



西南财经大学

SOUTHWESTERN UNIVERSITY OF FINANCE AND ECONOMICS

# 2020 届 本科毕业论文（设计）

论文题目：面板数据线性均值回归与分位数回归模型  
参数估计方法的比较

学生姓名：张嘉伟

所在学院：经济数学学院

专    业：数学与应用数学（金融数学创新实验班）

学    号：41634005

指导教师：林谦

成    绩：

2020 年 6 月

# 西南财经大学

## 本科毕业论文原创性及知识产权声明

本人郑重声明：所呈交的毕业论文是本人在导师的指导下取得的成果，论文写作严格遵循学术规范。对本论文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。因本毕业论文引起的法律结果完全由本人承担。

本毕业论文成果归西南财经大学所有。

特此声明

毕业论文作者签名：

作者专业：

作者学号：

年 月 日

# 摘要

针对大多数研究者在实证分析中使用面板数据线性均值回归模型进行参数估计时所存在的问题，本文引入能提供更多关于数据结构方面的信息的分位数回归模型，并且对比分析了这两种模型的相关估计方法在估计的精确度与稳定性方面的差异，蒙特卡洛模拟结果显示分位数回归法是分析面板数据的有效手段，且在非正态误差条件下由于线性均值回归方法；本文还通过真实面板数据建模分析验证了相关的结论，这对于未来研究者在实证研究中选择恰当面板回归模型进行分析提供了参考。

**关键词：**面板数据；线性均值回归模型；分位数回归模型；固定效应；蒙特卡洛模拟比较

# Abstract

In view of the problems that most researchers use panel data linear mean regression model to estimate parameters in empirical analysis, this paper introduces the quantile regression model which can provide more information about data structure, and compares the differences between the two models in the accuracy and stability of estimation. The results of Monte Carlo simulation show that the quantile regression methods are effective in dealing with the panel data and do better than the linear mean regression methods when the error distribution is non-normal; This paper also verifies the relevant conclusions through the real panel data modeling analysis, which provides a reference for future researchers to select the appropriate panel regression model in empirical research.

**Key words:** Panel Data; Linear Mean Regression Model; Quantile Regression Model; Fixed Effects; Monte Carlo Simulation and Comparison

# 目录

1. 绪论.....	1
1.1 研究背景和意义.....	1
1.2 研究内容.....	2
1.3 论文的创新之处.....	3
1.4 论文的结构安排.....	3
2. 文献综述与研究现状.....	4
2.1 面板数据线性均值回归模型估计方法的文献综述.....	4
2.2 面板数据分位数回归模型估计方法的文献综述.....	5
3. 面板数据线性均值回归模型的估计方法.....	7
3.1 混合面板模型与其估计方法.....	7
3.1.1 混合面板模型.....	7
3.1.2 混合面板模型的估计方法.....	8
3.2 固定效应面板模型与其估计方法.....	8
3.2.1 固定效应面板模型.....	8
3.2.2 个体固定效应面板模型的估计方法.....	9
3.3 随机效应面板模型与其估计方法.....	11
3.3.1 随机效应面板模型.....	11
3.3.2 随机效应模型的估计方法.....	11
4. 面板数据分位数回归模型的估计方法.....	13
4.1 面板数据分位回归的基本模型.....	13
4.1.1 分位数与最优化原理.....	13
4.1.2 面板数据分位数回归模型.....	16
4.2 面板数据分位数回归模型的估计方法.....	17
4.2.1 一阶差分分位回归法.....	18
4.2.2 固定效应变换分位回归法.....	19
4.2.3 模式搜索分位回归法.....	20
最优化问题的求解—模式搜索方法.....	20
分段线性损失函数最小化问题—模式搜索法.....	20
5. 面板数据线性均值回归与分位数回归方法的比较.....	22
5.1 中位点模拟结果比较.....	23
5.1.1 正态误差情况下的比较.....	23
5.1.2 其他误差分布情况下的比较.....	25
5.2 关于面板数据线性均值回归与分位数回归方法的结论.....	26
6. 实证分析.....	27
6.1 实证设计.....	27
6.2 实证结果比较.....	28
7. 结论与展望.....	30
参考文献.....	32
致谢.....	34

# 1. 绪论

## 1.1 研究背景和意义

在计量经济学领域，面板数据(Panel Data)是非常重要的一类数据类型。早在 20 世纪 60 年代中期，美国 NLS(National Longitudinal Surveys of Labor Market Experience)就开始搜集不同年龄组的劳动力数据信息；另外一个著名的面板数据集是密歇根大学的 PSID(Panel Study of Income Dynamics)。之后关于面板数据模型的研究也逐渐兴起，因其是既有时间维度和横截面维度的二维数据，与传统的时间序列数据或者横截面数据相比，在挖掘更多有效的信息、避免出现多重共线性并且提高参数估计精确性方面具有诸多的优势，所以面板数据模型被越来越广泛的应用在经济问题的研究中。

在现有面板数据模型建模过程中多数假设变量之间的关系符合线性均值回归，对参数进行求解时大多采用最小二乘法或者广义最小二乘法。然而一些研究者发现传统的面板数据分析方法在应用中存在某些限制性，比如现有面板数据分析方法通常假设误差项服从正态分布，这其实是在大多数情形下都很难实现的，并且线性均值回归难以提供更多的关于数据全面的结构特征。Koenker(2004)最早展开了对面板数据分位数回归模型的研究，近年来，国内外关于面板数据分位数回归的探究逐渐深入，但是由于模型提出的时间较晚，目前关于这类模型的理论研究仍处于起步阶段，对于面板数据分位数回归模型的参数估计方法不唯一且较少，面板数据分位数回归模型的应用相对于线性均值回归模型而言较少。在各领域问题的实证分析中，大多数研究者仍然使用面板数据线性均值回归模型进行分析，他们依旧坚持使用传统的线性均值回归模型进行数据分析，这是否会使得研究的实证结果与真实值之间存在较大的偏差呢？在什么情况下研究者需要考虑面板分位数回归模型进行实证分析，而不是“固执”地采用传统面板数据分析方法呢？

针对上述问题，本文对于面板数据传统的线性均值回归模型与分位数回归模型之间的差异进行探究，并利用真实数据进行实证分析比较，有助于当今实证研究者针对不同的问题采用适当的模型进行分析，进而提高结果的准确性。

## 1.2 研究内容

本文在传统的面板数据线性均值回归模型与近年来兴起的分位数回归模型的基础之上，论述了适用于面板数据线性均值回归模型的最小二乘法、广义最小二乘法等，介绍了面板数据分位数回归模型中的分段线性损失函数，并根据解决最小化问题的思路将参数估计方法划分为两类，一类是利用消除模型中的固定效应进而实现参数的估计，如一阶差分分位回归法和固定效应变换分位回归法等，另一类是针对目标函数的特性，引入最优化原理中的模式搜索法求解参数估计。为了比较面板数据线性均值回归与分位数回归模型估计方法的精确度和稳定性，本文进行的相关蒙特卡洛模拟以及真实数据的实证分析对比，具体研究工作包括：

### ① 面板数据线性均值回归模型与分位数回归模型的组间比较

为了比较面板数据线性均值回归模型与分位数回归模型在参数估计的精确度与稳定性方面的差异，本文通过蒙特卡洛模拟计算出面板数据线性均值回归模型的三种最小二乘估计以及分位数回归模型的四种中位数估计的偏差 Bias 以及均方根误差 RMSE，并且比较了这七种参数估计结果在不同误差分布条件下无偏性与有效性的变化，并得出结论：在非正态误差条件下建立面板数据分位数回归模型有利于提高参数估计的精确度与稳定性。

### ② 面板数据线性均值回归与分位数回归方法的组内比较

本文除了对面板数据线性均值回归与分位数回归模型的比较外，还进一步探究了适用于面板数据线性均值回归模型的估计方法之间的差异，结果表明当模型中存在固定效应时，离差最小二乘法优于一阶差分最小二乘法，且都要优于混合最小二乘法；而通过比较面板数据分位数回归模型的估计方法之间的差异，本文发现：面板数据的固定效应变换分位回归法在 3 种误差分布条件下，其估计结果的稳定性和精确度都较高，而一阶差分分位回归法(FDQR)在误差出现偏态分布时效果较好，基于最优化原理的模式搜索分位回归法(HJQR)在随机误差服从正态分布和对称厚尾分布条件下的表现优于随机误差在偏态条件下表现，但不如面板数据的固定效应变换分位回归法。

### ③ 面板数据线性均值回归与分位数回归方法在实证分析中的应用

本文进一步应用真实面板数据来检验有关于面板数据线性均值回归于分位数

回归方法之间的比较结论，并对我国各地区城镇居民的消费支出与可支配收入之间的关系进行了实证分析，实证结果验证了本文的主要结论，并且还利用面板数据分位数回归方法进一步探究了微观经济学中“边际消费倾向递减”的现象。

### 1.3 论文的创新之处

本文的创新之处包括：

①针对大多数实证研究中选择面板数据回归模型所存在的问题，本文通过对比传统的面板数据线性均值回归模型与分位数回归模型之间的差异，结合蒙特卡洛模拟结果以及真实数据的实证结果为研究者在面板数据分析时选择实证研究模型提供了参考，这有助于提高实证结果与结论的可靠性与精确性。

②本文进行了参数估计方法的组内对比：对比了几种面板数据线性均值回归模型参数估计方法在精确性与稳定性方面的差异，并且考虑了不同误差分布情况下以及是否考虑模型中的固定效应等因素，得出了离差最小二乘估计方法优于一阶差分最小二乘法的结论；进一步，本文对比了几种面板数据线性均值回归模型参数估计方法在中位数回归时的差异，并探究了几种方法在不同分位点处回归的结果差异，并发现了固定效应变换分位回归法的精确性与稳定性。

③本文还进行了相同方法应用于不同模型时在参数估计的精确度与稳定性方面的对比：利用离差变换消除固定效应的方法应用于面板数据线性均值回归模型与面板数据分位数回归模型的效果相当，而利用一阶差分消除固定效应的方法应用于面板数据分位数回归模型的效果不如其应用于面板数据线性均值回归模型，且均不如离差变换法。

### 1.4 论文的结构安排

论文的研究共包括 7 个部分，各部分的主要内容如下：

第一章，绪论。介绍了本文的研究背景、研究意义，主要的研究内容和论文的创新之处。

第二章，文献综述与研究现状。对面板数据线性均值回归模型与分位数回归模型的提出过程、发展过程以及研究现状与应用情况进行了梳理与介绍，通过对研



究现状的分析得出启示，确定文本的研究内容与方向。

第三章，介绍面板数据线性均值回归模型的估计方法。概括了面板数据线性均值回归模型的类型，包括混合效应模型、固定效应模型和随机效应模型等，以及详细介绍了针对这三种线性均值回归模型的估计方法，包括常用的最小二乘法和广义最小二乘法等。

第四章，介绍面板数据分位数回归模型的估计方法。首先引出了面板数据分位数回归模型以及分段线性损失函数，其次详细介绍了两类参数估计的方法，一类是通过数据转换消除模型种固定效应的方法，另一类应用求解多维无约束最小化问题的方法。

第五章，面板数据线性均值回归与分位数回归方法的比较。利用蒙特卡洛模拟方法对面板数据线性均值回归与分位数回归方法进行比较，并且展示了在不同时期数和不同误差分布条件下估计结果的偏差 Bias 和均方根误差 RMSE。

第六章，真实数据分析。利用相关的真实面板数据并分别使用面板数据的线性均值回归和分位数回归方法估计未知参数，并比较了估计方法之间的差异。

第七章，结论与展望。总结论文研究的理论方法和主要的研究成果，提出论文种存在的不足与尚待解决的问题，对未来的相关研究进行展望。

## 2. 文献综述与研究现状

### 2.1 面板数据线性均值回归模型估计方法的文献综述

面板数据最早出现是在 1968 年，美国 NLS(National Longitudinal Surveys of Labor Market Experience)就开始为了研究美国的贫困特征进而开始搜集不同年龄组劳动力的面板数据信息，之后关于面板数据模型的研究也逐渐兴起，而现今面板数据模型已经成为经济金融领域分析中极为重要的组成部分。Trognon(2000)认为，面板数据模型日益受到研究者的重视以及广泛的运用，一方面是因为借助面板数据比利用单纯的横截面或者是时间序列数据能更好的回答各领域重点关注

的问题；另一方面是因为研究者计算能力的快速提升以及大数据时代下信息数据的可获取性不断提高。Arellano(2003)、Baltagi(2008)、Hsiao(2014)等人的著作对面板数据模型的相关理论与方法做了详细的介绍。

关于面板数据线性均值回归模型及估计方法，Robinson(1978)提出了面板数据随机系数回归模型，Mairesse 和 Griliches(1990)对于混合效应估计模型假设了解释变量对于被解释变量的影响与个体无关，并在该假设前提下利用混合最小二乘法进行了相关的面板数据分析，最终发现参数估计的结果是稳健的；借助于 Laplace(1818)提出的最小二乘方法(OLS)，他们针对不同的面板数据线性均值回归模型，并试图通过消除模型中个体固定效应进而得到解释变量回归系数的估计量，建立了面板数据线性回归模型的传统最小二乘估计方法、广义最小二乘估计方法等。Kao(1999)应用面板数据模型和传统最小二乘估计方法对贸易协定与贸易成员之间贸易的影响进行了相关的实证研究，随着之后面板数据的广泛应用，一些实证研究者逐渐发现传统的面板数据分析和模型估计方法在应用中存在某些局限性：一方面，现有面板数据分析方法通常假设误差项服从正态分布，如果数据分布违背假设，比如数据存在尖峰或者厚尾时，那么模型的估计结果可能不再优良和稳健；另一方面，现有面板数据模型建模过程中多数假设变量关系符合线性均值回归，对参数进行求解时多采用最小二乘法或者广义最小二乘法，其回归结果虽然可以估计或者检验均值效应，却难以刻画变量在数据尾部可能存在的相关关系，因此无法提供数据全面的结构特征。特别地，当不同的分布具有相同的均值和方差时，使用均值回归分析得到的结果无法体现其差异性，因此不足以提供有价值的提示，甚至可能会遗漏某些重要的信息。

## 2.2 面板数据分位数回归模型估计方法的文献综述

作为一种精确地描述解释变量对于被解释变量的变化范围以及条件分布影响的统计方法，分位数回归模型的概念最早由 Koenker 和 Bassett(1978)提出，借助于 Laplace(1818)提出的最小绝对残差估计思想，他们针对线性均值回归方法的某些缺陷，创建了线性分位数回归理论。Bassett(1986)、Powell(1986)和 Chernozhuko(2002)等人在此基础上进行了深入的研究，陆续解决了分位数回归的线性假设检验、异

方差的稳健性检验、估计量的一致性和线性规划解法等应用方面的难题，使得其成为了近几十年来发展迅速、应用广泛的回归模型方法。

面板数据分位数回归模型作为两种计量方法的有机结合，既能够深入挖掘面板数据大样本蕴含的大量信息，又能够有效发挥分位数回归的作用，是一致更为灵活的回归分析方法；在 He(2003)等人提出了中位数回归模型的基础上，陆续出现了对面板数据分位回归模型的研究，关于模型构建、参数估计、参数检验等各方面问题的国内外文献大量涌现。Koenker(2004)首次提出了基于固定效应面板数据模型的惩罚分位数回归模型，对于固定效应面板数据模型，在进行分位数回归时通过目标函数中加入  $L_1$  惩罚项来缩小个体效应，同时能够有效缩小参数估计量的方差，而且可以使得估计量保持渐进正态性。Wang 和 Shao(2009)研究了分位数回归模型的诱导光滑估计，借助于构造光滑函数得到参数的区间估计，并通过蒙特卡洛模拟得出：与线性均值回归模型相比，该方法的表现良好。Canay(2011)通过一种简单的数据转换剔除面板数据模型中的固定效应，从而可以方便地进行分位数回归模型的估计，并且该方法得到的参数估计量是一致的和渐进正态的。朱建平，朱万闯(2012)提出两阶段面板数据分位回归法，并将其与固定效应变换分位回归法进行比较，结果认为两种方法的中位数回归都是稳定的且拟合精度较高，但是两类方法的应用范畴不同。张元杰和田茂再(2015)针对固定效应面板数据模型，提出了两阶段 K 步差分分位回归方法，并推导了其大样本性质。也有文献转向研究模型求解算法与软件应用模块的结合。Baker(2016)指出在固定效应面板数据分位回归模型中，个体固定效应的存在会使得对变量估计系数的解释产生改变，因此在 Stata 环境下创建 qregpd 模块来估计面板分位回归模型。广义分位估计可以通过 Stata 模块 genqreg 实现，而面板数据分位数估计只是广义分位数估计的一个特殊例子。

近年来，面板数据分位数回归模型在经济领域有着非常重要的应用。Covas et al(2014)利用面板数据的分位数回归模型评估银行控股公司(BHCs)在预先设定的宏观经济状态下的资本短缺，发现模型预测的资本缺口远远高于利用线性均值回归模型得到的结果，这表明该方法有可能发现金融系统中出现的漏洞。Paniagua et al.(2015)利用双边 FDI 面板数据的分位数回归模型研究了个体企业对于 FDI 流动性

的影响。张晶,王淼晗等(2015)利用面板数据固定效应变换分位回归法,分析了我国城乡居民收入的主要作用因素,研究结果显示:在农村中人均受教育年限的收益回报更高,就业率在城乡中影响较为显著。张所地,范新英(2015)使用我国多个城市的面板数据,利用惩罚分位回归法分析了收入水平、利率水平等因素对于房价波动的影响,分析结果认为我国房价的上涨主要来源于居民收入水平提高拉动造成的,利率对房价上涨的影响并不显著,并且房价上涨越高,收入水平的拉动作用越明显,而其他因素的影响就越小。

### 3. 面板数据线性均值回归模型的估计方法

利用面板数据建立的线性均值回归模型通常包含混合面板模型、固定效应模型和随机效应模型,而对于这三种模型的参数估计,有着不同类型的最小二乘估计方法、广义最小二乘估计方法等,本章对面板数据线性均值回归模型以及主要的参数估计方法进行介绍。

#### 3.1 混合面板模型与其估计方法

##### 3.1.1 混合面板模型

如果一个面板数据模型定义为:

$$Y_{it} = \alpha + X_{it}'\beta + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

其中  $Y_{it}$  为模型的被解释变量,  $\alpha$  表示模型的截距项,  $X_{it}$  为  $K \times 1$  阶解释变量列向量(包括  $K$  个解释变量),  $\beta$  为  $K \times 1$  阶解释变量回归系数列向量(包括  $K$  个回归系数),  $\varepsilon_{it}$  为模型的误差项。即:

$$X_{it} = \begin{pmatrix} X_{it,1} \\ X_{it,2} \\ \vdots \\ X_{it,k} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}$$

则称此模型为混合面板模型,其特点是该模型与一般的回归模型无本质区别,

无论对于任何个体和截面，截距项  $\alpha$  与解释变量回归系数  $\beta$  都相同。即从时间上看，不同个体之间不存在显著性差异，从截面上看，不同截面之间也不存在显著性差异。

### 3.1.2 混合面板模型的估计方法

混合最小二乘估计方法是在时间（T）上与截面（N）上把  $N \times T$  个观测值混合在一起，并利用最小二乘法来估计模型的参数。

对于给定的混合面板数据模型：

$$Y_{it} = \alpha + X_{it}'\beta + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

若模型是正确设定的，并且模型的解释变量与误差项不相关： $\text{Cov}(X_{it}, \varepsilon_{it}) = 0$

那么可以使用混合最小二乘估计方法来估计参数  $\alpha$  与  $\beta$ ，并借助于求解极小化问题来寻找参数的混合最小二乘估计量  $\hat{\alpha}$  与  $\hat{\beta}$ ：

$$\min_{\alpha, \beta} \sum_{i=1}^N \sum_{t=1}^T \varepsilon_{it}^2 = \min_{\alpha, \beta} \sum_{i=1}^N \sum_{t=1}^T (Y_{it} - \alpha - X_{it}'\beta)^2$$

虽然对于混合面板数据模型的参数估计方法有许多（包括平均数最小二乘估计方法等），但本文在后续对于混合面板数据模型的最小二乘估计方法与分位数回归方法的差异比较中，我们优先采用混合最小二乘法进行参数估计。

## 3.2 固定效应面板模型与其估计方法

### 3.2.1 固定效应面板模型

固定效应模型分为三种类型，即个体固定效应模型、时点固定效应模型与个体时点双固定效应模型；本文主要利用个体固定效应面板模型进行线性均值回归与分位数回归方法间的对比，故只介绍有关个体固定效应模型的概述：

如果一个面板数据模型定义为：

$$Y_{it} = \alpha_i + X_{it}'\beta + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

如果截距项  $\alpha_i$  为随机变量，表示对于  $i$  个不同个体由  $i$  个不同的截距项，且其变化与模型的解释变量  $X_{it}$  相关； $X_{it}$  为  $K \times 1$  阶解释变量列向量（包括  $K$  个解释变

量),  $\beta$  为  $K \times 1$  阶解释变量回归系数列向量(包括  $K$  个回归系数),  $Y_{it}$  为模型的被解释变量,  $\varepsilon_{it}$  为模型的误差项, 则称该模型为个体固定效应模型。

个体固定效应模型的重要强假定条件是:

$$E(\varepsilon_{it} | \alpha_i, X_{it}) = 0, \quad i = 1, 2, \dots, N;$$

截距项  $\alpha_i$  作为随机变量描述不同样本个体建立的模型间的差异, 因而  $\alpha_i$  是不可观测的, 且与可观测的解释变量  $X_{it}$  的变化相联系, 所以称为个体固定效应模型; 个体固定效应面板模型的参数估计方法有许多, 但核心的思想是设法去除截距项  $\alpha_i$  对于参数估计的影响与干扰, 从而保证回归系数  $\beta$  估计量的一致性。

### 3.2.2 个体固定效应面板模型的估计方法

#### (1) 离差变换最小二乘估计方法(Within)

离差变换最小二乘估计方法的原理是先将面板数据中的每个个体的观测值变换为其平均数的离差观测值, 然后利用离差变换数据估计模型的参数。

具体步骤是, 对于下列个体固定效应面板模型:

$$Y_{it} = \alpha_i + X_{it}'\beta + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

首先对面板数据中的每个个体观测值求平均数, 并且建立下列模型:

$$\bar{Y}_i = \alpha_i + \bar{X}_i'\beta + \bar{\varepsilon}_i, \quad i = 1, 2, \dots, N;$$

其中:

$$\bar{Y}_i = \frac{1}{T} \sum_{t=1}^T Y_{it}, \quad \bar{X}_i = \frac{1}{T} \sum_{t=1}^T X_{it}, \quad \bar{\varepsilon}_i = \frac{1}{T} \sum_{t=1}^T \varepsilon_{it}$$

将上述两式相减, 消去截距项  $\alpha_i$ , 并得到下列回归方程等式:

$$Y_{it} - \bar{Y}_i = (X_{it} - \bar{X}_i)'\beta + (\varepsilon_{it} - \bar{\varepsilon}_i), \quad i = 1, 2, \dots, N;$$

称此模型为离差变换数据模型, 并对此应用普通最小二乘估计方法, 则相应的参数估计量被称作离差变换最小二乘估计量; 对于个体固定效应面板模型, 由于离差变换最小二乘法消去了截距项  $\alpha_i$ , 从而消除了其对于参数估计的影响与

干扰,从而保证回归系数 $\beta$ 的离差最小二乘估计量的一致性;如果随机误差项 $\varepsilon_{it}$ 还满足独立同分布条件,则回归系数 $\beta$ 的离差最小二乘估计量不但具有一致性,还具有有效性,所以对于个体固定效应面板模型的参数估计通常采用离差变换最小二乘法,而时点固定效应模型与个体时点固定效应模型不能采用离差变换最小二乘法来估计模型的参数。

如果对于固定效应截距项 $\alpha_i$ 感兴趣,离差最小二乘估计方法也提供了对于截距项的估计,其中 $\hat{\beta}$ 为回归系数 $\beta$ 的离差最小二乘估计量:

$$\alpha_i = \bar{Y}_i - \bar{X}_i' \hat{\beta}, \quad i = 1, 2, \dots, N;$$

## (2) 一阶差分最小二乘估计方法(First Difference)

在短期面板的条件下,一阶差分最小二乘估计方法是对个体固定效应模型中的解释变量与被解释变量的差分变量构成的模型的参数进行最小二乘估计。

对于下列个体固定效应面板模型:

$$Y_{it} = \alpha_i + X_{it}' \beta + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

其滞后一期的关系表达式为:

$$Y_{i,t-1} = \alpha_i + X_{i,t-1}' \beta + \varepsilon_{i,t-1}, \quad i = 1, 2, \dots, N; \quad t = 2, 3, \dots, T$$

上述两式相减,消去模型的截距项 $\alpha_i$ ,得到一阶差分模型:

$$Y_i - Y_{i,t-1} = (X_i - X_{i,t-1})' \beta + (\varepsilon_{it} - \varepsilon_{i,t-1}), \quad i = 1, 2, \dots, N; \quad t = 2, 3, \dots, T$$

对该一阶差分模型应用最小二乘估计方法得到的 $\beta$ 的估计量称为一阶差分最小二乘估计量,并且是一致估计量。

相比于离差最小二乘估计方法,一阶差分最小二乘估计方法不能估计个体固定效应面板模型的截距项 $\alpha_i$ ,并且当随机误差项 $\varepsilon_{it}$ 满足独立同分布条件时,回归系数 $\beta$ 的一阶差分最小二乘估计量 $\hat{\beta}_2$ 不如离差最小二乘估计量 $\hat{\beta}_1$ 有效,即:

$$Var(\hat{\beta}_1) < Var(\hat{\beta}_2)$$

所以在后续对于个体固定效应面板数据模型的最小二乘估计方法与分位数回

归方法的差异比较中，我们优先采用离差最小二乘估计方法进行参数估计。

### 3.3 随机效应面板模型与其估计方法

#### 3.3.1 随机效应面板模型

对于一个面板数据模型：

$$Y_{it} = \alpha_i + X_{it}'\beta + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

如果截距项  $\alpha_i$  为随机变量，其分布状况与解释变量  $X_{it}$  无关； $X_{it}$  为  $K \times 1$  阶解释变量列向量(包括  $K$  个解释变量)， $\beta$  为  $K \times 1$  阶解释变量回归系数列向量(包括  $K$  个回归系数)， $\varepsilon_{it}$  为模型的误差项，则称该模型为个体随机效应面板模型，或者随机截距模型、随机分量模型等。该模型的假定条件为截距项和误差项独立同分布，但是并未限定具体的分布，即：

$$\alpha_i \sim i.i.d(\alpha, \sigma_\alpha^2), \quad \varepsilon_{it} \sim i.i.d(0, \sigma_\varepsilon^2)$$

同理也可以定义时点随机效应面板模型和个体时点随机效应面板模型，但是在实证研究中个体随机效应面板模型最为常见；对于个体随机效应面板模型， $E(\alpha_i | X_{it}) = E(\alpha_i) = \alpha$ ， $E(Y_{it} | X_{it}) = \alpha + X_{it}'\beta$ ，被解释变量  $Y_{it}$  在解释变量给定时可以识别，所以当采用混合最小二乘法估计参数时，随机效应模型参数的混合 OLS 估计量具有一致性，但是它忽略了随机误差项的结构信息，因而不具有有效性。

#### 3.3.2 随机效应模型的估计方法

##### (1) 平均数最小二乘估计方法(Between)

平均数最小二乘估计方法的步骤是首先对面板数据的每个个体求平均数，共得到  $N$  个平均数估计值，然后利用  $X_{it}$  与  $Y_{it}$  的  $N$  组观测值并采用最小二乘方法来估计模型的参数。

以下列随机效应面板模型为例：

$$Y_{it} = \alpha_i + X_{it}'\beta + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

首先对面板数据中的每个个体观测值求平均数，并且建立下列模型：



$$\bar{Y}_i = \alpha_i + \bar{X}_i' \beta + \bar{\varepsilon}_i, \quad i = 1, 2, \dots, N;$$

其中：

$$\bar{Y}_i = \frac{1}{T} \sum_{t=1}^T Y_{it}, \quad \bar{X}_i = \frac{1}{T} \sum_{t=1}^T X_{it}, \quad \bar{\varepsilon}_i = \frac{1}{T} \sum_{t=1}^T \varepsilon_{it}$$

对上述回归方程进行变换，得到：

$$\bar{Y}_i = \alpha + \bar{X}_i' \beta + (\alpha_i - \alpha + \bar{\varepsilon}_i), \quad i = 1, 2, \dots, N;$$

称此回归方程模型为平均数模型，并对此应用普通最小二乘估计方法，则相应的参数估计量被称作平均数最小二乘估计量。

如果随机变量  $\bar{X}_i$  与  $(\alpha_i - \alpha + \bar{\varepsilon}_i)$ ,  $i = 1, 2, \dots, N$  相互独立，则模型中截距项与解释变量回归系数的平均数最小二乘估计量是一致估计量，由此看出平均数最小二乘估计方法适用于混合短期面板数据模型与个体随机效应模型，对于个体固定效应模型来说，由于截距项  $\alpha_i$  与模型的解释变量  $X_{it}$  相关，也即  $\alpha_i$  与随机变量  $\bar{X}_i$  相关，则  $\bar{X}_i$  与  $(\alpha_i - \alpha + \bar{\varepsilon}_i)$ ,  $i = 1, 2, \dots, N$  不是相关独立的，所以利用平均数最小二乘法得到的平均数最小二乘估计量是非一致估计量。

## (2) 可行广义最小二乘估计方法

对于随机效应面板模型：

$$Y_{it} = \alpha_i + X_{it}' \beta + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

若截距项  $\alpha_i$  与随机误差项  $\varepsilon_{it}$  独立同分布，则对上述回归方程作如下变换：

$$Y_{it} - \hat{\lambda} \bar{Y}_i = (1 - \hat{\lambda}) \cdot \mu + (X_{it} - \hat{\lambda} \bar{X}_i)' \beta + v_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

其中：

$$\bar{Y}_i = \frac{1}{T} \sum_{t=1}^T Y_{it}, \quad \bar{X}_i = \frac{1}{T} \sum_{t=1}^T X_{it}$$

$$\hat{\lambda} = 1 - \frac{\sigma_{\varepsilon}}{\sqrt{\sigma_{\varepsilon}^2 + T\sigma_{\alpha}^2}}$$

$v_{it} = (1 - \hat{\lambda})\alpha_i + (\varepsilon_{it} - \hat{\lambda}\bar{\varepsilon}_i)$  渐进服从独立同分布，对该变换后的回归方程应用普通最小二乘估计法，所得到的估计量称为随机效应估计量或者可行广义最小二乘估计量；对于随机效应面板模型，可行广义最小二乘估计量不但是一致估计量，而且是有效估计量，但是对于混合面板模型于个体固定效应模型，可行广义最小二乘估计量不是一致估计量，不能利用此方法来估计对应模型的参数。

(1) 当  $\hat{\lambda} = 1 - \frac{\sigma_\varepsilon}{\sqrt{\sigma_\varepsilon^2 + T\sigma_\alpha^2}} = 0$  时，即  $\sigma_\alpha = 0$ ，模型的截距项为常量时，变换的

回归方程变为：

$$Y_{it} = \mu + X_{it}'\beta + (\alpha_i + \varepsilon_{it}), \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

可以看出此方法等价于适用于混合面板模型的混合最小二乘估计。

(2) 当  $\hat{\lambda} = 1 - \frac{\sigma_\varepsilon}{\sqrt{\sigma_\varepsilon^2 + T\sigma_\alpha^2}} = 1$ ，即  $\sigma_\varepsilon = 0$ ，模型的误差项为常量时，变换的

回归方程变为：

$$Y_{it} - \bar{Y}_i = (X_{it} - \bar{X}_i)'\beta + (\varepsilon_{it} - \bar{\varepsilon}_i), \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

可以看出此方法等价于个体固定效应面板模型的离差变换最小二乘估计方法。

## 4. 面板数据分位数回归模型的估计方法

### 4.1 面板数据分位回归的基本模型

#### 4.1.1 分位数与最优化原理

定义：假设  $Y$  为一连续的随机变量，其分布函数为  $CDF: F_Y(y) = P(Y \leq y)$ ，那么对于任意的  $0 < \tau < 1$ ， $\tau$  分位数定义为：

$$q_\tau = \inf\{y: F_Y(y) \geq \tau\}$$

当  $\tau = 1/2$  时，即为最常见的中位数。根据分位数的定义能清楚了解随机变量  $Y$  的数值分布情况，即在变量  $Y$  的所有取值中，在分位数  $q_\tau$  之下的取值所占比例为  $\tau$ ，同时在分位数  $q_\tau$  之上的取值所占比例为  $(1-\tau)$ 。

分位数可视为某些最小化问题的最优解，为理解分位数与最优化之间的关系，我们首先考虑随机变量  $Y$  的均值  $\mu$  与最小化的关系：

假设我们通过平方差  $(Y-\mu)^2$  来测量随机变量  $Y$  某一观测值与均值  $\mu$  的距离，通过其期望值  $E(Y-\mu)^2$  来计算  $Y$  与  $\mu$  的平均方差，并且当随机变量  $Y$  的均值  $\mu = EY$  时使得均方差  $E(Y-\mu)^2$  取得最小值，证明如下：

$$\begin{aligned} E(Y-\mu)^2 &= EY^2 - 2E(Y) \cdot \mu + \mu^2 \\ &= [EY^2 - E^2(Y)] + [E^2(Y) - 2E(Y) \cdot \mu + \mu^2] \\ &= \text{Var}(Y) + (EY - \mu)^2 \geq \text{Var}(Y) \end{aligned}$$

因为上述表达式中的  $\text{Var}(Y)$  是固定不变的，所以通过最小化  $(EY - \mu)^2$  可使得均方差  $E(Y-\mu)^2$  取得最小值  $\text{Var}(Y)$ （当且仅当  $\mu = EY$ ）。

接下来我们考虑随机变量  $Y$  的中位数  $m$ 。假设根据绝对距离  $|Y-m|$  来衡量  $Y$  的某一观测值与常数  $m$  的离差，并且通过平均绝对距离  $E|Y-m|$  来衡量随机变量  $Y$  偏离  $m$  的程度，同样在最小化  $E|Y-m|$  的情况下，求得  $m$  得值。

假设随机变量  $Y$  为连续随机变量，其概率密度函数  $PDF: f(y)$ ，那么：

$$\begin{aligned} E|Y-m| &= \int_{-\infty}^{+\infty} |y-m| \cdot f(y) dy \\ &= \int_{-\infty}^m (m-y) f(y) dy + \int_m^{+\infty} (y-m) f(y) dy \end{aligned}$$

为了求解  $E|Y-m|$  得最小值，我们借助于导函数等于零求解  $m$  的值，对上述表达式中的第一项求导如下：

$$\frac{d}{dm} \int_{-\infty}^m (m-y) f(y) dy$$

$$\begin{aligned}
&= \frac{d}{dm} (m \int_{-\infty}^m f(y) dy - \int_{-\infty}^m y f(y) dy) \\
&= F(m) + m f(m) - m f(m) = F(m)
\end{aligned}$$

同理对第二项求导可得：

$$\begin{aligned}
&\frac{d}{dm} \int_m^{+\infty} (y-m) f(y) dy \\
&= -m f(m) - \frac{d}{dm} (m \int_m^{+\infty} f(y) dy) \\
&= -m f(m) - [1 - F(m)] + m f(m) = F(m) - 1
\end{aligned}$$

合并上述两部分，并令导函数为零，则有：

$$\frac{d}{dm} E|Y-m| = 2F(m) - 1 = 0$$

解得  $F(m) = 1/2$ ，即当  $m$  为中位数时，离差  $E|Y-m|$  取得最小值。

最后，我们考虑随机变量  $Y$  的分位数  $q_\tau$ 。对于任意的  $\tau \in (0,1)$ ，随机变量  $Y$  到分位数  $q$  的距离用加权绝对距离来衡量：根据  $Y$  的观测值是位于分位数  $q$  的左边还是右边赋予不同的权重：

$$d(y, q) = \begin{cases} (1-\tau) \cdot |y-q|, & y < q \\ \tau \cdot |y-q|, & y > q \end{cases}$$

可以证明当分位数  $q$  是第  $\tau$  分位数时，即  $q = q_\tau$  时，可使得  $Ed(y, q)$  取得最小值：

$$Ed(y, q) = (1-\tau) \int_{-\infty}^q (q-y) f(y) dy + \tau \int_q^{+\infty} (y-q) f(y) dy$$

为了求解使得上述表达式最小化的  $q$  值，可利用上式关于  $q$  的导函数：

$$\begin{aligned}
\frac{d}{dq} Ed(y, q) &= (1-\tau) F(q) - \tau [1 - F(q)] \\
&= F(q) - \tau
\end{aligned}$$

显然当  $q$  满足  $F(q) - \tau = 0$  时，可以使得  $Ed(y, q)$  取到最小值，因此，满足集合：

$\{q : F(q) = \tau\}$  的任一元素都能使得  $Ed(y, q)$  最小化；并且如果解是唯一的，那么

$q = F^{-1}(\tau)$ ，如果没有唯一解，那么可以得到一个“ $\tau$ 分位数区间”，选择其中最小的元素作为 $q$ 的值即可，此时分位数函数依然满足单调不减性。

若记损失函数为如下分段线性函数：

$$\rho_{\tau}(\mu) = \mu(\tau - I(u < 0))$$

其中， $I(\cdot)$ 代表指示函数，当括号里的不等式成立时取值为1，否则为0，则有：

$$\rho_{\tau}(y - q) = \begin{cases} (y - q)(\tau - 1), & y < q \\ (y - q) \cdot \tau, & y > q \end{cases}$$

通过上述关于最优化的过程描述可知：当 $q$ 为第 $\tau$ 分位数时，能够使得 $E\rho_{\tau}(y - q)$ 取得最小值。

#### 4.1.2 面板数据分位数回归模型

假设有面板数据模型：

$$Y_{it} = \alpha_i + X_{it}'\beta + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

其中， $X_{it}$ 为 $K \times 1$ 阶解释变量列向量(包括 $K$ 个解释变量)， $\beta$ 为 $K \times 1$ 阶解释变量回归系数列向量(包括 $K$ 个回归系数)，在该模型中假设 $\alpha_i$ 为截面固定效应，表示未被包括在回归模型中的但是跟个体变化相关的难以预测的因素。对于任意的分位点 $\tau$ ，可建立被解释变量 $Y_{it}$ 对应的条件分位数函数：

$$Q_{y_{it}}(\tau | X_{it}) = \alpha_i(\tau) + X_{it}'\beta(\tau), \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

该模型中需要估计的参数为 $\alpha_i(\tau)$ 和 $\beta(\tau)$ ，如果观测值足够多，我们可以对每个个体在 $\tau$ 分位数下的固定效应 $\alpha_i(\tau)$ 进行估计，但是当模型涉及的个体截面数量较多时，所需估计的 $\alpha_i(\tau)$ 的数目会非常多，在实际应用中，样本往往存在个体数较多而时期数较少的情况，此时对 $\alpha_i(\tau)$ 进行估计容易导致“维数灾难”问题，可能无法得到未知参数的估计值，即使能够得到参数的估计结果，但是该结果往往不精确。

为了减少待估参数的个数，假设个体固定效应  $\alpha_i$  的值与分位数  $\tau$  无关，仅有模型的解释变量  $X_{it}$  对被解释变量产生的影响即回归系数  $\beta(\tau)$  随分位点变化而变化，并将上述条件分位数函数修改为下列的条件分位数函数：

$$Q_{y_{it}}(\tau | X_{it}) = \alpha_i + X_{it}' \beta(\tau), \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

对上述模型的参数估计，可通过求解如下加权绝对距离之和最小化问题来实现：

$$\min_{(\alpha, \beta)} \sum_{k=1}^q \sum_{i=1}^N \sum_{t=1}^T w_k \cdot \rho_{\tau_k}(Y_{it} - \alpha_i - X_{it}' \beta(\tau_k))$$

其中， $\rho_{\tau}(\mu) = \mu(\tau - I(u < 0))$  是分段线性分位数损失函数（Koenker 和 Bassett, 1978）。 $I(\cdot)$  代表指示函数，当括号里的不等式成立时取值为 1，否则为 0。即：

$$\rho_{\tau}(\mu) = \begin{cases} \mu \cdot (\tau - 1), & \mu < 0 \\ \mu \cdot \tau, & \mu \geq 0 \end{cases}$$

而  $w_k$  是权重系数，控制  $q$  个分位数  $\tau_1, \tau_2, \dots, \tau_q$  在估计参数时的相对影响程度，可以将所有的权重系数设置为相等的值，也可以参考 Mosteller(1946)提出的离散 L-统计量加权方法来进行设置。通过求解上式可以同时得到所有截面个体的固定效应  $\alpha_i$  以及任意分位数下的模型解释变量回归系数  $\beta(\tau)$ 。

上述最小化问题的求解时模型参数估计的关键所在，目前主要存在两类求解方法：一类方法侧重于估计自变量系数  $\beta(\tau)$  而忽略个体固定效应  $\alpha_i$ ，其基本思想是通过对模型中的解释变量与被解释变量的转换（例如差分转换，均值转换等），消除个体固定效应  $\alpha_i$ ，然后利用最小二乘法估计每个分位点下的系数  $\beta(\tau)$ ；另一类方法则是从最小化问题整体出发，该最小化问题没有解析解，通过数值迭代算法可求其数值解。

## 4.2 面板数据分位数回归模型的估计方法

Koenker(2004)提出了固定效应面板分位回归模型，并利用加入惩罚项的方法实

现参数估计，之后陆续有研究者对这类模型的参数估计问题进行探究，提出了一阶差分分位回归法、固定效应变换分位回归法等。其中，一阶差分分位回归法、固定效应变换分位回归法均是通过数据转换消除模型中的个体固定效应，进而利用求解给定分位点条件下的极小化问题来估计模型的参数。这类方法操作起来较为简单，但是都忽略了对个体固定效应的估计。而惩罚分位回归法根据模型本身的特点进行求解，使用内点法进行计算未知参数，此方法可以得到所以参数的估计(包括个体固定效应)，但实现起来较为困难，惩罚项中调节参数的选择也增加了模型的复杂性，所以本文忽略惩罚分位回归法与其他参数估计方法的比较。

#### 4.2.1 一阶差分分位回归法

对于下列个体固定效应模型：

$$Y_{it} = \alpha_i + X_{it}'\beta + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

由于相同的个体其固定效应相同，所以对于特定的个体  $i$ ，我们可以取相邻期的差分进而消去个体固定效应  $\alpha_i$ ，即：

$$\begin{aligned} \Delta Y_{i,1} &= Y_{i,2} - Y_{i,1} \\ \Delta Y_{i,2} &= Y_{i,3} - Y_{i,2} \\ &\vdots \\ \Delta Y_{i,t-1} &= Y_{i,t} - Y_{i,t-1} \end{aligned}$$

从而有：

$$Y_{it} = \alpha_i + X_{it}'\beta + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

$$Y_{i,t-1} = \alpha_i + X_{i,t-1}'\beta + \varepsilon_{i,t-1}, \quad i = 1, 2, \dots, N; \quad t = 2, 3, \dots, T$$

两式相减得到：

$$(Y_{it} - Y_{i,t-1}) = (X_{it} - X_{i,t-1})'\beta + (\varepsilon_{it} - \varepsilon_{i,t-1}), \quad i = 1, 2, \dots, N; \quad t = 2, 3, \dots, T$$

$$\Delta Y_{it} = \Delta X_{it}'\beta + \Delta \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T-1$$

经过以上变换，虽然样本数据减少了  $N$  个，只剩下  $N(T-1)$  个样本数据，但我们需要估计的参数只有解释变量回归系数  $\beta$ ，这对于实践中往往截面个体数量  $N$

很大, 而时期数  $T$  很小也可以给出回归系数  $\beta$  较为精确有效的估计, 于是考虑  $\Delta Y_{it}$  的条件分位数函数:

$$Q_{\Delta Y_{it}}(\tau | \Delta X_{it}) = \Delta X_{it}' \beta(\tau), \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T-1$$

通过求解极小化问题:  $\min_{\beta} \sum_{i=1}^N \sum_{t=1}^{T-1} \rho_{\tau}(\Delta Y_{it} - \Delta X_{it}' \beta)$  可获得解释变量回归系数  $\beta$

的  $\tau$  分位点估计, 并称此估计方法为一阶差分分位回归估计 FDQR(First-Differenced Quantile Regression Estimator)。

#### 4.2.2 固定效应变换分位回归法

另一种消除个体固定效应模型的方法是固定效应变换分位回归法, 对于个体固定效应模型:

$$Y_{it} = \alpha_i + X_{it}' \beta + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

现在对每个个体  $i$  求其在时期上的平均数, 并得到:

$$\bar{Y}_i = \alpha_i + \bar{X}_i' \beta + \bar{\varepsilon}_i, \quad i = 1, 2, \dots, N;$$

其中:

$$\bar{Y}_i = \frac{1}{T} \sum_{t=1}^T Y_{it}, \quad \bar{X}_i = \frac{1}{T} \sum_{t=1}^T X_{it}, \quad \bar{\varepsilon}_i = \frac{1}{T} \sum_{t=1}^T \varepsilon_{it}$$

因为对于每个个体  $i$  其固定效应  $\alpha_i$  在不同时期内固定不变, 故对于每一时期  $t$  利用上述两个回归方程相减, 得到:

$$Y_{it} - \bar{Y}_i = (X_{it} - \bar{X}_i)' \beta + (\varepsilon_{it} - \bar{\varepsilon}_i), \quad t = 1, 2, \dots, T$$

记:  $Y'_{it} = Y_{it} - \bar{Y}_i$ ,  $X'_{it} = (X_{it} - \bar{X}_i)$ ,  $\varepsilon'_{it} = (\varepsilon_{it} - \bar{\varepsilon}_i)$ , 则有:

$$Y'_{it} = X'_{it}' \beta + \varepsilon'_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

于是考虑  $Y'_{it}$  的条件分位数函数:

$$Q_{Y'_{it}}(\tau | X'_{it}) = X'_{it}' \beta(\tau), \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$



通过求解极小化问题： $\min_{\beta} \sum_{i=1}^N \sum_{t=1}^T \rho_{\tau}(Y'_{it} - X'_{it}\beta)$  可获得解释变量回归系数  $\beta$  的分位点估计，并称此估计方法为固定效应分位回归估计法 FETQR(Fixed Effects Transformation Quantile Regression Estimator)。

#### 4.2.3 模式搜索分位回归法

##### 最优化问题的求解--模式搜索方法

Hooke 和 Jeeves 在 1961 年提出了原始的模式搜索算法，1997 年由 Torczon 将其加以推广，引入模式矩阵的概念(包括基本矩阵和生成矩阵)，提出了 GPS 算法，并证明了它关于无约束最优化问题的收敛性。近年来，模式搜索算法与其他方法相结合，一定程度上克服了其在接近稳定点时不成功的迭代增加从而使搜索进程变慢的缺陷。例如，在模式搜索中引入单纯性导数，根据他的信息对搜索方向进行编号；利用拟牛顿法中近似计算的 Hessian 矩阵作为模式搜索算法中的一组方向集，以及蚁群算法在模式搜索中的应用。

模式搜索法是一种解决最优化问题的直接搜索算法<sup>1</sup>，在求解最优化问题中，当目标函数较为复杂或者不能利用变量显函数描述时，可采用直接搜索的方法并经过若干次迭代搜索到最优点，而模式搜索法是直接搜索法其中的一类方法，其是对当前搜索点按固定模式和步长探索移动，以寻求可行下降方向的直接搜索法，迭代过程只要找到相对于当前点的改善点，则步长递增，并从该点开始进入下一次迭代；否则步长递减，在当前点继续搜索。模式搜索法的基本思想是从初始基点开收，交替实施两种搜索：轴向搜索和模式搜索。轴向搜索依次沿  $n$  个坐标轴的方向进行，用来确定新的基点和有利于函数值下降的方向；模式搜索则是沿着相邻两个基点的连线方向进行，试图使得函数值下降更快，两种移动模式交替运行，沿着函数值下降最快的方向搜索最优解向量。

##### 分段线性损失函数最小化问题--模式搜索法

面板分位数回归参数估计的关键是对分段线性损失函数最小化问题的求解，考虑到其本身是一个多维无约束的最优化问题，并且目标函数为分段函数而无法求导，因此可以选择多维无约束优化理论的直接法--“模式搜索法”来进行求解，该

<sup>1</sup> 基于启发式方法的只利用目标函数值信息的无约束优化方法，简称直接法，常见的直接法有坐标轮换法、鲍威尔法、模式搜索法等。

方法通过数值迭代计算未知参数的数值解，不需要在模型中添加惩罚项，从而使模型的计算过程相对简单，并且可以同时得到个体固定效应和解释变量回归系数的估计值。下面对该方法的具体步骤进行介绍：

Step1 给定参数  $\alpha_i$  和  $\beta(\tau)$  的初始值，构成  $n$  维初始解向量  $x^0$ ；给定初始步长向量  $\delta^0 = (\delta_1^0, \delta_2^0, \dots, \delta_n^0)^T > 0$ ，加速系数  $\gamma > 0$ ，收缩系数  $\theta \in (0,1)$  以及精度  $\varepsilon > 0$ ；令  $k=0$ 。

Step2 令  $y = x^k$ ，记分段线性函数极小值问题中的目标函数为  $f(x)$ ；

Step3 从  $y$  出发，依次作平行于单位矢量  $e^j (j=1,2,\dots,n)$  的轴向搜索；

①正向搜索：若  $f(y + \delta_j^k e^j) < f(y)$ ，则令  $y = y + \delta_j^k e^j$ ，否则做负向搜索

②负向搜索：若  $f(y - \delta_j^k e^j) < f(y)$ ，则令  $y = y - \delta_j^k e^j$ ，否则令  $y=y$

Step4 令  $x^{k+1} = y$ ，若  $f(x^{k+1}) < f(x^k)$ ，则对  $x^{k+1}$  沿着加速方向  $l^k = x^{k+1} - x^k$  作模式搜索，令  $y = x^{k+1} + \gamma l^k$ ， $\delta^{k+1} = \delta^k$ ， $k = k+1$ ，转 Step3，否则转 step5；

Step5 若  $|\delta^k| < \varepsilon$ ，则停止迭代，输出  $x^k$ ，否则当  $x^{k+1} \neq x^k$  时，令  $y = x^{k+1}$ ， $\delta^{k+1} = \delta^k$ ， $k = k+1$ ，转 Step3；当  $x^{k+1} = x^k$  时，令  $y = x^{k+1}$ ， $\delta^{k+1} = \theta \delta^k$ ， $k = k+1$ ，转到 Step3。

从以上模式搜索算法过程可看出，只要给定参数  $\alpha_i$  和  $\beta(\tau)$  的初始值，通过不断的轴向搜索与模式搜索，寻找到最小的函数值，就可以得到参数  $\alpha_i$  和  $\beta(\tau)$  的最优解，即得到参数的模式搜索分位回归估计量。关于初始点的选择问题，在模拟过程中通过反复试验发现，选择任意初始点计算结果都相差不大，也就是说模式搜索分位回归法对于参数初始值的设定没有特殊要求。

## 5. 面板数据线性均值回归与分位数回归方法的比较

由于在大多数研究中我们主要关心的参数为解释变量回归系数 $\beta$ ，所以在本节的蒙特卡洛模拟中，本文仅对各个分位点处解释变量回归系数估计结果进行比较，首先比较在中位点出的估计值，然后对其他分位点处的估计值进行对比；为了比较面板数据线性均值回归模型与分位数回归模型的估计方法之间的差异，本文将固定效应面板数据模型、混合面板数据模型和混合分位数回归模型三种模型在不同估计方法下的估计结果一同进行比较，这样用于比较的估计方法共有以下 7 种：

① 混合面板数据模型估计：完全忽略模型中的个体固定效应，直接使用混合数据的混合最小二乘估计，并记相应的估计量为 PLS 估计；

② 混合面板数据分位数回归估计：完全忽略个体固定效应，直接使用混合数据的分位数回归估计，并记相应的估计量为 PQR 估计；

③ 固定效应面板数据线性均值回归模型的估计：包含个体固定效应，并使用离差最小二乘估计，并记相应的估计量为 FE\_Within；

④ 固定效应面板数据线性均值回归模型的估计：包含个体固定效应，并使用一阶差分最小二乘估计，并记相应的估计量为 FE\_FD

⑤ 固定效应面板数据分位数回归模型的估计：考虑个体固定效应变换分位回归估计，并记相应的估计量为 FETQR

⑥ 固定效应面板数据分位数回归模型的估计：考虑个体固定效应一阶差分分位回归估计，并记相应的估计量为 FDQR

⑦ 固定效应面板数据分位数回归模型的估计：运用最优化原理中的模式搜索法求解分段线性损失函数取得最小值时回归系数的估计，并记相应的估计量为 HJQR

根据含固定效应面板数据结构产生模拟数据的二变量面板数据模型为：

$$y_{it} = \alpha_i + \beta \cdot x_{it} + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

其中  $x_{it} = \gamma_i + v_{it}$ ,  $\gamma_i \sim N(0, \sigma_1^2)$ ,  $v_{it} \sim N(0, \sigma_2^2)$ , 且  $\gamma_i$  和  $v_{it}$  是相互独立的。不失一般

性，本文在模拟中设定相关的初始值： $\beta = 2, \sigma_1 = 1, \sigma_1 = 1, a = 0.5, \alpha_i = a + \gamma_i$ ，即横截面共同的截距为 0.5，每个个体有不同的个体效应 $\gamma_i$ ，需要说明的是这些参数的设定值是可以任意调整的，这将不会影响到本文对于不同估计方法进行模拟的比较结果。对于随机误差项 $\varepsilon_{it}$ ，经典假设认为其服从标准正态分布 $\varepsilon_{it} \sim N(0,1)$ ，但是当模型中的随机误差项违背经典假设时，也选取了两种情况进行估计并对结果进行了对比分析，因此总共讨论随机误差项服从三种不同分布时的三种情况：标准正态分布 $\varepsilon_{it} \sim N(0,1)$ ，对称厚尾分布（t 分布） $\varepsilon_{it} \sim t(3)$ 以及偏态分布（卡方分布） $\varepsilon_{it} \sim \chi^2(3)$ 。模拟过程中，考虑到我国经济研究中常用的面板数据大多是以省市地区作为横截面的，所以在模拟中本文固定截面个数为  $N=30$ ，而时期数  $T$  则分别考虑取较短和较长期数为 5 和 15。

蒙特卡洛模拟过程中共生成 50 组面板数据，在以上 7 中不同的参数估计方法下逐次进行线性均值回归或者分位数回归并得到 50 个回归系数估计值 $\hat{\beta}_i, (i=1,2,\dots,50)$ ，并在不同的误差分布影响下分别计算 7 种估计方法对于回归系数 $\beta$ 的偏差 Bias 和均方根误差 RMSE，具体定义如下：

$$\text{偏差 Bias} = \frac{1}{n} \sum_{i=1}^n (\hat{\beta}_i - \beta)$$

$$\text{均方根误差 RMSE} = \frac{1}{n} \sqrt{\sum_{i=1}^n (\hat{\beta}_i - \beta)^2}$$

其中， $n$  为模拟重复次数(50)，回归系数真实值 $\beta = 2$ ，偏差 Bias 主要用于衡量估计量的无偏性，RMSE 用于衡量估计量的稳定性，其绝对值越小说明估计值越接近真实值，为了便于观察，本文在后续模拟结果的展示中将同样条件下 7 中估计中偏差 Bias 和均方根误差 RMSE 绝对值最小的用星号\*标记出来。

## 5.1 中位点模拟结果比较

### 5.1.1 正态误差情况下的比较

首先假设随机误差项服从标准正态分布： $\varepsilon_{it} \sim N(0,1)$ ，将三种面板数据线性均值回归模型的估计结果和四种分位数回归模型估计方法在中位点估计结果进行对比，估计的偏差 Bias 和 RMSE 值见下表 1：

**表 1：正态误差分布下中位点处估计量的 Bias 和 RMSE 值**

正态误差		面板数据线性均值回归			面板数据分位数回归(中位数)			
$\varepsilon_{it} \sim N(0,1)$		PLS	FE_Within	FE_FD	PQR	FETQR	FDQR	HJQR
T=5	Bias	0.4565	-0.0195	-0.1013	0.4236	-0.0188*	-0.0417	-0.0352
	RMSE	0.0732	0.0114*	0.0477	0.0691	0.0135	0.0173	0.0171
T=15	Bias	0.5340	0.0176	0.0641	0.4675	0.0142*	0.0384	0.0246
	RMSE	0.0775	0.0096*	0.0131	0.0742	0.0103	0.0124	0.0137

通过表 1 的结果可以看出，当模型的随机误差项服从正态分布并且时期数 T 较短时，从估计量的无偏性来看，偏差 Bias 值最小的是 FETQR，其次是 HJQR 和 FE\_Within，从估计的有效性来看，均方根误差 RMSE 最小的是 FE\_Within，其次是 FETQR 和 HJQR。当时期数 T 较长时，从估计的无偏性来看，偏差 Bias 值最小的是 FETQR，其次是 FE\_Within 和 FE\_FD，然后是 HJQR，而从估计的有效性来看，最好的是 FE\_Within，较为稳定。所以通过蒙特卡洛模拟结果可以看出，估计效果较好的三种方法是 FETQR，FE\_Within 和 HJQR，其中 FE\_Within 与 FETQR 的主要区别在于 FE\_Within 法估计的模型是面板数据线性均值回归模型，而 FETQR 法估计的是面板数据分位数回归模型，所以无法提供各个分位点上的估计系数，但是它们的原理都是先将面板数据中的每个个体的观测值变换为其平均数的离差观测值，然后利用离差变换数据估计模型的参数；FE\_Within 与 HJQR 法相比其待估参数较少，所以其估计精度相对较高，均方根误差表现较为稳定。所以总的来讲，当随机误差项服从正态分布时，固定效应变换分位回归法(FETQR)与固定效应面板数据的离差变换最小二乘法(FE\_Within)的估计效果相差不大，但明显优于其它面板数据线性均值回归方法。

从表 1 的结果还可以看出，PLS 和 PQR 法在不同时期数条件下估计的偏差绝对值均超过 0.4，甚至最高超过了 0.5，其对应的 RMSE 值也是几种估计方法中明

显最大的，这应该是直接忽略模型中的固定效应造成的，这说明即使是在正态误差的条件下，不考虑面板数据中可能存在的个体固定效应而做出的估计会有很大的偏差和风险，从而降低了参数估计的有效性。

### 5.1.2 其他误差分布情况下的比较

在考虑了随机误差服从正态分布后，本文探究了随机误差项来自对称厚尾分布和偏态分布的情况，模拟的比较结果如表 2：

**表 2：非正态误差分布下中位点处的 Bias 和 RMSE 值**

非正态误差		面板数据线性均值回归模型			面板数据分位数回归模型(中位数)			
$\varepsilon_{it} \sim t(3)$		PLS	FE_Within	FE_FD	PQR	FETQR	FDQR	HJQR
T=5	Bias	0.5237	0.0269	0.0842	0.4942	0.0142*	0.0171	0.0296
	RMSE	0.0782	0.0352	0.0964	0.0815	0.0139*	0.0258	0.0403
T=15	Bias	0.5029	0.0216	0.1053	0.4953	0.0069*	0.0086	0.0145
	RMSE	0.0742	0.0334	0.0652	0.0719	0.0138	0.0110*	0.0244
$\varepsilon_{it} \sim \chi^2(3)$		PLS	FE_Within	FE_FD	PQR	FETQR	FDQR	HJQR
T=5	Bias	0.5247	0.0418	0.2813	0.4832	0.0121	0.0090*	0.0108
	RMSE	0.0831	0.0293	0.0525	0.0431	0.0174	0.0129	0.0121*
T=15	Bias	0.4967	0.0324	0.1954	0.3731	0.0171	0.0102*	0.0125
	RMSE	0.0831	0.0457	0.0345	0.4902	0.0105*	0.0141	0.0188

表 2 的结果显示，当随机误差项具有厚尾分布时，面板数据分位数回归估计方法 FDQR 和 FETQR 要优于传统的线性均值回归模型的估计方法 FE\_Within 和 FE\_FD，因为面板分位数回归估计的偏差明显小于前三种面板数据线性均值回归估计的偏差，特别是 FETQR 方法，不仅偏差 Bias 较小，而且 RMSE 在所有情况下都是最小的，表现较为稳定。另外还可以看到，不考虑个体固定效应的 PQR 法虽然表现不如其他考虑个体固定效应的估计方法，但是在随机误差厚尾分布之下，比起同样不考虑个体效应的 PLS 法，其无偏性和有效性表现都更好一些。所以，

当随机误差服从对称厚尾分布时，面板数据分位数回归方法在估计的精确度和稳定性上都要优于线性均值回归方法。

当随机误差项服从偏态分布时，表现最好的是 FDQR 和 FTEQR 方法，特别是 FDQR 的 Bias 值是 7 种方法当中最小的，而 FE\_Within 和 FE\_FD 表现不佳。在有效性方面，FDQR、FETQR 和 HJQR 表现得较为稳定，而 PLS 和 PQR 两种方法的 RMSE 值相差不大且都比较高，但是 PQR 的偏差要明显小于 PLS。所以综合比较结果，当随机误差出现偏态分布时，FDQR 法的估计精度和稳定性都是最较好的，而面板数据线性均值回归方法不如其在正态随机误差条件下的表现，这进一步说明面板数据线性均值回归模型的参数估计方法对随机误差项分布的要求较高，当误差项时厚尾或者偏态分布时，对其估计结果的影响较大，所以综上所述三种情况，当随机误差项服从正态分布时，面板数据线性均值回归方法与分位数回归方法在精确度和稳定性方面差异不大，但当误差项出现对称厚尾分布或者偏态分布时，面板数据分位数回归方法比线性均值回归方法在估计时更精确和稳定。

## 5.2 关于面板数据线性均值回归与分位数回归方法的结论

综合 5.1 的蒙特卡洛模拟结果，可以得出下列结论：

(1) 在面板数据分析中，建立分位数回归模型进而采取不同的参数估计方法是一种较为有效的方式，能够在不同误差分布情形下对回归参数做出精确稳定的估计，并且能够提供不同分位数水平下的参数估计值，突出了局部之间的相关关系，能够全面的刻画分布的特征。

(2) 当模型随机误差项满足正态条件假设时，分位数回归方法与传统的均值回归方法相差不大，但是当不是正态误差时，分位数回归方法比线性均值回归方法在估计精度和稳定性上都要更好。

(3) 不管是面板数据线性均值回归方法还是分位数回归方法，如果直接忽略面板数据种存在的个体固定效应，会使得参数估计值与真实值之间有较大的偏差，并且估计也会变得不够稳定，所以在面板数据进行建模分析前，要判断模型中是否存在个体固定效应。

(4) 利用离差变换消除固定效应的方法应用于面板数据线性均值回归模型与

面板数据分位数回归模型的效果相当，而利用一阶差分消除固定效应的方法应用于面板数据分位数回归模型的效果不如其应用于面板数据线性均值回归模型，且均不如离差变换法。

(5) 面板数据的固定效应变换分位回归法(FETQR)在 3 种误差分布条件下，其估计结果的稳定性和精确度都较高，而一阶差分分位回归法(FDQR)在误差出现偏态分布时效果较好，基于最优化原理的模式搜索分位回归法(HJQR)在随机误差服从正态分布和对称厚尾分布条件下的表现优于随机误差在偏态条件下表现，但不如面板数据的固定效应变换分位回归法。

## 6. 实证分析

### 6.1 实证设计

本文利用 2014-2018 年我国各地区城镇居民人均可支配收入  $x$ (元)和消费支出  $y$ (元)的面板数据为例，分别建立面板线性均值回归模型和分位数回归模型，并采用在蒙特卡洛模拟中表现较好的离差最小二乘法(FE\_Within)和固定效应变换分位回归法(FETQR)对近 5 年来我国城镇居民收入消费模式进行相关的实证分析，探讨我国居民消费水平与人均可支配收入水平之间的关系，并且为了验证忽略个体固定效应给参数估计带来的影响，本文还引入了在蒙特卡洛模拟中表现较差的混合最小二乘法(PLS)和混合数据分位数回归法(PQR)作为对比。所有数据来源于《中国统计年鉴(2015-2019)》，消费支出  $y$  与人均可支配收入的散点图如下：

通过对总体数据作散点图容易看到  $y$  与  $x$  之间有明显较明显的线性关系，所以可以考虑采用线性模型来刻画。而通过散点图可以看出各地区  $y$  与  $x$  的斜率变化不大，但是截距却有明显的不同，即各地区居民平均边际消费倾向差异不大，但是自发消费水平存在较大差异，说明不应该忽略模型中的个体固定效应<sup>2</sup>。所建立的面板

---

<sup>2</sup> 为了作比较，本文还是求解了直接忽略模型中的固定效应后的参数估计结果



数据线性均值回归模型与分位数回归模型如下：

(1) 线性均值回归模型：个体固定效应模型

$$y_{it} = \alpha_i + \beta \cdot x_{it} + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

(2) 分位数回归模型

$$Q_{y'_{it}}(\tau | x'_{it}) = \beta(\tau) \cdot x'_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

$$\text{其中：} \quad y'_{it} = y_{it} - \frac{1}{T} \sum_{t=1}^T y_{it}, \quad x'_{it} = x_{it} - \frac{1}{T} \sum_{t=1}^T x_{it}, \quad \varepsilon'_{it} = \varepsilon_{it} - \frac{1}{T} \sum_{t=1}^T \varepsilon_{it}$$

## 6.2 实证结果比较

本文使用不同估计方法对参数  $\beta$  的估计结果如下：

**表 3：面板数据线性均值回归与分位数回归结果比较**

参数估计	面板数据线性均值回归		面板数据分位数回归(中位数)	
	PLS	FE_Within	PQR	FETQR
系数 $\beta$	0.677	0.631	0.654	0.633
标准误差	0.0294	0.012	0.0419	0.0087
t 统计量	61.458	51.746	41.2541	57.636
F 统计量	3777.1	10406.4	3590.3	11340.5
可决系数	0.8931	0.9884	0.6317	0.9896

通过表 3 可知：FE\_Within 估计的 F 值和可决系数明显比 PLS 估计高，这说明直接忽略个体固定效应会造成参数估计的精确性下降，FE\_Within 估计时能够极大地改进 PLS 估计，并且从残差分析图来看 FE\_Within 估计方法能够减小异方差的干扰，残差呈现正态分布。而在 FE\_Within 与 FETQR 估计的对比中，我们发现二者得到的系数  $\beta$  的估计结果较为接近，这也验证了正态误差条件下面板线性均值回归与分位数回归的效果相当的结论，但是二者都比 PLS 估计的结果小，这也可能是模型中存在个体固定效应而造成的。为了利用 FETQR 法在固定效应模型较好

的参数估计精确度与稳定性，本文以固定效应变换分位回归法来分析不同消费水平居民的收入支出模型是否存在差异：

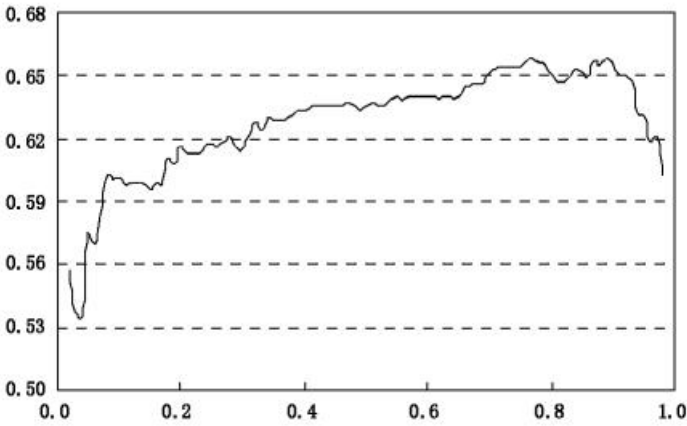


图 4：FEQR 估计随分位点变化函数图

图 4 给出了边际消费倾向 $\beta$ 随分位点 $\tau$ 变化而变化的函数图，我们发现， $\beta$ 值随着 $\tau$ 的变化呈现较大波动，当 $\tau$ 从 0.02  $\rightarrow$  0.10 时， $\beta$ 值快速增加，而在 $\tau$ 从 0.10  $\rightarrow$  0.77 时， $\beta$ 值则缓慢平稳增加，直至达到最大值 0.659，而之后则渐趋平稳并在  $\tau \approx 0.90$  后 快速下降，到  $\tau = 0.98$  时  $\beta$  值已降至为 0.603。边际消费倾向最高的是消费水平在 0.6 至 0.9 分位点之间的居民，即中高消费群体；最低的是 0.2 分位点以下的居民，即低消费群体，不过虽然他们的平均边际消费水平是最低的，但当他们收入增加时，其平均边际消费倾向增加的意愿却是最强烈的。另外对于 0.9 分位点以上的高消费居民，虽然他们的边际消费倾向不是最低的，但其收入的再增加并不会增加他们的边际消费倾向，反而是有明显的降低边际消费倾向的意愿，这也验证了微观经济学中边际消费倾向递减的结论。以上发现都是传统线性均值回归方法无法捕捉到的信息，但这些信息却为我们当前制定刺激消费的宏观经济政策提供了有利参考：即有效的刺激消费政策应该努力提高中低消费居民的收入水平。

## 7. 结论与展望

针对采用面板数据线性均值回归模型进行实证分析时所存在的诸多问题,本文引入能提供更多关于数据结构方面的信息的分位数回归模型,对比了面板数据线性均值回归方法与分位数回归方法之间的差异,其中面板数据线性均值回归方法包括最小二乘法以及广义最小二乘法等,分位数回归方法包括一阶差分分位回归法、固定效应变换分位回归法和模式搜索分位回归法;进一步探究了这些估计方法在是否存在个体固定效应以及误差分布是否为正态分布等情形下的具体表现,具体的发现以及结论包括以下三个方面:

第一,在面板数据线性均值回归模型与分位数回归模型的组间比较中,本文发现:(1)建立面板数据分位数回归模型进而采取不同的参数估计方法是一种较为有效的方式,能够在不同误差分布情形下对回归参数做出精确稳定的估计,并且能够提供不同分位数水平下的参数估计值,突出了局部之间的相关关系,能够全面的刻画分布的特征。(2)当模型随机误差项满足正态条件假设时,分位数回归方法与传统的均值回归方法相差不大,但是当不是正态误差时,分位数回归方法比线性均值回归方法在估计精度和稳定性上都要更好。

第二,在面板数据线性均值回归与分位数回归模型的多种估计方法之间的组内比较中,本文发现:(1)当模型中存在固定效应时,离差最小二乘法优于一阶差分最小二乘法,且都要优于直接忽略固定效应的混合最小二乘法;(2)面板数据的固定效应变换分位回归法(FETQR)在3种误差分布条件下,其估计结果的稳定性和精确度都较高,而一阶差分分位回归法(FDQR)在误差出现偏态分布时效果较好,基于最优化原理的模式搜索分位回归法(HJQR)在随机误差服从正态分布和对称厚尾分布条件下的表现优于随机误差在偏态条件下表现,但不如面板数据的固定效应变换分位回归法,但都优于直接忽略个体固定效应的混合分位数回归方法。

第三,在相同方法应用于不同回归模型时的差异比较中,本文发现:利用离差变换消除固定效应的方法应用于面板数据线性均值回归模型与面板数据分位数回归模型的效果相当,而利用一阶差分消除固定效应的方法应用于面板数据分位数回归模型的效果不如其应用于面板数据线性均值回归模型,且均不如离差变换法。

本文在面板数据线性均值回归与分位数回归方法的比较过程中,由于研究者的

时间和能力有限，还存在诸多的问题未能解决，主要问题包括：

(1) 针对不同的面板数据线性均值回归与分位数回归方法，本文将对应的参数估计量的无偏性与有效性通过蒙特卡洛模拟进行比较，但参数检验与置信区间的计算还未涉及，对于参数估计量的一致性、渐进性也未能进行推证，这是今后研究中有待解决的问题。

(2) 此外，本文只对比了面板数据线性均值回归与分位数回归模型的几种常见的参数估计方法，并未涉及分位数回归技术的各种前沿方法，将 Logistic 分位回归、协整分位回归等方法推广至面板数据样本条件下，将更好地比较面板数据线性均值回归模型与分位数回归模型在参数估计精确度与稳定性方面的差异。

## 参考文献

- [1] 徐进, 王淑超. 中国居民支出与收入的关系研究——基于分位数回归方法[J]. 北方经贸, 2012(5):47-50.
- [2] 罗幼喜, 田茂再. 面板数据的分位回归方法及其模拟研究[J]. 统计研究, 2010, 27(10):81-87.
- [3] 李群峰. 基于分位数回归的面板数据模型估计方法[J]. 2011:24-26.
- [4] 王娜, 任燕燕. 面板数据分位数回归模型求解的模式搜索法[J]. 数理统计与管理, 2016,35(3):435-444.
- [5] 齐晓丽, 金善女, 梁慧超等. 基于面板数据分分位数回归及实证研究[J]. 河北工业大学学报, 2010(3):98-101.
- [6] 刘焕鹏, 严太华. 面板数据分位数回归模型研究综述[J]. 统计与决策, 2014(17):82:84.
- [7] 罗幼喜, 李翰芳, 田茂再, 正列. 基于双惩罚分位回归的面板数据理论与实证研究[J]. 武汉大学学报. 2016,39(6):462-467.
- [8] 罗玉波. 分位数回归模型及其应用[M]. 知识产权出版社,2009.
- [9] 张晶, 王淼晗, 方匡南. 我国城乡居民收入影响因素研究——基于省级面板分位数回归分析[J]. 数理统计与管理, 2015,34(4):571-579.
- [10] 朱建平, 朱万闯. 中国居民消费的特征分析——基于两阶段面板分位数回归[J]. 数理统计与管理, 2012,(7):680-688.
- [11] 张所地, 范新英. 基于面板分位数回归模型的收入、利率对房价的影响关系研究[J]. 数理统计与管理, 2015,34(6):1057-1065.
- [12] Koenker R. and Bassett G. Regression Quantile[J]. Econometrica, 1978(46):33-50.
- [13] Koenker R. Quantile Regression for Longitudinal Data[J]. Journal of Multivariate Analysis, 2004(91):74-89.
- [14] Chernozhukov V, Hong H. 3-step censored quantile regression[J]. Journal of the American Statistical Association, 2002,97(459):872-897.
- [15] Chernozhukov V. Extremal quantile regression[J]. The Journal of Statistics , 2005,33(2):806-839.

- [16] Wang Y G, Shao Q, Zhu M. Quantile regression without the curse of unsmoothness[J]. Computational Statistics & Data Analysis, 2009,53(10):3696-3705.
- [17] Robinson P M. Statistical inference of a random coefficient autoregressive model[J]. Scandinavian Journal of Statistics, 1978,5(3):163-168.
- [18] Mairesse J and Z.Griliches, Essays in honor of Edmond Malinvaud[J]. 1990.
- [19] Kao. Scurious regression and residual-based tests for conintergration in panel data[J]. Journal of Econometrics, 1999,90.
- [20] He X, Fu B, Fung W K. Median regression for longtitudinal data[J]. Statistic in Medicine, 2003,22(1):3655-3669.
- [21] Canay I A. A simple approach to quantile regression for panel data[J]. The Econometrics Journal, 2011,14(3):368-386.
- [22] J. Paniagua, E. Figueiredo, and J. Sapena. Quantile regression for the FDI gravity equation. Journal of Business Research, 2015,68(7):1512-1518.
- [23] F.B Covas, B.Rump. Stress-testing us bank holding companies: A dynamic panel quantile regression approach. International Journal of Forecasting, 2014,30(3):691-713.

## 致谢

东流逝水，叶落纷纷。不知不觉在西南财经大学四年的美好时光就这样流逝了，回忆四年来的本科生涯，有低落、有笑容、有迷惘、有坚持、有黑暗、有光明。值此论文付梓之际，我谨向所有关心我、陪伴我、支持我的老师、同学、朋友、家人致以最诚挚的感谢与最美好的祝愿！

本论文最终得以顺利完成，非常感谢我的指导老师林谦副教授。在论文的完成与修改过程中，她都给予我尽心尽力的指导与帮助，其严谨的治学和学术态度深深影响着我，对我在研究生阶段的学习、工作、生活都必将产生积极的影响。借此机会，特向林谦老师表示最诚挚的感谢！

同时，感谢经济数学学院的所有领导和老师。他们严谨的态度、渊博的知识以及诲人不倦的品格一直感染和激励着我不断进步，使我大学四年的时光充实而有意义。“经世济民，孜孜以求”，西财精神在每位老师身上得以完美诠释。

在本论文的初步写作中，我也参照了大量的著作与文献，许多学者的科研成果与写作思路给我很大的启发，在此向这些学者们表示由衷的感谢。感谢我的家人、朋友、同学对我的关爱与支持，使我能够继续追求自己的人生理想与目标，感谢所有关心、帮助和支持我的人！

最后，衷心感谢在百忙之中抽出时间审阅本论文的各位老师，谢谢！

