysing farmland rental rates using Bayesian geoaddivitive quantile regression, *European Association of Agricultural Economists - EAAE, 2014 International Congress, August 26-29, 2014, Ljubljana, Slovenia*.

Erklärung über den geleisteten Eigenanteil an der Arbeit

Hiermit erkläre ich den geleisteten Anteil an den in die Dissertationsschrift aufgenommenen Beiträgen.

Im ersten Beitrag mit dem Titel *'Analysing farmland rental rates using Bayesian geoadditive quantile regression'*, der in Zusammenarbeit mit Frau Dr. Nadja Klein, Herrn Prof. Dr. Thomas Kneib sowie Herrn Prof. Dr. Oliver Mußhoff verfasst wurde, sind folgende Bereiche von mir übernommen worden: Idee und Konzeption des Beitrags, Durchführung der Datenfernverarbeitung bei den Forschungsdatenzentren (FDZ) der Statistischen Ämter des Bundes und der Länder, Durchführung der Berechnungen in Zusammenarbeit mit Herrn Prof. Dr. Thomas Kneib und Frau Dr. Nadja Klein sowie Interpretation der Ergebnisse und Verfassen des Beitrags in Zusammenarbeit mit Herrn Prof. Dr. Oliver Mußhoff.

Im zweiten Beitrag mit dem Titel *'Intergenerational social mobility in the United States: a multivariate analysis using distributional regression'*, der in Zusammenarbeit mit Frau Dr. Nadja Klein, Herrn Prof. Dr. Thomas Kneib sowie Herrn Prof. Dr. Oliver Mußhoff verfasst wurde, sind folgende Bereiche von mir übernommen worden: Idee und Konzeption des Beitrags, Durchführung der Datenbeschaffung und Datenaufbereitung, Durchführung der Berechnungen in Zusammenarbeit

mit Herrn Prof. Dr. Thomas Kneib und Frau Dr. Nadja Klein sowie Interpretation der Ergebnisse und Verfassen des Beitrags in Zusammenarbeit mit Herrn Prof. Dr. Oliver Mußhoff.

Im dritten Beitrag mit dem Titel *'Experimentelle Untersuchung der Optimalität von Investitionsentscheidungen'*, der in Zusammenarbeit mit Frau Dr. Gesa Holst sowie Herrn Prof. Dr. Oliver Mußhoff verfasst wurde, sind folgende Bereiche von mir übernommen worden: Datenbereinigung, Beschreibung des theoretischen Regressionsmodells, Durchführung der Berechnungen sowie Interpretation und Beschreibung der Ergebnisse in Zusammenarbeit mit Frau Dr. Gesa Holst. Des Weiteren wurde die Überarbeitung des Beitrags in Zusammenarbeit mit Frau Dr. Gesa Holst vorgenommen.

Eidesstattliche Erklärungen

Hiermit erkläre ich eidesstattlich, dass:

1. ich diese Arbeit weder in gleicher noch in ähnlicher Form bereits anderen Prüfungsbehörden vorgelegen habe.
2. ich mich an keiner anderen Hochschule um einen Doktorgrad beworben habe.

Göttingen, den

.....
(Unterschrift)

Hiermit erkläre ich eidesstattlich, dass diese Dissertation selbständig und ohne unerlaubte Hilfe angefertigt wurde.

Göttingen, den

.....
(Unterschrift)