Essays on Convex Regression and Frontier Estimation

Sheng Dai
Essays on Convex Regression and Frontier Estimation

Sheng Dai
Supervising professor
Professor Timo Kuosmanen, Turku School of Economics, University of Turku, Finland

Thesis advisor
Professor Timo Kuosmanen, Turku School of Economics, University of Turku, Finland

Preliminary examiners
Professor Mika Kortelainen, Turku School of Economics, University of Turku, Finland
Professor Vincent Charles, School of Management, University of Bradford, United Kingdom

Opponent
Professor Mika Kortelainen, Turku School of Economics, University of Turku, Finland
Abstract

Convex regression is increasingly popular in economics, finance, operations research, machine learning, and statistics. In the productivity and efficiency analysis field, convex regression and its latest development have bridged the long-standing gap between the conventional deterministic nonparametric and stochastic-parametric methods. This dissertation presents three scientific essays that contribute to the field of convex regression and frontier estimation from methodological, computational, and empirical aspects.

Essay I develops a new $L_0$-norm regularization approach to the convex quantile and expectile regressions for subset selection. In the paper, I show how to use mixed-integer programming to solve the proposed $L_0$-norm regularization approach in practice and build a link to the commonly used $L_1$-norm regularization approach. A Monte Carlo study is performed to compare the finite sample performances of the proposed $L_0$-penalized convex quantile and expectile regression approaches with the $L_1$-norm regularization approaches. The proposed approach is further applied to benchmark the sustainable development performance of the OECD countries and empirically analyze the accuracy in the dimensionality reduction of variables. The results from the simulation and application illustrate that the proposed $L_0$-norm regularization approach can more effectively address the curse of dimensionality than the $L_1$-norm regularization approach in multidimensional spaces.

Essay II is motivated by computational and pedagogical needs. The heavy computational burden and the lack of powerful, reliable, and fully open access computational packages have slowed down the diffusion of these advanced estimation techniques to empirical practice. The Python package pyStoNED aims to address this challenge by providing a freely available and user-friendly tool for multivariate convex regression, convex quantile regression, convex expectile regression, and stochastic nonparametric envelopment of data. This paper presents a tutorial on the pyStoNED package and illustrates its application.

Essay III contributes to convex quantile and expectile regressions to more accurately estimate shadow prices and marginal abatement costs of bad outputs. Specifically, using panel data of 30 Chinese provinces during 1997–2015, we first estimate the marginal CO$_2$ abatement costs using a novel data-driven approach, convex quantile regression. Based on the marginal abatement cost estimates and China’s plans regarding carbon intensity reduction and economic growth, we present a forward-looking assessment of the abatement costs for Chinese provinces for 2016–2020. Our main finding is that all the Chinese provinces have a negative abatement cost, which means these provinces can benefit from an increase in the absolute level of CO$_2$ emissions despite the constraint on carbon intensity.

Keywords Abatement cost; Convex regression; Convex quantile regression; Frontier estimation; Regularization; Subset selection

ISSN (printed) 1799-4934 ISSN (pdf) 1799-4942
Location of publisher Helsinki Location of printing Helsinki Year 2022
Luckily, I have come to the final stage of my doctoral study that I have been expecting for several years. It is also the time to look back and recall how I can complete the present dissertation. Take this opportunity to express my warmest thanks to those who have provided advice and support throughout these years.

I thank my supervisor Professor Timo Kuosmanen for his unconditional support and continuous encouragement at different stages of my study. Recall that in December 2015, I got in touch with Professor Kuosmanen for the first time to ask some StoNED-related questions. Ever since, he has been a valuable guide and supporter in my academic career. The constructive and timely feedback has always improved my research work. It is also appreciated that his numerous ideas have enlarged my vision and motivated me to explore different areas. More importantly, I like and enjoy the happy and free working atmosphere with Professor Kuosmanen.

Apart from my Ph.D. supervisor, I would like to extend my deep gratitude to my former supervisor Professor Jinhua Cheng from China University of Geosciences. Without his unconditional support, I could never have challenged myself and pursued a doctoral degree abroad. Particularly, thanks for providing many opportunities to attend various academic conferences at home and abroad during my master’s study. Also, I would like to thank Professors Xiaolin Wang, Jianzhong Xiao, and Jun Chen from my home university for their valuable advice on my academic path.

My preliminary examiners, Professors Mika Kortelainen and Charles Vincent, I thank for their constructive comments, which have greatly helped me improve the presentation of this dissertation. I also feel privileged to have Professor Pekka Malo as the chairperson for the public defense.

I want to thank my co-authors, Dr. Xun Zhou, Professor Chia-Yen Lee, Jiahui Yi, Zhiqiang Liao, Iaroslav Kriuchkov, and Yu-Hsueh Fang, for their significant contribution to our collaborations, even though some of the work are not included in this dissertation. Working with you is full of joy and is a rich source of inspiration for me as a researcher.

I gratefully acknowledge the financial support from the Foundation for Economic Education (Liikesivistysrahasto), the HSE Support Foundation, and the research project from the Prime Minister’s Office of Finland (VN TEAS). The calculations presented in the dissertation were performed using computational resources provided
by the Aalto Science-IT project, which deserves my acknowledgment.

I am also grateful to the Head of the Dept. of Information and Service Management, Professor Markku Kuula, for creating a research environment that facilities my research work. My gratitude extends to Professors Juuso Liesiö, Eeva Vilkkumaa, and Ilkka Leppänen, for their comments and suggestions on my research work at our unit seminars, and Dr. Xiangping Hu from NTNU for sharing his academic experiences, and other outstanding colleagues and friends at Aalto University, Peng Xu, Yanqing Lin, Jie Li, Lina Siltala-Li, Taeyong Kee, Wenjie Fan, and Weimu You. I would also like to thank the administrative staff in the department and the Doctoral Programme of Business, Merja Mäkinen, Heli Solla, Hanne Lehtinen, and other members.

Most importantly, I would like to thank my family for their love, support, and encouragement.

Espoo, August 2022

Sheng Dai
### Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>1</td>
</tr>
<tr>
<td>Contents</td>
<td>3</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>5</td>
</tr>
<tr>
<td>List of Essays</td>
<td>7</td>
</tr>
<tr>
<td>Author’s Contribution</td>
<td>9</td>
</tr>
<tr>
<td>1. Motivation</td>
<td>11</td>
</tr>
<tr>
<td>2. Literature review</td>
<td>13</td>
</tr>
<tr>
<td>2.1 Production function hypotheses</td>
<td>13</td>
</tr>
<tr>
<td>2.2 Methods</td>
<td>15</td>
</tr>
<tr>
<td>2.3 Applications</td>
<td>21</td>
</tr>
<tr>
<td>3. Overview of the essays</td>
<td>25</td>
</tr>
<tr>
<td>3.1 Essay I</td>
<td>25</td>
</tr>
<tr>
<td>3.2 Essay II</td>
<td>26</td>
</tr>
<tr>
<td>3.3 Essay III</td>
<td>27</td>
</tr>
<tr>
<td>References</td>
<td>29</td>
</tr>
<tr>
<td>Essays</td>
<td>35</td>
</tr>
</tbody>
</table>
List of Abbreviations

CER  Convex Expectile Regression
CNLS Convex Nonparametric Least Squares
C$^2$NLS Corrected Convex Nonparametric Least Squares
CO$_2$ Carbon dioxide
COLS Corrected Ordinary Least Squares
CQR Convex Quantile Regression
DEA Data Envelopment Analysis
NLP Nonlinear Programming
QP Quadratic Programming
SDGs Sustainable Development Goals
SFA Stochastic Frontier Analysis
StoNED Stochastic Nonparametric Envelopment of Data
List of Essays

This thesis consists of an overview and of the following scientific essays which are referred to in the text by their Roman numerals.


Author’s Contribution

Essay I: “Variable selection in convex quantile regression: $L_1$-norm or $L_0$-norm regularization?”

The author of the dissertation is the sole author of this essay.

Essay II: “pyStoNED: A Python package for convex regression and frontier estimation”

The author of the dissertation is the first author of this essay.

**Sheng Dai:** Writing - original draft, Conceptualization, Methodology, Formal analysis, Validation, Visualization, Software. **Yu-Hsueh Fang:** Writing - original draft, Formal analysis, Validation, Visualization, Software. **Chia-Yen Lee:** Writing - review & editing, Conceptualization, Methodology, Supervision. **Timo Kuosmanen:** Writing - review & editing, Conceptualization, Methodology, Supervision.

Essay III: “Forward-looking assessment of the GHG abatement cost: Application to China”

The author of the dissertation is the first author of this essay.

**Sheng Dai:** Writing - original draft, Conceptualization, Methodology, Formal analysis, Visualization, Software. **Xun Zhou:** Writing - review & editing, Conceptualization, Methodology, Formal analysis, Software. **Timo Kuosmanen:** Writing - review & editing, Conceptualization, Methodology, Supervision.
1. **Motivation**

Nonparametric regression refers to a category of regression analysis where we do not assume a priori functional form for an object being estimated. Typically, nonparametric regression imposes less restrictive assumptions such as smoothness and moment restrictions on functional forms, which are actually rarely known and even misspecified (Li & Racine, 2007). The flexible nonparametric techniques can help capture the underlying complex dependence structure and bring a great deal of benefits to economics, econometrics, statistics, and other disciplines (see, e.g., Yatchew, 1998).

A classic approach to nonparametric regression is to build upon global shape constraints such as monotonicity (Brunk, 1955, 1958) and/or concavity/convexity of the regression function (Hildreth, 1954). However, extending such approaches from the univariate setting to the more general multivariate regression has proved a vexing challenge. By deriving an explicit piece-wise linear characterization for regression function, Kuosmanen (2008) introduces convex nonparametric least squares (CNLS), a multivariate convex regression approach, which later attracts great interest in econometrics, statistics, operations research, and machine learning (e.g., Magnani & Boyd, 2009; Seijo & Sen, 2011; Lim & Glynn, 2012; Hannah & Dunson, 2013; Yagi et al., 2020; Bertsimas & Mundru, 2021). The recent studies by Dai (2021) and Dai et al. (2022) impose additional regularization on convex regression to address its various intrinsic problems.

Convexity and monotonicity are commonly seen constraints in economic theory. Typically, the duality theory of production and consumption directly implies certain monotonicity and convexity/concavity properties for many functions of interest (e.g., Afriat, 1967, 1972; Varian, 1982, 1984). For example, the cost function of the firm must be monotonic increasing and convex with respect to the input prices. Similar to the fact that a density function must be non-negative and its definite integral is equal to one, the cost function must satisfy the monotonicity and convexity properties implied by the theory. Otherwise it is not really a cost function at all. The recent developments in the convex regression enable researchers to impose the concavity or convexity constraints implied by the theory to estimate the functions of interest without any parametric functional form assumptions.

Convex regression has been widely applied to estimate the cost or production
frontiers in multidisciplinary fields such as banking, environment, health care, energy, transportation, and utilities (see a detailed literature review in Section 2.3). However, the cost or production frontiers are more often estimated by other two competing conventional paradigms: Data Envelopment Analysis (DEA) (Charnes et al., 1978), a deterministic, fully nonparametric approach; and Stochastic Frontier Analysis (SFA) (Aigner et al., 1977; Meeusen et al., 1977), a probabilistic, fully parametric approach. To bridge the gap between these two paradigms, Stochastic Nonparametric Envelopment of Data (StoNED) (Kuosmanen, 2006; Kuosmanen & Kortelainen, 2012) is proposed as a unified framework that combines virtues of DEA and SFA, encompassing both approaches as its restricted special cases.

With the in-depth understanding of convex regression and frontier estimation, we have noticed several drawbacks that need to be further investigated. We will comment on each in turn.

1) **Curse of dimensionality.** The convergence rate of nonparametric models is directly related to the number of variables involved; see the convex regression case in Kur et al. (2020). High dimensional nonparametric model thus yields a slow rate of convergence even for moderate datasets and the low prediction accuracy and exploratory power. How to model the production/cost frontier using convex regression with a larger number of inputs is definitely a tricky challenge.

2) **Computation.** One solution to the curse of dimensionality is to increase the sample size. However, the computation burden of a sufficiently large sample size has the use of convex regression and frontier estimation, since the sample will generate $O(n^2)$ linear constraints (Lee et al., 2013; Mazumder et al., 2019). A more efficient algorithm is needed. Further, lacking reliable tools restricts the applications of convex regression to frontier estimation, which calls for a powerful, reliable, and open access computational tool.

3) **Impacts of inefficiency and noise.** When estimating shadow price, the conventional full frontier approaches (e.g., DEA or StoNED) tend to neglect the impact of inefficiency. Such ignorance will potentially result in an overestimation, which further induces the inefficient policy implications. Moreover, the full frontier estimations are sensitive to noise and heteroscedasticity. How to mitigate the effects of inefficiency and noise in shadow price and even abatement cost estimation is obviously a further step to increase the attractiveness of convex regression.

Other deficiencies such as nonunique subgradients and quantile crossings are observed and solved in our recent studies (see, e.g., Dai et al., 2022). The overfitting problem in convex regression has also been noticed by Mazumder et al. (2019).

The overarching objective of this dissertation is to address the above three problems. To pursue this main goal, this dissertation incorporates three essays that propose a new convex regression model for subset selection, develop an open source Python package, and present an empirical application using quantile estimation.
Consider the following nonparametric regression model with observations \( \{(x_i, y_i)\} \)

\[ y_i = f(x_i) + \varepsilon_i, \quad \text{for } i = 1, \ldots, n, \]

where \( y_i \in \mathbb{R} \) is the output variable, \( x_i = (x_{i1}, \ldots, x_{id})' \in \mathbb{R}^d \) is the \( d \)-dimensional input variables, \( \varepsilon_i \) is a random variable satisfying \( E[\varepsilon_i | x_i] = 0 \), and \( f : \mathbb{R}^d \rightarrow \mathbb{R} \) is an unknown regression function to be estimated. In the context of productivity and efficiency analysis, production function can be characterized by \( f \) under some conditions (see, e.g., Afriat, 1972; Varian, 1984).

2.1 Production function hypotheses

To estimate production function, the regression function \( f \) in (2.1) needs to satisfy the following three increasingly restrictive possibilities (Afriat, 1972)

(P1) \( f \) non-decreasing (free disposal of input);
(P2) \( f \) non-decreasing concave (classical);
(P3) \( f \) non-decreasing concave conical (classical constant returns to scale).

Such possibilities are also consistent with the classical axioms of the benchmark technology (see, e.g., Koopmans, 1951; Färe & Primont, 1995) and further demonstrated by Afriat (1972) with the corresponding representation theorems. Specifically, for any given \( (x_i, y_i) \), the functions \( f \) satisfying the above possibilities can be characterized as a class of piece-wise linear function

\[ F_{P1}(x) = \max \{ y_i | x_i \leq x \}; \]

\[ F_{P2}(x) = \max \left\{ \sum_{i=1}^{n} y_i \lambda_i \mid \sum_{i=1}^{n} x_i \lambda_i \leq x, \sum_{i=1}^{n} \lambda_i = 1, \lambda_i \geq 0 \right\}; \]

\[ F_{P3}(x) = \max \left\{ \sum_{i=1}^{n} y_i \lambda_i \mid \sum_{i=1}^{n} x_i \lambda_i \leq x, \lambda_i \geq 0 \right\}. \]
Varian (1984) further demonstrates that the function $f$ needs to satisfy the following weak axiom of profit maximization theorem

$$y^j \leq y^i + (w^j/p^j)(x^j - x^i) \quad \forall i, j,$$

where $w$ and $p$ are inputs and outputs price, respectively. This theorem in fact means that the function $f$ is a continuous, concave, monotonic function, which is also a piece-wise linear function.

Note that those theorems state that if there exists a function that can "rationalize" the observed data, then there exists a monotonic and concave piece-wise linear function that also rationalizes the data. That is, even if the true function $f$ is smooth, one cannot ever distinguish it from a piece-wise linear function. Furthermore, some studies (see, e.g., Boyd & Vandenberghe, 2004) assume that the true function $f$ is piece-wise linear, which appears to be redundant and misleading. Building on Afriat-Varian results, Kuosmanen (2008) shows that the piece-wise linear functional form can represent any arbitrary monotonic and concave function. Also the upper and lower bounds are piece-wise linear functions. Recent work (see, e.g., Yagi et al., 2020) indicates that while kernel smoothing is introduced to the objective function, the estimated production function remains piece-wise linear.

Following Kuosmanen (2008), we can easily transform the primal characterization of $F_{P2}(x)$ or $F_{P3}(x)$ to its dual form

$$f_j = \min_{\alpha \in \mathbb{R}, \beta \in \mathbb{R}^d} \{ \alpha + \beta' x \mid \alpha + \beta' x_i \geq f_i \quad \forall i, j \},$$

(2.2)

where $f_j$ is the estimated function value $f(x_j)$, $\beta_j$ represents supporting hyperplanes to epigraph $f$, and $\alpha_j = f_j - \beta_j' x_j$. From the dual characterization (2.2), we can obtain a class of continuous, monotonic, and concave functions, i.e., Afriat inequalities (Afriat, 1967; Varian, 1984)

$$\beta'_i (x_i - x_j) \leq f_i - f_j \quad \forall i, j;$$

or $\alpha_i + \beta'_i x_i \leq \alpha_j + \beta'_j x_i \quad \forall i, j.$

Such Afriat inequalities are essential for our purpose due to that they bridge the gap between production function estimation and convex regression (see, e.g., Yatchew, 1998; Kuosmanen, 2008). We thus safely use the Afriat inequalities to model the concavity constraints in convex regression.\(^1\) So far, we are ready to estimate the production function using convex regression.\(^2\)

---

\(^1\)We could revise the inequality sign in Afriat’s inequalities to impose the convexity constraints.

\(^2\)Other approaches to estimating the production function are reviewed in Kuosmanen & Johnson (2010) and Keshvari & Kuosmanen (2013).
2.2 Methods

2.2.1 Isotonic regression

If the production function $f$ is assumed to be non-decreasing, i.e., satisfying the possibility (P1), we can use the least-squares based isotonic regression approach (Brunk, 1955; Ayer et al., 1955) to estimate this unknown monotonic function. In the univariate setting ($d = 1$), isotonic regression is built based on the order of the single $x$ (i.e., $x_1 \leq x_2 \leq \ldots \leq x_n$), while in the multivariate setting ($d > 1$), it is built based on a partial ordering of $x$.

Let $\chi := \{x_i \in \mathbb{R}_+^d\}$ be a non-empty set with $d$ distinct elements in a metric space with a partial order. Consider the production function $f$ is isotonic with respect to a partial ordering on $\chi$. If $x_i, x_j \in \chi$, then $x_i \preceq x_j$ implies $f(x_i) \leq f(x_j)$. The least-squares based isotonic regression is

$$\min \sum_{i=1}^{n} (y_i - f(x_i))^2$$

s.t. $f : \mathbb{R}_+^d \rightarrow \mathbb{R}$ is monotonic

where if the partial ordering is defined as the standard dominance relation (i.e., $x_i \preceq x_j$ if $x_i \leq x_j$), then the non-decreasing production function $f$ satisfies the monotonicity, that is, isotonicity is equivalent to monotonicity. However, the partial ordering can be defined using other criteria (e.g., revealed preference information), where now isotonicity is not exactly the same as monotonicity.

Isotonic regression is the base of convex regression. The univariate isotonic regression is quite similar to the univariate convex regression; see more details in the following section. The multivariate isotonic regression can be easily extended to the isotonic convex regression or isotonic convex quantile regression (see, e.g., Keshvari & Kuosmanen, 2013; Dai et al., 2021). Other appealing features and properties of isotonic regression (2.3) are summarized in Groeneboom & Jongbloed (2014). However, the classical isotonic regression is known to be inconsistent at boundaries (i.e., the “spiking” problem) (see, e.g., Pal, 2008; Wu et al., 2015). Moreover, when $d > 1$, the estimator (2.3) can overfit the data (see, e.g., Luss & Rosset, 2014; Wu et al., 2015).

2.2.2 Convex regression

When the production function $f$ is assumed to be non-decreasing concave, i.e., satisfying the possibility (P2), convex regression for function $f$ is obtained as the optimal solution to the following infinite dimensional optimization problem

$$\min \sum_{i=1}^{n} (y_i - f(x_i))^2$$

s.t. $f : \mathbb{R}_+^d \rightarrow \mathbb{R}$ is monotonic and concave
In the univariate setting \((d = 1)\), convex regression (2.4) can be solved by sorting the data in ascending order according to the value of \(x\) (i.e., \(x_1 \leq x_2 \leq \ldots \leq x_n\)) and be transformed to the finite dimensional problem (Hildreth, 1954)

\[
\min \sum_{i=1}^{n} (y_i - f_i)^2 \quad \text{(2.5)}
\]

\[
\text{s.t. } f_i \leq f_{i+1} \quad \forall i, \ldots, n - 1 \\
\frac{f_{i+1} - f_i}{x_{i+1} - x_i} \geq \frac{f_{i+2} - f_{i+1}}{x_{i+2} - x_{i+1}} \quad \forall i, \ldots, n - 1
\]

The statistical properties (e.g., consistency and rate of convergence) of the estimator (2.5) have been well investigated in existing literature (see, e.g., Hanson & Pledger, 1976; Mammen, 1991; Groeneboom et al., 2001). Note that since there exists the constraint on the derivative of \(f\) in problem (2.5), univariate convex regression (2.5) is quite similar to isotonic regression (2.3) with \(d = 1\) (see, e.g., Brunk, 1955; Hall & Huang, 2001; Neelon & Dunson, 2004; Shively et al., 2009).

While Hildreth (1954) briefly discusses the multivariate setting with \(d > 1\), Hildreth’s attempt to estimate the multivariate model fails because there is no unique ordering of a vector of explanatory variables \(x\). Several authors (e.g., Holloway, 1979) propose other algorithms for estimating the multivariate model, but those algorithms are subject to the same fundamental flaw as Hildreth’s.

Building on the results of Afriat (1967, 1972), Kuosmanen (2008) develops the first operational solution to the multivariate convex regression problem. Specifically, Kuosmanen considers the least squares estimator of the multivariate convex regression model (2.4) and proves that the optimal solution to this infinite dimensional optimization problem is always equivalent to the optimal solution to the following quadratic programming (QP) problem (also known as the CNLS problem)

\[
\min_{\alpha, \beta, \varepsilon} \sum_{i=1}^{n} \varepsilon_i^2 \quad \text{(2.6)}
\]

\[
\text{s.t. } y_i = \alpha_i + \beta_i' x_i + \varepsilon_i \quad \forall i \\
\alpha_i + \beta_i' x_i \leq \alpha_j + \beta_j' x_i \quad \forall i, j \\
\beta_i \geq 0 \quad \forall i
\]

The first constraint in (2.6) can be interpreted as the multivariate regression equation. The second constraint, a set of Afriat inequalities, enforces concavity of the piecewise linear regression function \(f\). The third constraint imposes monotonicity on regression function \(f\). When the production function \(f\) satisfies a more strict assumption (P3), we additionally impose \(\alpha_i = 0\). Furthermore, the uniqueness, consistency, unbiasedness, and rate of convergence of the CNLS estimator (2.6) have been investigated by Seijo & Sen (2011), Lim & Glynn (2012), and Kur et al. (2020).

While the literature on convex regression focuses almost exclusively on the additive model introduced above, in economic applications, it is often more natural to posit a
multiplicative model (Kuosmanen & Kortelainen, 2012)

\[ y_i = f(x_i) \exp(\varepsilon_i), \quad (2.7) \]

Taking logarithm of both sides yields\(^3\)

\[ \ln y_i = \ln f(x_i) + \varepsilon_i. \]

The corresponding multiplicative CNLS estimator is thus rephrased as the following nonlinear programming (NLP) problem

\[
\min_{\alpha, \beta, \varepsilon} \sum_{i=1}^{n} \varepsilon_i^2 
\text{s.t.} \quad \ln y_i = \ln(\alpha_i + \beta'_i x_i) + \varepsilon_i \quad \forall i \\
\alpha_i + \beta'_i x_i \leq \alpha_j + \beta'_j x_i \quad \forall i, j \\
\beta_i \geq 0 \quad \forall i \\
\varepsilon_i^+ \geq 0, \quad \varepsilon_i^- \geq 0 \quad \forall i
\]

While CNLS is completely free of tunable parameters, it is sensitive to random noise and heteroscedasticity. By analogy with linear quantile regression (Koenker & Bassett, 1978), CNLS is extended to estimate the conditional medians or other quantiles/expectiles of dependent variables (Wang et al., 2014; Kuosmanen et al., 2015). The nonparametric model (2.1) is rephrased as the following nonparametric conditional quantile model

\[ Q_{y_i}(\tau \mid x_i) = f_{\tau}(x_i) + F_{\varepsilon_i}^{-1}(\tau), \quad \text{for} \quad i = 1, \ldots, n \]

where \( \tau \in (0, 1) \) represents the order of quantile and \( F_{\varepsilon_i} \) represents the cumulative distribution function of the errors.

Following Wang et al. (2014), given a quantile \( \tau \), we estimate the \( \tau^{\text{th}} \) quantile production function \( Q_{y_i}(\tau \mid x_i) \) using the following convex quantile regression (CQR) approach

\[
\min_{\alpha, \beta, \varepsilon^+ \varepsilon^-} \tau \sum_{i=1}^{n} \varepsilon_i^+ + (1 - \tau) \sum_{i=1}^{n} \varepsilon_i^- 
\text{s.t.} \quad y_i = \alpha_i + \beta'_i x_i + \varepsilon_i^+ - \varepsilon_i^- \quad \forall i \\
\alpha_i + \beta'_i x_i \leq \alpha_j + \beta'_j x_i \quad \forall i, j \\
\beta_i \geq 0 \quad \forall i \\
\varepsilon_i^+ \geq 0, \quad \varepsilon_i^- \geq 0 \quad \forall i
\]

where the objective function minimizes the asymmetric absolute deviations from quantile function and the error term \( \varepsilon \) is now decomposed into two non-negative

\(^3\)Note that the multiplicative formulation is not just an inconsequential data transformation due to that \( \ln f(x_i) \neq f(\ln(x_i)) \).
components $\varepsilon_i^+, \varepsilon_i^-$. However, the estimated quantile functions $\hat{Q}_{y_i}(\tau \mid x_i)$ are non-unique. To guarantee the uniqueness of estimated quantile functions, Kuosmanen et al. (2015) propose an indirect estimation of quantiles through expectile regression (Newey & Powell, 1987) by replacing the objective function of (2.9) with the quadratic objective function $\min \sum \varepsilon_i^+ + (1 - \tau) \sum \varepsilon_i^-$ (referred to as convex expectile regression (CER)). The quantile and expectile properties are demonstrated in Wang et al. (2014) and Kuosmanen & Zhou (2021).\footnote{The corresponding multiplicative CQR and CER estimators are available in Kuosmanen, Zhou & Dai (2020) and Dai et al. (2020).} Note that the CNLS estimator is a special variant of the CER estimator, i.e., expectile $\tilde{\tau} = 0.5$.

Recently, several other extensions have been proposed to address such longstanding problems as overfitting, the curse of dimensionality, and quantile crossings in convex regression. Regularization is the main applied technique used in these extensions. We proceed to review the latest development of convex regression.

First, convex regression would lead to an overfitting problem when confronted with a limited data sample size in multidimensional spaces, especially when the estimated production function approaches the boundary of the convex hull of the production set (Mazumder et al., 2019). To ameliorate this problem, following Mazumder et al. (2019), we can introduce a bounded Lipschitz norm to CNLS

$$\min_{\alpha, \beta, \varepsilon} \sum_{i=1}^{n} \varepsilon_i^2 \quad (2.10)$$

subject to

$$y_i = \alpha_i + \beta_i' x_i + \varepsilon_i \quad \forall i$$

$$\alpha_i + \beta_i' x_i \leq \alpha_j + \beta_j' x_i \quad \forall i, j$$

$$\beta_i \geq 0 \quad \forall i$$

$$\|\beta_i\|_2 \leq L \quad \forall i$$

where $\|\cdot\|_2$ represents the standard Euclidean norm (i.e., $L_2$-norm) and the parameter $L > 0$ is the Lipschitz constant.

Second, the curse of dimensionality is an acknowledged challenge in nonparametric estimation. To mitigate the effects of dimensionality, Dai (2021) introduces two different penalty functions, $L_1$ and $L_0$ norms, to CQR and CER. The $L_1$-CQR approach is formulated as

$$\min_{\alpha, \beta, \varepsilon^+, \varepsilon^-} \sum_{i=1}^{n} \varepsilon_i^+ + (1 - \tau) \sum_{i=1}^{n} \varepsilon_i^- + \lambda \sum_{i=1}^{n} \|\beta_i\|_1 \quad (2.11)$$

subject to

$$y_i = \alpha_i + \beta_i' x_i + \varepsilon_i^+ - \varepsilon_i^- \quad \forall i$$

$$\alpha_i + \beta_i' x_i \leq \alpha_j + \beta_j' x_i \quad \forall i, j$$

$$\beta_i \geq 0 \quad \forall i$$

$$\varepsilon_i^+ \geq 0, \varepsilon_i^- \geq 0 \quad \forall i$$
And the $L_0$-CQR approach is

$$\min_{\alpha, \beta, \epsilon^+, \epsilon^-} \tau \sum_{i=1}^{n} \epsilon_i^+ + (1 - \tau) \sum_{i=1}^{n} \epsilon_i^-$$

s.t.  

$$y_i = \alpha_i + \beta_i^i x_i + \epsilon_i^+ - \epsilon_i^- \quad \forall i$$

$$\alpha_i + \beta_i^i x_i \leq \alpha_j + \beta_j^j x_i \quad \forall i, j$$

$$\beta_i \geq 0 \quad \forall i$$

$$\epsilon_i^+ \geq 0, \epsilon_i^- \geq 0 \quad \forall i$$

$$\|\beta_i\|_0 \leq k \quad \forall i$$

where $\lambda$ and $k$ are predetermined tuning parameters. By replacing the linear objective function with quadratic objective function, we can obtain the corresponding $L_1$-CER and $L_0$-CER approaches. Note that problem (2.12) can be converted to a computational mixed integer linear programming problem (mixed integer quadratic programming problem in the $L_0$-CER problem).

Third, when separately estimating multiple quantiles to obtain a family of conditional quantile functions, two or more quantile curves may cross or even overlap when the distribution functions and their associated inverse functions are not monotone increasing (He, 1997). To address such quantile crossing problem, Dai et al. (2022) propose the following penalized CQR approach

$$\min_{\alpha, \beta, \epsilon^+, \epsilon^-} \tau \sum_{i=1}^{n} \epsilon_i^+ + (1 - \tau) \sum_{i=1}^{n} \epsilon_i^- + \gamma \sum_{i=1}^{n} \|\beta_i\|_2^2$$

s.t.  

$$y_i = \alpha_i + \beta_i^i x_i + \epsilon_i^+ - \epsilon_i^- \quad \forall i$$

$$\alpha_i + \beta_i^i x_i \leq \alpha_j + \beta_j^j x_i \quad \forall i, j$$

$$\beta_i \geq 0 \quad \forall i$$

$$\epsilon_i^+ \geq 0, \epsilon_i^- \geq 0 \quad \forall i$$

where $\|\cdot\|_2^2$ is the square of the $L_2$-norm and $\gamma$ is the Tikhonov form tuning parameter. Note that the tuning parameters involved in the above penalized convex regression models can be determined by the standard cross-validation approach from a wide range of candidates. Another appealing feature of Tikhonov regularization is to guarantee the uniqueness of $\beta_i$ estimates. Such regularization can also be introduced to CNLS to ensure unique estimation.

In practice, there is another concern when using this series of convex regression approaches. Estimating a larger sample becomes excessively expensive in convex regression due to the $O(n^2)$ linear constraints (see, e.g., Lee et al., 2013; Mazumder et al., 2019). For example, if the sample contains 500 observations, the total number of linear constraints amounts to 250000. Undoubtedly, the computational burden becomes a barrier to the application of convex regression under larger sample sizes and calls for a more efficient algorithm.
Recent work has shown several promising algorithms from the perspective of computational burden (see, e.g., Lee et al., 2013; Balázs et al., 2015; Mazumder et al., 2019; Bertsimas & Mundru, 2021; Lin et al. 2020). It is worth highlighting two of them here, namely, the CNLS-G algorithm proposed by Lee et al. (2013) and the cutting-plane algorithm proposed by Balázs et al. (2015) and extended by Bertsimas & Mundru (2021). Although both algorithms use the relaxed Afriat inequality constraint set and iteratively introduce new ones as necessary, the cutting-plane algorithm is more efficient to solve the convex regression problems due to its fast identification of violating constraints (Dai, 2021).

2.2.3 Frontier estimation

Frontier estimation relies on the specification of the random variable $\varepsilon_i$ in (2.1), which is usually assumed to be $\varepsilon_i = v_i - u_i$, where $v_i \sim N(0, \sigma_v^2)$ denotes stochastic noise and $u_i \sim N^+(0, \sigma_u^2)$ is inefficiency. If the noise $v_i$ is assumed away, then the production model (2.1) is a deterministic model, stochastic model otherwise.

We can estimate the deterministic frontier using either sign-constrained CNLS or corrected CNLS. Following Kuosmanen & Johnson (2010), the sign-constrained CNLS problem is formulated as

$$\min_{\alpha, \beta, \varepsilon} \sum_{i=1}^{n} \varepsilon_i^2 \quad \text{(2.14)}$$

subject to

$$y_i = \alpha_i + \beta_i' x_i + \varepsilon_i \quad \forall i$$

$$\alpha_i + \beta_i' x_i \leq \alpha_j + \beta_j' x_i \quad \forall i, j$$

$$\beta_i \geq 0 \quad \forall i$$

$$\varepsilon_i \leq 0 \quad \forall i$$

where problem (2.14) is a variant of problem (2.6) by imposing additional constraint on residuals (i.e., the last constraint). The estimated $\hat{f}_i$ is the deterministic production frontier. Note that if the sign in Afriat inequalities (i.e., the second constraint) is $\geq$, then the estimated frontier is the cost frontier. The corresponding inefficiency can be estimated by $\theta_i = 1 - \hat{\varepsilon}_i / y_i$.

Corrected CNLS ($C^2\text{NLS}$) is analogy to corrected ordinary least squares (COLS), where CNLS replaces the first-stage ordinary least squares (OLS) regression (Kuosmanen & Johnson, 2010). Similar to COLS, the $C^2\text{NLS}$ approach

- Estimates $E[y_i | x_i]$ by solving the CNLS model, e.g., Problem (2.6). Denote the CNLS residuals by $\varepsilon_i^{CNLS}$;

- Shifts the residuals analogous to the COLS procedure. The $C^2\text{NLS}$ estimator is

$$\hat{\varepsilon}_i^{C^2\text{NLS}} = \varepsilon_i^{CNLS} - \max_j \varepsilon_j^{CNLS}$$
where values of $\hat{\varepsilon}_i^{C2NLS}$ range from $[0, +\infty]$ with 0 indicating efficient performance. Similarly, we adjust the CNLS intercepts $\alpha_i$ as

$$\hat{\alpha}_i^{C2NLS} = \alpha_i^{CNLS} + \max_j \varepsilon_j^{CNLS}$$

where $\alpha_i^{CNLS}$ is the optimal intercept for firm $i$ in the above CNLS problem and $\hat{\alpha}_i^{C2NLS}$ is the C$^2$NLS estimator.

The estimated C$^2$NLS frontier is then computed by

$$\hat{f}_i^{C2NLS} = \hat{f}_i^{CNLS} + \max_j \varepsilon_j^{CNLS},$$

and the inefficiency (i.e., output efficiency scores) is calculated by

$$\theta_i = y_i / \hat{f}_i^{C2NLS}.$$ 

Note that the estimated C$^2$NLS frontier always envelops the sign-constrained CNLS frontier (i.e., DEA frontier) (Kuosmanen et al., 2015).

In practice, however, external stochastic shocks (e.g., natural disasters) often occur such that assuming away the stochastic noise $v$ from random residuals may cause bias in the estimated frontier. Furthermore, the conventional stochastic frontier (i.e., the SFA frontier) requires a priori functional form. Combining virtues of SFA and DEA in a unified framework, StoNED (Kuosmanen, 2006; Kuosmanen & Kortelainen, 2012) considers the stochastic shock ($v$) and does not require the functional form assumption. Analogous to the COLS/C$^2$NLS estimators, the StoNED estimator consists of the following four steps:

1) Estimate the conditional mean $E[y_i | x_i]$ using CNLS estimator;
2) Estimate the unconditional expected inefficiency $\mu$ based on the residuals $\varepsilon_i^{CNLS}$ using the parametric method of moments/quasi-likelihood estimation or the nonparametric deconvolution approach;
3) Estimate the StoNED frontier $\hat{f}_i^{StoNED}$ based on the $\hat{\mu}$ using

$$\hat{f}_i^{StoNED} = \hat{f}_i^{CNLS} + \hat{\mu};$$

4) Estimate firm-specific inefficiencies $E[u_i | \varepsilon_i^{CNLS}]$ (i.e., conditional mean) using the parametric JLMS estimator (Jondrow et al., 1982) or the nonparametric deconvolution approach (Horrace & Parmeter, 2011).

See a more detailed description of the above steps in Kuosmanen et al. (2015) and Dai et al. (2021). Note that the quantile and expectile related estimators cannot be integrated into the StoNED framework at present. This is a potential fascinating avenue for future research.

### 2.3 Applications

With the development of convex regression and frontier estimation, CNLS/StoNED and their extensions have been widely applied to efficiency analysis, productivity analysis, and shadow pricing (Table 1). This section starts to summarize the typical applications of convex regression and frontier estimation.
Besides acting as a unified framework, StoNED has another feature that it can directly incorporate contextual variables to frontier estimation. Most applications resort to one-stage CNLS/StoNED to investigate the impact of contextual variables on efficiency (e.g., Kuosmanen, 2012; Eskelinen & Kuosmanen, 2013; Molinos-Senante & Maziotis, 2022). Compared to the traditional two-stage DEA approach (i.e., DEA + regression), StoNED can reduce the estimated bias in contextual variables, particularly when contextual variables are correlated with inputs (Johnson & Kuosmanen, 2011; Johnson & Kuosmanen, 2012). Recently, Kuosmanen et al. (2021) incorporate contextual variables to the CER estimator to evaluate the performance for English National Health Service hospitals in combating COVID-19 pandemic.

Table 1. Typical applications of convex regression and frontier estimation.

<table>
<thead>
<tr>
<th>Field</th>
<th>Method</th>
<th>Sample</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency analysis</td>
<td>StoNED</td>
<td>Electric grid utilities in China</td>
<td>Li et al. (2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Electricity distribution networks in Finland</td>
<td>Kuosmanen (2012); Kuosmanen et al. (2013); Dai &amp; Kuosmanen (2014); Dai (2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bank branches in Finland</td>
<td>Eskelinen &amp; Kuosmanen (2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Water and sewerage companies in Chile</td>
<td>Molinos-Senante &amp; Maziotis (2021)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wastewater treatment plants in Chile</td>
<td>Molinos-Senante &amp; Maziotis (2022)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NHS hospitals in the UK</td>
<td>Kuosmanen, Tan &amp; Dai (2020)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Countries</td>
<td>Cordero et al. (2020); Lu et al. (2021)</td>
</tr>
<tr>
<td></td>
<td>CNLS</td>
<td>Registered passenger cars in Finland</td>
<td>Zhou (2018); Zhou &amp; Kuosmanen (2020)</td>
</tr>
<tr>
<td></td>
<td>$L_0$-CER</td>
<td>OECD countries</td>
<td>Dai (2021)</td>
</tr>
<tr>
<td>Productivity analysis</td>
<td>StoNED</td>
<td>OECD countries</td>
<td>Kuosmanen (2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Electricity distribution companies in Norway</td>
<td>Cheng, Bjørndal &amp; Bjørndal (2015); Cheng, Bjørndal, Lien &amp; Bjørndal (2015)</td>
</tr>
<tr>
<td>Shadow pricing</td>
<td>StoNED</td>
<td>Bituminous coal power plants in the US</td>
<td>Mekaroonreung &amp; Johnson (2012); Mekaroonreung &amp; Johnson (2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Provincial regions in China</td>
<td>Shen &amp; Lin (2017)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coal-fired power plants in China</td>
<td>Lee &amp; Wang (2019)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transportation modes</td>
<td>Chen et al. (2021)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Iron and steel industry in China</td>
<td>Xian et al. (2022)</td>
</tr>
<tr>
<td></td>
<td>CER</td>
<td>Electric power plants in the US</td>
<td>Kuosmanen &amp; Zhou (2021); Zhao &amp; Qiao (2022)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OECD countries</td>
<td>Kuosmanen, Zhou &amp; Dai (2020)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Provincial regions in China</td>
<td>Dai et al. (2020)</td>
</tr>
</tbody>
</table>
The applications in productivity analysis mainly include three aspects: full frontier productivity analysis (e.g., Kuosmanen, 2013), quantile performance evaluation (e.g., Dai, 2021), and index decomposition analysis (e.g., Zhou & Kuosmanen, 2020). Similar to the DEA Malmquist index, the StoNED Malmquist index also relies on the full frontier to gauge productivity growth. However, such full frontier estimation ignores the impact of inefficiency and is sensitive to the choice of the direction vector and outliers (Lee et al., 2002; Kuosmanen & Zhou, 2021), easily leading to an overestimation of the Malmquist index. As a compelling extension to the full frontier StoNED Malmquist index, it is also possible to evaluate the quantile performance using the CER approach (Dai, 2021).

Similar to DEA and parametric programming, StoNED is another essential paradigm for shadow pricing (see, e.g., Mekaroonreung & Johnson, 2012; Xian et al., 2022). While the stochastic stocks are considered in StoNED when estimating shadow prices, StoNED requires \textit{a priori} distribution for inefficiency $u$ (e.g., half norm distribution) and noise $v$ (e.g., norm distribution) to identify that full frontier. Moreover, the full frontier based shadow price estimation ignores the impact of inefficiency. To estimate shadow price more accurately, Kuosmanen & Zhou (2021) recently introduce the quantiles to shadow pricing estimation, where we do not need any parametric distributional assumptions and can take inefficiency explicitly into account.

Overall, convex regression offers a new paradigm shift for productivity and efficiency analysis and is applied to a large number of real-world cases, which, in turn, motivate further methodological innovations.
3. Overview of the essays

This dissertation consists of three distinct essays, which contribute to the field of convex regression and frontier estimation from methodological, computational, and empirical aspects. In summary, this dissertation has

1) proposed a new $L_0$-norm regularization approach to convex quantile and expectile regressions for subset variable selection;
2) developed a powerful, reliable, and fully open access computational Python package, pyStoNED, for estimating convex regression and frontier estimation;

We next briefly summarize each in turn.

3.1 Essay I

Essay I is motivated by an empirical application to sustainable development goals (SDGs). The 2030 Agenda for Sustainable Development consists of 231 unique indicators that are designed to measure 169 targets and monitor progress toward the achievement of the 17 SDGs. Consequently, the integrated assessment approaches based on such a high number of indicators are immensely challenging to be implemented due to correlation among the indicators and the different units of measurement. The need and challenge to better benchmark the degree of sustainable development of countries triggered the emerging integration of SDG assessment and production economics.

In general, high-dimensional data space weakens prediction accuracy and exploratory power of nonparametric estimators such as CNLS and CQR/CER. Improving the performance of these estimators with sparse data has motivated scholars to use different variable extraction and variable selection techniques over the past decades. Recently, $L_1$- and $L_0$-norm regularization are used to choose a subset of variables in the field of productivity and efficiency analysis. However, the existing literature on production models does not address the question of which regularization would be a better subset selection approach via, e.g., an effectiveness comparison.
We develop a new $L_0$-norm regularization approach to CQR/CER for subset variable selection. We show how to use mixed integer programming to solve the proposed $L_0$-norm regularization approach in practice and build a link to the commonly used $L_1$-norm regularization approach. A Monte Carlo study is performed to compare the finite sample performances of the proposed $L_0$-CQR/CER approaches with the $L_1$-norm approaches. We also evaluate the progress of SDGs for OECD countries and determine what causes inequality in sustainable development. Furthermore, to mitigate the computational burden, we adapt the cutting-plane algorithm to solve the $L_1$- and $L_0$-CQR/CER approaches. The results of the simulation and application illustrate that the proposed $L_0$-norm regularization approach can more effectively address the curse of dimensionality than the $L_1$-norm regularization approach in multidimensional spaces. The constraints of Afriat inequalities have been extensively reduced in the new algorithm, and the superiority in performance of new algorithm has also been verified.

3.2 Essay II

Essay II was originally motivated by pedagogical needs. In April 2020, I prepared computational tutorials for the online course Productivity and Efficiency Analysis (30E00300) at Aalto University School of Business. Professor Timo Kuosmanen suggested that an open-source Python package would lower the barrier for students and increase the popularity of convex regression and related methods. I started to develop Python codes for the course tutorials, and we subsequently published the first version of the pyStoNED package in GitHub. Professor Chia-Yen Lee and Mr. Yu-Hsueh Fang joined the initiative and contributed to the further development of the pyStoNED package. Essay II was written together as a detailed documentation to illustrate how to use this powerful package.

Convex regression and its extensions have bridged the gap between the traditional deterministic nonparametric and stochastic parametric approaches and are applied in various fields such as econometrics, operations research, and management science. In practice, however, convex regression and StoNED are computationally demanding approaches, requiring a user to solve a mathematical programming problem subject to a large number of linear constraints. For example, the additive CNLS model is a QP problem, whereas the multiplicative model requires solving an NLP problem. Therefore, most empirical applications published thus far have made use of commercial QP and NLP solvers, which can be coded using high-level mathematical computing languages such as GAMS or MATLAB. The lack of a comprehensive, powerful, reliable, and fully open access computational package for convex regression has slowed down its diffusion to the empirical practices.

pyStoNED, an open source Python package, was first introduced in April 2020 to lower the barrier to convex regression for practitioners. It has been proven to

1Recently, the R package, Benchmarking, provides a new function StoNED() to estimate the additive CNLS model.
be a user-friendly tool for the multivariate CNLS and StoNED methods. Its latest edition 0.5.5 includes modules for convex regression, convex quantile/expectile regression, isotonic regression, and graphical illustration. It also facilitates efficiency measurement using the conventional DEA and free disposable hull approaches. The pyStoNED package allows practitioners to estimate these models in an open access environment under a GPL-3.0 License.

Essay II presents a tutorial of the developed pyStoNED package, briefly reviews the alternative models supported, and illustrates the tutorials. We focus on the estimation of frontier cost and production functions, which currently forms the main application area of these techniques, emphasizing that the various modules of the pyStoNED package are directly applicable for (semi-)nonparametric regression analysis in any other contexts as well. To our knowledge, the pyStoNED package has been used in various research (including master thesis and research articles), showing its popularity and potential in our community.

### 3.3 Essay III

Essay III contributes to a series of papers on convex quantile/expectile regression approach to more accurate estimation of shadow prices and marginal abatement costs of bad outputs. When I joined the research team of Professor Timo Kuosmanen at Aalto University School of Business, Kuosmanen and Xun Zhou (doctoral student at that time) were together developing a methodology paper that was later published in European Journal of Operational Research (Kuosmanen & Zhou, 2021). I contributed to the follow-up paper Kuosmanen, Zhou & Dai (2020), where we applied the proposed approach to assess the economic cost of climate policy in the OECD counties. In Essay III we turn attention from ex-post evaluation towards expected future costs.

Carbon reduction target has become an essential policy in combating climate change. Various mandatory or voluntary abatement targets have been adopted worldwide. For example, China aims to decrease its carbon dioxide (CO₂) intensity by 40–45% from 2005 levels by 2020 and 60–65% by 2030. However, large scale actions on CO₂ emissions reduction would become a new economic burden due to the fact that economic growth is still tightly coupled with energy consumption in most countries. Moreover, implementing climate actions also needs additional high financial expenditure. Therefore, reliable cost estimation for a certain reduction target calls for a new empirical application.

Using the Work Plan for Controlling GHG Emissions During the 13th Five-Year Plan Period (2016–2020) as the benchmark, Essay III presents a forward-looking assessment of CO₂ abatement costs for Chinese provinces. Specifically, we first estimate the marginal abatement cost of CO₂ emissions for Chinese provinces from 1997 to 2015 by using the convex quantile regression approach. We further provide a forward-looking assessment of the abatement costs for Chinese provinces during the 13th Five-Year Plan Period based on their preassigned abatement targets.

The main advantage of convex quantile regression used in Essay III is that it employs
multiple quantiles without making any parametric distributional assumption, takes inefficiency explicitly into account, and is robust to noise and outliers. The empirical result shows that the estimated negative abatement costs for Chinese provinces indicate that they can benefit from the CO$_2$ emissions reduction targets. Accordingly, the magnitudes of economic benefits exhibit a significant regional disparity due to the different CO$_2$ emissions targets. Further policy recommendations are proposed for the government to pursue a more ambitious climate target.
References


References


