

2019年（第六届）全国大学生统计建模大赛

# 中国经济发展与碳排放库兹涅茨曲线的验证研究

——基于省域面板数据的非参数回归模型

参赛单位： 中央民族大学

参赛者姓名： 许发明 于子涵 李文博

# 目 录

摘要.....	I
Abstract.....	II
一、引言.....	1
二、文献综述.....	2
(一) 国内外 EKC 曲线研究综述.....	2
(二) 研究方法综述.....	3
(三) 综述总结.....	4
三、数据来源与分析.....	4
(一) 变量选择.....	4
(二) 面板数据介绍与优缺点.....	5
(三) 数据来源.....	6
四、模型研究与建立过程.....	8
(一) 面板数据模型研究.....	8
(二) 非参数回归模型的估计方法.....	9
五、实证分析与模型求解.....	12
(一) 检验与模型选择.....	13
(二) 利用交叉检验法和创新法确定最佳窗宽.....	14
(三) Poolability 检验.....	14
(四) 消除固定效应.....	14
(五) 变窗宽局部线性估计.....	15
(六) 结果分析.....	16
六、结论与建议.....	18
参考文献.....	20
附 录.....	22
致 谢.....	26

## 表格和插图清单

表 1	八种主要能源的二氧化碳排放系数.....	6
表 2	2007——2016 各年窗宽.....	14
表 3	人均 GDP 对 CO <sub>2</sub> 排放量的预测.....	16
图 1	2007 年各省份碳排放总量.....	7
图 2	2016 年各省份碳排放总量.....	7
图 3	典型省份人均 GDP 随年份的变化图.....	7
图 4	北京天津 2007 年到 2016 年碳排放总量随时间变化的折线图.	8
图 5	数据单位根检验结果展示.....	13
图 6	协整性检验结果展示.....	13
图 7	基于省际面板数据碳排放量的 EKC 曲线.....	17
图 8	碳排放量的 EKC 曲线（平滑拟合）.....	17

## 摘 要

“绿水青山就是金山银山”，绿色经济成为国家可持续发展的重要基石，在绿色发展的深化改革中，减少碳排放量一直是全球性的热点话题和棘手问题。为了更好地解决碳减排的问题，需要对碳排放来源进行分析，并且量化其与经济规模之间的时空演变趋势，从而给出相应对策。

通过查看已有文献发现 1991 年提出的环境库兹涅茨曲线，可以用来描述碳排放量与经济发展水平之间的“倒 U 型”关系。关于 EKC 曲线的参数估计研究完善，但是非参数面板数据估计验证上文献很少。

本文主要建立非参数面板数据优化模型，验证环境库兹涅茨曲线的分支中国碳排放量的库兹涅茨曲线是否成“倒 U 型”分布，进一步探求两者之间的关系。采用 2006—2017 年的省际碳排放量的面板数据和人均 GDP 面板数据，首先进行 Poolability 检验确定最佳模型，其次利用交叉检验法和创新算法相结合确定最佳窗宽，接着用一阶差分、组内变换消除固定效应，最后运用局部线性估计和大量编程软件（R 语言、MATLAB 等）进行模型求解，并画出 EKC 曲线。

我们发现运用非参数面板数据优化模型给出的 EKC 曲线，不再是“倒 U 型”的，而是呈现为“N 型”，这说明环境不会随着经济增长而自发改善，对处理我国环境问题有着非常重要的借鉴意义。

**关键词：**二氧化碳；经济增长；非参数面板数据模型；局部线性回归；EKC 曲线

## Abstract

"Lucid waters and lush mountains are invaluable assets.", the green economy has become an important cornerstone of the country's sustainable development. In the deepening reform of green development, reducing carbon emissions has always been a global hot topic and a difficult issue. In order to better solve the problem of carbon emission reduction, it is necessary to analyze the sources of carbon emissions and quantify the temporal and spatial evolution trends between them and the scale of the economy, so as to provide corresponding countermeasures.

By looking at the existing literature, the environmental Kuznets curve proposed in 1991 can be used to describe the inverted "U" relationship between carbon emissions and economic development levels. The parameter estimation of the EKC curve is well researched, but the literature on nonparametric panel data estimation is rare.

This paper mainly establishes a non-parametric panel data optimization model to verify whether the Kuznets curve of the carbon emissions of the branch of the environmental Kuznets curve is inverted U-shaped distribution, and further explores the relationship between the two. Using the panel data of inter-provincial carbon emissions from 2006 to 2017 and the per capita GDP panel data, the first model is determined by the Poolability test, and then the cross-test method and the innovative algorithm are combined to determine the optimal window width, followed by the first-order difference. The intra-group transformation eliminates the fixed effect. Finally, the local linear estimation and a large number of programming software (R language, MATLAB, etc.) are used to solve the model, and the EKC curve is drawn.

We find that the EKC curve given by the non-parametric panel data optimization model is no longer an inverted "U" type, but an "N" type, which means that the environment does not spontaneously improve with economic growth. Environmental issues have very important implications.

**Key words:** carbon dioxide; economic growth; nonparametric panel data model; local linear regression; EKC curve

## 一、引言

近年来，随着社会经济的飞速增长与科学技术的日新月异，环境问题越发引起了人们的广泛关注，环境污染成为全球性问题。其中讨论较多的是二氧化碳排放的问题，过度的二氧化碳排放会导致温室效应、全球变暖等一系列严重问题，而化石燃料燃烧与人为活动是引发二氧化碳过度排放的主要原因。中国作为全球人口大国，同样也逐步成为碳排放大国。数据显示，20世纪90年代，中国二氧化碳排放量仅占全球总排放量的11%，而2013年中国人均碳排放接近法国，达到7.6吨/人。2014年中国碳排放总量达到世界第一，占全世界总排放量的26%。预测2020年中国碳排放量约占全球比重达到三分之一左右，由此表明中国未来将面临更大减排压力。

与此同时，中国也在节能减排方面推出了相应政策，做出了相应行动。2009年11月，国务院决定将“2020年单位国内生产总值二氧化碳比2005年下降40%-45%”作为约束性指标纳入国民经济和社会发展中长期规划。2014年11月12日，中美共同发表了具有历史意义和里程碑式的《中美气候变化联合声明》，中国政府提出在2030年左右二氧化碳排放达到峰值且将努力早日达峰。国民经济和社会发展第十三个五年规划纲要也指出“坚持减缓与适应并重，主动控制碳排放，落实减排承诺，增强适应气候变化能力，深度参与全球气候治理，为应对全球气候变化作出贡献。”从世界范围来看，从《京都协定书》到《巴黎协定》的签订，中国已逐步成为全球气候治理的领军者。相关政策的出台与相关协议的签订彰显了中国人民对“绿色、低碳、环保”社会的美好愿景，同时也彰显了中国作为世界大国的责任感与使命感。

在此背景下，研究我国飞速增长的经济与二氧化碳排放量之间的关系变得尤为重要。环境库兹涅茨曲线（environmental Kuznets curve, EKC），又称倒U曲线，是上个世纪50年代诺贝尔奖获得者、经济学家库兹涅茨于1991年所提出的用于研究收入分配状况与经济发展过程之间变化关系的曲线。EKC基本涵义可以理解为：在经济发展初期，环境污染程度不高，但此时人口增长较快而技术相对落后，由此导致资源浪费，从而造成环境的不断恶化。随着经济不断增长，科学技术的进步与人们环保意识的加强，污染物的排放得到控制，当经济发展达到一定水平，即拐点时，环境质量逐渐改善，环境污染开始由高趋低。即经济增长与环境压力之间呈现出“倒U型”关系。理论上来看，中国各省二氧化碳排放量与中国各省人均GDP之间可能也存在这样的“倒U型”关系，那实际情况是否如此？我国碳排放是否呈现出良好的EKC形态？

本文通过收集2007年至2016年我国各省8种化石燃料的终端消费量，采用

联合国间气候变化专门委员会 (IPCC) 提出的碳排放量计算方法 (IPCC, 2008) 计算这 10 年来我国各省 (西藏, 港澳台缺失数据除外) 碳排放量, 再收集这 10 年来各省人均 GDP 值, 基于面板数据采用非参数回归方法分析实际情况下我国经济状况和二氧化碳排放量的关系, 分析环境库兹涅茨曲线的存在性, 并据此得出结论, 提出相应建议与规划。

## 二、文献综述

自库兹涅茨曲线提出以来, 国内外针对 EKC 曲线的研究层出不穷, 并取得了许多卓越的成果, 而针对 EKC 曲线的研究, 由其研究方法、研究对象的不同, 结论也会因此不一致。

### (一) 国内外 EKC 曲线研究综述

Lindmark (2002) 通过对瑞典 1970-1997 年二氧化碳排放量与收入之间的关系得出“倒 U 型”曲线。Giles 和 Mosk (2003) 通过对新西兰地区 CH<sub>4</sub> 排放量和收入之间的关系, 也发现了“倒 U 型”曲线的存在。<sup>[15]</sup>

然而 EKC 本质来说只是一个经验现象, 而不是任何一种情况下都遵守的一个定律。Holtz-Eakin 和 Selden (1995) 通过对二氧化碳与收入之间的研究得出两者之间没有呈现出 EKC 形式, 而是二氧化碳排放随着收入单调递增。Shafik (1994) 通过对 1960-1999 不同环境质量指标与人均 GDP 之间的研究, 验证出二氧化碳与人均 GDP 之间呈线性关系。

Kaufmann et al (1998) 通过研究二氧化硫与收入之间的关系得出“正 U 型”曲线, 与 EKC 曲线正好相反。肖彦, 王金叶 (2006)<sup>[14]</sup> 等人通过选取 14 年来工业“三废”的数据研究广西省库兹涅茨曲线得出广西省人均 GDP 与工业“三废”的排放量之间也呈“正 U 型”曲线。

与上述不同的是, 张捷、张玉梅 (2008) 通过对广东 1985-2003 年间工业“三废”排放量和人均 GDP 增长量之间关系的研究, 得出其中存在“N 型”关系, 究其原因, 研究者们发现霍夫曼比率和 MS 比率与“N 型”曲线之间关联性很高。

Friedl 和 Getzner (2003) 通过对奥地利 1960-1999 年经济发展状况和二氧化碳排放情况关系的研究, 得出二氧化碳与收入之间呈现出“N 型”关系, 与通常情况下的“倒 U 型”关系不同。研究者们认为经济增长水平与环境压力之间的分离状态并不会长期持续, 当经济增长到一定水平时, 两者之间的关系重新组合, 由此呈现出“N 型”。

国内的研究大部分分为从全国范围、各省之间进行比较以及分东、中和西部的比较研究三种。杨万平和袁晓玲（2009）<sup>[12]</sup>通过对中国总体面板数据的研究得出废水、废气和固体废弃物排放量与经济之间遵循“倒 U 型”曲线。许士春和何正霞（2007）<sup>[10]</sup>通过中国 28 个省份的面板数据发现大多数省份遵循“倒 U 型”曲线。

周茜（2011）、李飞和庄宇（2012）、许广月和宋德勇（2010）将全国分为东、中和西部，由此对 EKC 曲线开展探讨，研究者们发现东部和中部呈现出明显的“倒 U 型”，而西部地区则呈现出相反的“正 U 型”。在对三个地区经济增长与环境压力之间的研究结果显示两者之间呈“N 型”曲线，并由于三个地区经济增长的不平衡因素，三个地区到达的拐点也就不同，其中，东部>中部>西部。

苏伟、刘景双（2007）通过使用吉林省 1986-2004 年人均 GDP 与相关环境数据，建立经济与典型环境指标关系的计量模型，得出工业废水排放量与人均 GDP 之间呈现出明显的线性关系，而工业废气排放量与人均 GDP 之间则呈现出二次“正 U 型”曲线，工业固体废物排放量与人均 GDP 之间呈现三次“N 型”曲线，即根据不同研究对象得出了不同的结果。

## （二）研究方法综述

目前针对 EKC 曲线的研究大多采用的是参数模型，即事先设计模型的形式，并假设其中的随机干扰项服从正态分布，但问题在于绝大部分研究在套用模型之前并没有对数据的随机特征进行检验的模型也没有对其假设是否合理进行判断的模型，即忽略了模型的稳健性，由此造成因参数设置不同而出现结果不同的问题。

在二氧化碳排放与经济增长关系研究方面，大部分国内外学者采用了参数模型，其中包括线性、二次多项式、三次多项式模型，处理方式则普遍采取对数据进行对数化，理由是以此去除异方差和其波动趋势。<sup>[11]</sup>

Shafik（1994）同时采用了线性、二次和三次的个体固定效应模型，但是选择模型时并未进行参数的合理性检验。Schmalensee 等（1998）通过 1950-1990 年间 141 个国家的数据，采用考虑了个体和时间固定效应模型的分段的线性函数进行拟合，以此估计碳排放的收入效应，结果发现呈现出明显的“倒 U 型”曲线。Holtz-Eakin 和 Selden（1995）<sup>[16]</sup>根据 130 个国家 1951-1986 年人均 GDP 数值和人均二氧化碳排放量，采用二次函数个体与时间的固定效应模型，得出结论样本外的库兹涅茨曲线近似为一条直线。Koop 和 Tole（1999）则使用了随机系数模型，结果发现拟合效果不佳。



除此之外,国内研究者在研究方法上也有类似研究,王锋(2010)采用 Divisia 指数分解法,把中国二氧化碳排放的增长率分解成了 11 种驱动因素的加权贡献,并对每一种因素分别研究,得出人均 GDP 为其主要驱动因素。刘华军(2011)采用协整检验方法,发现碳排放的库兹涅茨曲线呈“倒 N 型”。

由于多项式模型难以避免的递增或递减速率过快的问題,即使部分曲线在人均 GDP 合理范围之内,但是二氧化碳排放量的拟合值已经变为了负数。虽然对数据采取取对数之后可以减缓速度,但是曲线的整体形态已经因为取了对数而发生了变形。

因参数模型的缺陷性,有极少数研究者提出采用非参数或者半参数的方法展开研究。Luisito Bertinelli 和 Eric Strobl 采用半参数模型验证 EKC 是否存在,结果发现环境质量与经济发展自己呈现线性关系。Nguyen Van Phu 等(2008)<sup>[17]</sup>采用非参数模型对 100 多个国家从 1960 年到 1990 年的面板数据进行研究,结果与传统的参数模型结果相悖。

目前国内还鲜有研究者采用非参数方法对 EKC 进行检验和研究,付淼(2008)采用非参数的方法得到了我国废水、废气、固体污染物的 EKC,并根据分析提出了相关建议。<sup>[9]</sup>而有关二氧化碳排放和人均 GDP 的 EKC 研究方面仍是空白。

### (三) 综述总结

综上,国内外研究者广泛使用参数模型对二氧化碳排放与经济增长之间关系进行研究,虽然其存在可以进行预测的优点,但是由于没有对模型稳健性的相关检验致使得出的结论差距颇大。

而非参数模型尽可能的避免了参数模型的缺点,不仅可以避免模型设计上的误差,加上非参数模型的输出方式以图形为主,说明问题更加直白明了。并且于参数模型的不同点在于对于同样的数据样本,采取非参数模型得到的结果一般一致。考虑到目前国内外极少有研究者关于使用非参数方法针对中国省际面板数据与二氧化碳排放量的 EKC 曲线进行研究,因此,本文以此为切入点使用非参数模型对中国二氧化碳排放量与人均 GDP 的环境库兹涅茨曲线进行验证。

## 三、数据来源与分析

### (一) 变量选择

关于碳排放及其影响因素的研究一直是学术界关注的重点,大量文献表明碳排放量与经济规模有一定关联。在经济规模方面,环境库兹涅茨曲线(EKC)假

说认为，碳排放量随经济增长程度呈倒“U”型的变化，即在经济发展早期，碳排放量随人均 GDP 增长而增加，而在经济发展到一定程度后，环境质量会出现好转。经典线性计量经济学模型在分析时只利用了面板数据中的某些二维数据或者是利用截面数据，为了更好的利用非参数回归来估计人均 GDP 与碳排放量之间是否存在倒“U”型曲线关系，我们选用动态面板数据作为变量。

当选取的自变量过多时，会产生多重共线性，为了避免多重共线性的存在，更好地探讨碳排放与经济增长之间的关系，本文选取单一变量人均 GDP 作为自变量，选取碳排放量的时空动态面板数据作为因变量建立非参数空间估计模型，从而研究两者之间的关系。

## （二）面板数据介绍与优缺点

面板数据（panel data）是指在时间序列上取多个截面，在这些截面上同时取样本观测值所构成的样本数据。这种多维度的数据是把时间序列和截面数据融合在一起，方便进行多维度的回归和影响因素的研究，在研究人均 GDP 和碳排放量的模型建立中起到不可替代的重要作用，其也在多个领域，包括计量经济学领域有着重要发展和很多待发现的未知领域。<sup>[1]</sup>

下面是通过大量的文献查阅对面板数据的优点和局限性的分析。在分析优点的时候，和截面数据模型和只利用时间序列的模型进行对比，发现面板数据有以下优点：

1. 面板数据可以偶从多种层面分析关联问题和影响因素。
2. 面板数据通常提供给研究者大量的数据，也在无形中包含了更多的变量。这其实是一种很好的增加自由度的方式，减少了解释变量之间的共线性，从而增强了估计的准确和有效性。
3. 面板数据可以对个体不同进行有效的控制。
4. 面板数据能更好地识别和测量一些随机效应，这些是截面数据所不能衡量的。

同时，其也有一定的局限性：

1. 因为面板数据有二维属性，数据收集的错误和模型建立的偏差都会引起很大的震荡性事务。
2. 面板数据建立的模型的准确性和稳定性会受样本非随机造成观测值的偏差，使得回归统计有一定的偏差。

在文献综述中提到，在研究碳排放和经济增长规模的非参数回归中，大家普遍使用的是截面数据和时间序列相关数据，所以我们进行创新，基于面板数据的优势和局限性分析，我们将扬长避短，利用人均 GDP 和碳排放量的面板数据建立非参数空间模型求解问题。

### （三）数据来源

本文所有数据来源于国家统计局官方网站、《中国能源统计年鉴》以及部分省、市、自治区的不同年份的统计年鉴。由于一些省、市、自治区的具体能源消耗数据的缺失和不完整性，所以经过数据清洗与整理，我们获得了可用性高的 30 个省、市、自治区（不包括西藏）的相关数据。体现经济增长的规模数据中，我们选取了 30 个省、市、自治区的 2007—2016 年的人均生产总值(人均 GDP)；体现碳排放量的数据，我们选取煤炭消费量、焦炭消费量、原油消费量、汽油消费量、煤油消费量、柴油消费量、燃料油消费量、天然气消费量这八种能源在 2007—2016 年这十年的消耗量数据。

下面我们参考《综合能耗计算通则》和《省级温室气体清单编制指南》，利用 IPCC（政府间气候变化专门委员会）提供的二氧化碳排放量估算方法，估算了中国 30 个省市自治区 2007 年—2016 年的碳排放量。计算公式如（1）式所示：

$$C = \sum_{i=1}^n C_i = \sum_{i=1}^n E_i \times NCV_i \times CEF_i \times COF_i \times \frac{44}{12} \quad (1)$$

其中， $i=1,2,\dots,8$ ，分别代表煤炭、焦炭、原油、汽油、煤油、柴油、燃料油、天然气八种主要能源，C 代表二氧化碳排放， $E_i$  代表不同能源消耗量，NCV 代表平均低位发热量，CEF 代表单位热值含碳量，COF 代表碳氧化率，各对应值可见下表：

表 1 八种主要能源的二氧化碳排放系数

能源种类	煤炭	焦炭	原油	汽油
二氧化碳 排放系数	1.9003	2.8604	3.0202	2.9251
	煤油	柴油	燃料油	天然气
	3.0179	3.0959	3.1705	2.1622

（天然气单位为： $\text{kg}\cdot\text{CO}_2/\text{m}^3$ ，其他单位： $\text{kg}\cdot\text{CO}_2/\text{kg}$ ）

#### (四) 二氧化碳与人均 GDP 数据的初步分析

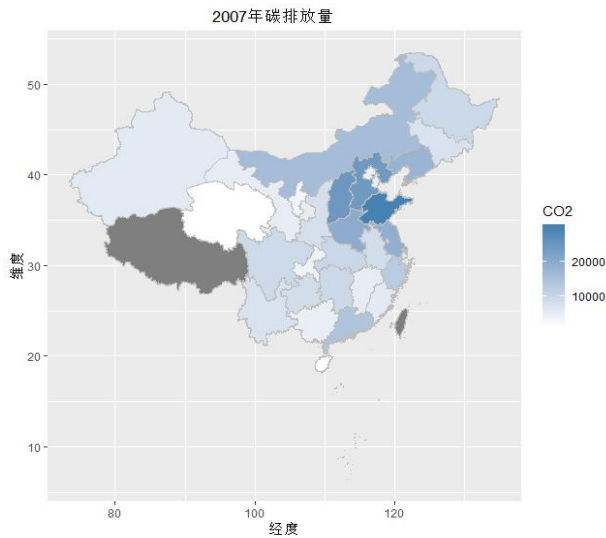


图1 2007年各省份碳排放总量

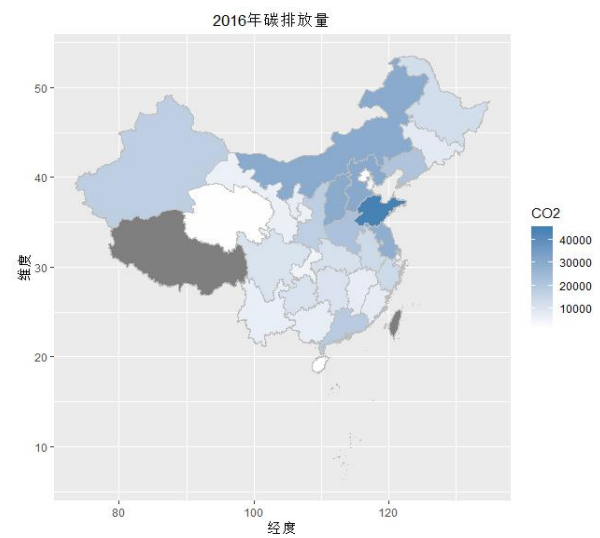


图2 2016年各省份碳排放总量

图1给出的是2007年各省份碳排放总量，图2给出的为2016年各省份碳排放总量，通过对比两图可以发现2016年碳排放量相较于2007年有明显上升(由于西藏，台湾数据缺失故在图中表示为灰色)，2007年全国碳排放最多的省份碳排放量没有超过30000万吨，但到2016年已有多个省份碳排放量超过30000万吨。通过图2可知，2016年，碳排放量超过30000万吨的省份有：山东省、内蒙古自治区、河北省，山西省，江苏省，碳排放量低于10000万吨的省份有：北京市、天津市、青海省以及海南省。先以它们为典型省份给出其2007年至2016年人均GDP变化曲线。<sup>[13]</sup>

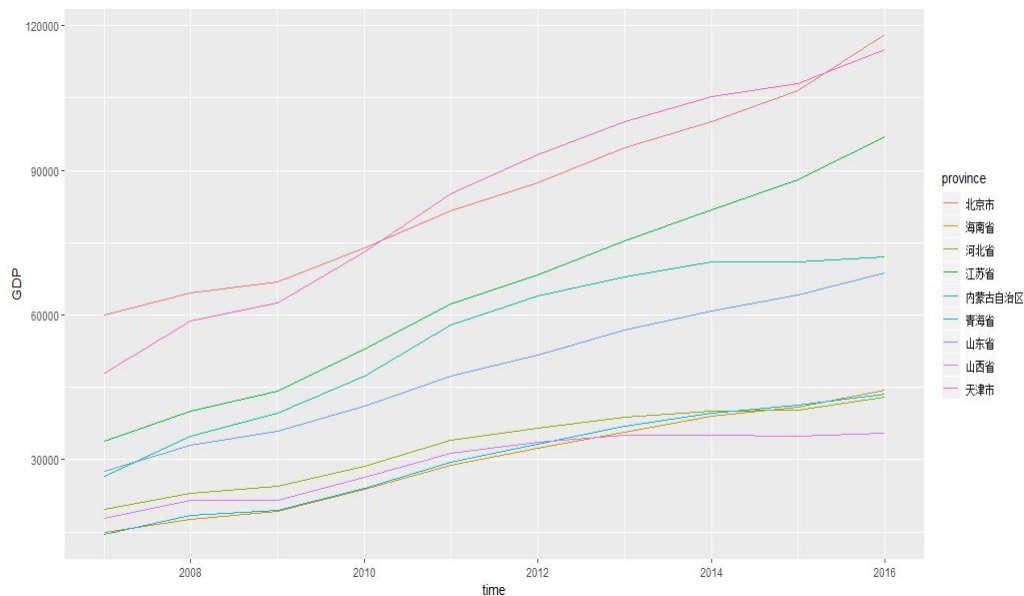


图3 典型省份人均GDP随年份的变化图

由人均 GDP 随年份的变化图（图 3）我们可以看出在 2007 年至 2016 年上述各省人均 GDP 都在逐步增加，这也可以体现碳排放量随着经济发展而增加。但北京与天津人均 GDP 远高于上述其他个省，碳排放量却也远低于其他省份。通过对北京，天津 2007 年到 2016 年碳排放总量随时间变画的折线图（图 4）的观察，我们可以发现北京市的碳排放在逐年下降。

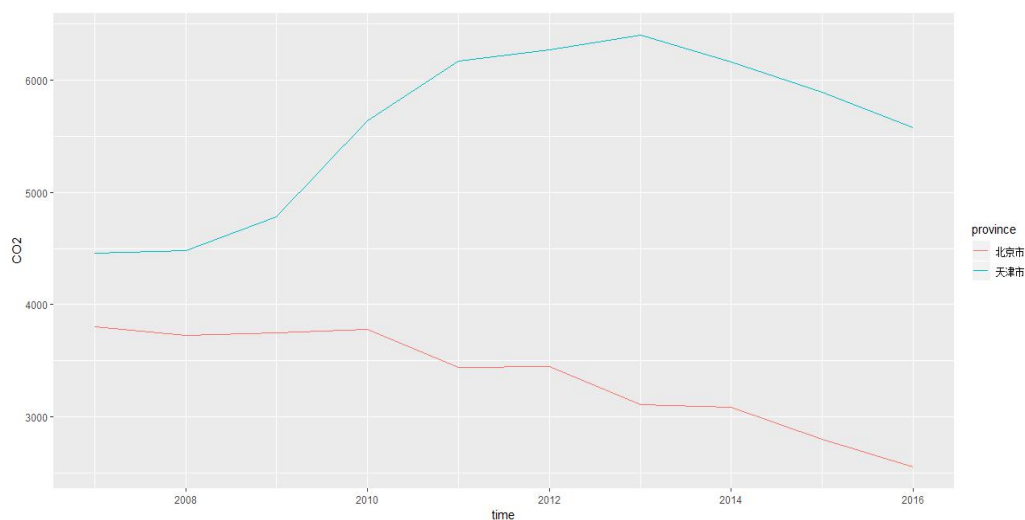


图 4 北京天津 2007 年到 2016 年碳排放总量随时间变画的折线图

天津市的碳排放量虽然在 2006 年到 2013 年有所上升但之后也呈下降趋势，这体现了碳排放量与经济发展的关系在北京市，天津市呈现的是随着经济增长碳排放量逐渐下降。

## 四、模型研究与建立过程

### （一）面板数据模型研究

#### 1. 面板数据概述

面板数据是指变量  $y$  关于  $N$  个不同个体的  $T$  个观测时间，所得到的二维数据，记为  $y_{it}$ ，其中  $i$  代表  $N$  个不同的对象， $t$  为观察时间内  $T$  个不同的截点。一般的面板数据模型表示如下：

$$y_{it} = \alpha + X_{it}^T \beta + \varepsilon \quad (2)$$

其中， $y_{it}$  表示因变量的列向量， $X_{it}$  表示解释的自变量的列变量， $i=1, 2, \dots, N$ ， $t=1, 2, \dots, T$ ； $\alpha$  为模型的截距项， $\beta$  为回归系数的列向量； $t$  表示面板数据的时间截点， $T$  对应时间截点的区间长度， $\varepsilon_{it}$  为随机误差项。

对于面板数据的分类可以分为静态面板数据和动态面板数据。其中静态面板数据包括固定效应模型、随机效应模型和混合回归模型；动态面板数据模型主要

是带有外生变量的线性动态模型。普通的面板数据模型用在计量经济学模型上运用最多的是固定效应和随机效应两类。其中随机效应是模型的结构参数对于不同的截面个体都存在着随机扰动的差异；而固定效应是模型的结构参数对于不同截面个体存在着实质上的差异。

## 2. 固定效应模型和随机效应模型

下面对于固定效应模型和随机效应模型给出简短定义和解释。

### ①固定效应模型

如果将面板数据模型定义为：

$$y_{it} = \alpha_i + X_{it}^T \beta + \varepsilon_{it}, i = 1, 2, \dots, N, t = 1, 2, \dots, T \quad (3)$$

其中， $\alpha_i$ 为随机变量，对于不同的个体其截距项各不相同，而且截距项 $\alpha_i$ 的变化和 $X_{it}$ 有关；针对不同个体 $i$ ,截距项 $\alpha$ 不同而回归系数 $\beta$ 相同。

### ②随机效应模型

若面板数据模型满足：

$$y_{it} = \alpha_i + X_{it}^T \beta + \varepsilon_{it}, i = 1, 2, \dots, N, t = 1, 2, \dots, T \quad (4)$$

其中， $\alpha_i$ 为随机变量，对于不同的个体其截距项各不相同，针对不同的个体 $i$ ，回归系数。通过观察我们可以看出来虽然形式与上文的固定效应模型很类似，但是值得注意的是其截距项 $\alpha_i$ 的变化和 $X_{it}$ 是没有关系的。

## (二) 非参数回归模型的估计方法

### 1.非参数回归模型

建立非参数面板数据模型形式，非参数回归分析不再以作者的主观臆断作为参考依据，而是通过数据作为回归依据。<sup>[4]</sup>设定形式如下：

$$y_{it} = G_i(x_{it}) + \varepsilon_{it}, t = 1, \dots, N, t = 1, \dots, T, \quad (7)$$

其中， $G(\cdot)$ 是未知函数； $\varepsilon_{it} = \mu_i + v_{it}$ 是随机误差项，我们将(7)式定义为多元非参数回归模型。非参数回归模型的估计方法有很多种，主要有局部线性估计、核估计、近邻估计、正交序列估计和多项式样条估计。<sup>[6]</sup>我们在求解和估计非参数回归的时候主要采用的是局部估计的方法，所以下面我们将介绍核估计方

法和局部线性估计方法，这两者都在后面实证分析中起到了很大的作用。

## 2.核函数

将光滑思想中的核函数极为  $K(t)$ ，且满足

$$\lim_{t \rightarrow \infty} K(t) = 0$$

此外给定的窗宽  $h(>0)$ ，记

$$K_h(t) = \frac{1}{h} K\left(\frac{t}{h}\right)$$

①Gauss 核函数:

$$K(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$$

显然，Gauss 核函数是均值为 0，方差为 1 的正态分布的概率密度函数。

②对称 Beta 函数族:

$$K(t) = \frac{1}{Be\left(\frac{1}{2}, \gamma+1\right)} (1-t^2)_+^\gamma$$

其中， $Be\left(\frac{1}{2}, \gamma+1\right)$  表示参数为 0.5 和  $\gamma+1$  的 Beta 函数， $(1-t^2)_+^\gamma$  表示函数  $(1-t^2)^\gamma$

的正部。

## 3.单变量窗宽的选择

在介绍核估计方法和局部线性估计方法之前，要对单变量窗宽的选择进行说明。通过理论探求可以发现，非参数的核估计值对核函数形式的变化相对没有那么敏感，但是却对于窗宽  $h$  的变化极为敏感。在计量经济学的实际应用中，大量文献叙述，如果窗宽过于大，就会使得给出的非参数回归函数的核估计接近于直线，这就说明我的局部加权平均的观测点太多，所以这个时候要减小这个窗宽  $h$ ；反之如果说，回归函数的核估计很不光滑，过于曲折的话，就应该让局部回归用到的加权平均的观测点变多，使窗宽  $h$  变大。这样看来，最佳窗宽有着合适的大小，窗宽是核估计的京都控制的一个重要参数。

理论上选择最优窗宽是从密度估计和真实密度之间的误差开始进行的, 因为分布密度函数是连续函数, 因而通常在积分均方差即 MISE 或大样本的渐近积分均方误差 AMISE 意义下, 求得最优窗宽。Rudemo(1982) 和 Bowman(1984) 提出了用交叉验证法 (cross-validation) 确定最终窗宽的方法, 使得窗宽的选择并不依赖总体的真实密度。<sup>[7]</sup>同时我们组成员也进行了进一步创新, 发现了一种探求窗宽的新方法, 将在下面进行陈述。

### ①交叉验证法

交叉验证法是一种常见的窗宽选择方法, 这是一个完全的数据驱动的方法, 由所知的样本数据来选择窗宽。<sup>[8]</sup>最小二乘交叉验证法的原理是选择使平方拟合误差最小的平滑参数  $h$ , 也就是对针对  $G(x)$  的所有的  $x$  求最佳窗宽。

具体来说,  $\hat{G}(x)$  与  $G(x)$  的差的平方的积分为:

$$\int [\hat{G}(x) - G(x)]^2 dx = \int \hat{G}(x)^2 dx - 2 \int \hat{G}(x)G(x)dx + \int G(x)^2 dx \quad (8)$$

我们发现  $\int G(x)^2 dx$  与窗宽  $h$  无关, 所以最小化式变为:

$$\int \hat{G}(x)^2 dx - 2 \int \hat{G}(x)G(x)dx$$

通过变换, 我们可以发现, 最小二乘交叉检验法其实是求使平方拟合误差  $CV_m(h)$  最小化的窗宽, 其表达式:

$$CV_m(h) = \frac{1}{n^2 h} \sum_{i=1}^n \sum_{j=1}^n \bar{k}\left(\frac{X_i - X_j}{h}\right) - \frac{2}{(n-1)h} \sum_{j=1, j \neq i}^n k\left(\frac{X_i - X_j}{h}\right) \quad (9)$$

这种计算方法是一种典型的数值搜索算法。值得注意的是, 在构建  $CV_m(h)$  时必须使用留一法核估计值, 如果没有使用的话, 这种最小二乘交叉检验的办法就会失效, 得到  $\hat{h}=0$ 。但是如果  $h=0$  使得目标函数最小, 那么就违反了一致性假设:  $n \rightarrow \infty$  时,  $nh \rightarrow \infty$ 。

### ②新方法求最佳窗宽

交叉验证法是常用的寻求最佳窗宽的方法, 我们小组成员通过大量的文献查阅, 发现在样本容量足够大的情况下, 核函数的选取对总体密度的估计的影响很小。据此, 我们给出两个核函数



$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^n K_1\left(\frac{x-X_i}{h}\right) \quad (10)$$

$$\check{f}_n(x) = \frac{1}{nh} \sum_{i=1}^n K_2\left(\frac{x-X_i}{h}\right) \quad (11)$$

既然核函数的选取对其总体密度的估计的影响微乎其微，那么两者的误差就不应该太大，我们选择最优窗宽  $h$  的原则也恰恰来源于此。考虑

$$MISE(h) = E \int (\hat{f}_n(x) - \check{f}_n(x))^2 dx \quad (12)$$

当其值最小时，求得窗宽  $h$  作为最优窗宽。

#### 4. 变窗宽的局部线性回归方法

局部线性估计不存在核估计的边界效应，这是因为局部线性估计的偏差和解释变量的密度函数是无关的。同时运用变窗宽的局部线性回归，它不仅估计出回归函数  $G(x)$ ，还可以估计出来回归函数的导函数  $G'(x)$ ，在对于后面计量经济学的分析弹性和乘数上起到不可忽视的作用。

$G(x)$  的变窗宽的局部线性估计矩阵为：

$$\hat{G}_n(x, h_n, \alpha) = e_1^T (X_x^T W_{x,\alpha} X_x)^{-1} X_x^T W_{x,\alpha} Y \quad (13)$$

其中， $e_1 = (1, 0, \dots, 0)^T$ ， $X_x = (X_{x,1}, X_{x,2}, \dots, X_{x,n})^T$ ， $X_{x,j} = (1, (X_i - x)^T)^T$ ，

$$W_x = \text{diag}\{K_{h_n/\alpha}(X_1), K_{h_n/\alpha}(X_2), \dots, K_{h_n/\alpha}(X_n)\}, \quad Y = [Y_1, \dots, Y_n]^T;$$

$\alpha(X_i)$  表示与解释变量的概率密度有关联的变窗宽函数。通过变窗宽的局部线性估计的性质可以知道，当窗宽为  $h_n = cn^{-1/(d+4)}$  时，局部线性估计的收敛速度在边界内点和边界上的点是相等的，所以也就不存在边界效应，可以说明变窗宽的局部线性估计是更胜一筹的，是一致估计量，且具有渐近正态性。

## 五、实证分析与模型求解

我们尝试建立一个非参数回归估计面板数据模型来检验中国 2007 年——2016 年碳排放量与人均 GDP 的关系。通过前面部分的数据调取、预处理和计算，我们已经得到了人均 GDP 的值  $y_{it}$ ，它表示第  $i$  个地区在  $t$  时间的数量值；而  $x_{it}$

表示第  $i$  个地区在  $t$  时间的碳排放的值。

## (一) 检验与模型选择

在使用面板数据前要对数据进行检验，首先进行单位根检验，验证数据是否平稳，通过 R 语言建模得到结果如下：

```
Augmented Dickey-Fuller Test

data: tlist_GDP
Dickey-Fuller = -5.7014, Lag order = 6, p-value = 0.01
alternative hypothesis: stationary

Augmented Dickey-Fuller Test

data: tlist_CO2
Dickey-Fuller = -5.9183, Lag order = 6, p-value = 0.01
alternative hypothesis: stationary
```

图 5：数据单位根检验结果展示

由上图可知，可知中国 2007 年——2016 年碳排放量与人均 GDP 数据均是平稳的。

其次，我们对模型进行协整性检验（格兰杰因果检验），为了判断我们选取的自变量和因变量之间是否存在因果关系。

结果如下

```
Granger causality test

Model 1: data$CO2 ~ Lags(data$CO2, 1:2) + Lags(data$GDP, 1:2)
Model 2: data$CO2 ~ Lags(data$CO2, 1:2)
  Res.Df Df      F    Pr(>F)
1     293
2     295 -2 19.814 8.494e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

图 6：协整性检验结果展示

由图 6 的检验结果，我们可知中国 2007 年——2016 年碳排放量与人均 GDP 这两个变量间存在因果关系，可以继续做回归分析。

最后是我们对于非参数回归模型的选择，由于我们选取了全国三十个省十年来的数据，可以认为样本量足够大且全面，所以我们认为应该建立固定效应模型。由此我们建立一个个体固定效应模型的非参数回归模型为：

$$y_{it} = G(x_{it}) + \mu_i + u_{it}, i = 1, 2, \dots, N, t = 1, 2, \dots, N \quad (14)$$

其中， $\mu_i$ 为固定效应， $u_{it}$ 为误差项，假定其为相互独立同分布的随机变量。

## (二) 利用交叉检验法和创新法确定最佳窗宽

$$\hat{G}_n(x) = \frac{1}{nh} \sum_{i=1}^n K_1\left(\frac{x - X_{it}}{h}\right) \quad (15)$$

$$\tilde{G}_n(x) = \frac{1}{nh} \sum_{i=1}^n K_2\left(\frac{x - X_{it}}{h}\right) \quad (16)$$

从而得到：

$$MISE(h) = E \int (\hat{G}_n(x) - \tilde{G}_n(x))^2 dx$$

当 MISE (h) 最小时得到窗宽。下面，通过 R 语言编程得到 2007——2016 年份各年度的窗宽结果。

表 2 2007——2016 各年窗宽

年份	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
窗宽 h	0.25	0.31	0.33	0.32	0.38	0.46	0.54	0.15	0.39	0.56

最后通过所有的各个年份的  $x_{it}$  值得到一个最佳窗宽  $h_0=1.99$ 。

## (三) Poolability 检验

在建模之前还要对 G 进行一个 poolability 检验,其目的目的是为了检验  $G_t$  不随时间变化。

原假设  $H_0: G_t = G$ , 对任何 t 成立

备择假设  $H_1: 存在 t 使得 G_t \neq G 成立$

通过检验发现我们的回归系数函数  $G(\cdot)$  不随时间改变。

## (四) 消除固定效应

如果说  $G(x)$  函数充分光滑，那么其在  $x$  点一定可以进行泰勒展开，展开式为：

$$y_{it} = G(x) + (x_{it} - x)G'(x) + \mu_i + u_{it} \quad (17)$$

我们要对 (17) 式进行组内的变换以此来进一步去消除固定效应，即：

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)G'(x) + (u_{it}(x) - \bar{u}_i(x)) \quad (18)$$

其中  $\bar{y}_i$  可以表示为：

$$\bar{y}_i = \frac{1}{T-1} \sum_{s=1, s \neq t}^T y_{is} \quad (19)$$

下面我们要给定一个假定， $E[u_{it} - u_{it-1} | x_{it}, x_{it-1}] = 0$ ，在这个的假设下，我们构建函数  $\psi_t(X_{it}) = E[y_{it} - y_{it-1} | x_{it}, x_{it-1}] = G_t(x_{it}) - G(x_{it-1})$ ，其中， $X_{it} = (x_{it}, x_{it-1})'$ 。

### (五) 变窗宽局部线性估计

通过前面对变窗宽局部线性估计模型的介绍，结合人均 GDP 和碳排放量的相关数据，进行如下非参数回归估计。

$$\hat{\psi}(x_0) = e_1^T (X_x^T W_{x,\alpha} X_x)^{-1} X_x^T W_{x_0} Y \quad (20)$$

其中， $e_1 = (1, 0, \dots, 0)^T$ ， $W_x = \text{diag}\{K_{h_n/\alpha(X_1)}, K_{h_n/\alpha(X_2)}, \dots, K_{h_n/\alpha(X_n)}\}$ ；众所周知，相比于窗宽  $h$  的选择，核函数  $K$  的选择准确性和合理性对模型的估计起到干涉作用会较小。我们运用高斯核函数，给出  $K_h(\tau_1, \tau_2) = \frac{1}{h_1} K(\frac{\tau_1}{h_1}) \frac{1}{h_2} K(\frac{\tau_2}{h_2})$  作为  $K$  的选择，同时在第二部分的基础上运用交叉检验和创新方法相结合的办法，给出最佳窗宽。

一旦， $\hat{\psi}(x_0)$  被局部线性估计方法确定获得后，我们就可以运用 1995 年 Linton 和 Nielsen 的方法去给出一个关于个体固定效应模型的回归系数函数  $G(x_{it})$  和  $G(x_{it-1})$ 。上面 Linton 和 Nielsen 的估计方法大概可以用下文的式子来描述。在描述之前，我们需要重新定义用  $x$  和  $y$  来定义  $\hat{\psi}$ 。我们可以写出：

$$E_y[\hat{\psi}(x, Y)] = \int \hat{\psi}(x, y) f(y) dy = G(x) - E_y[G(Y)] = G(x) - k \quad (21)$$

$$E_x[\hat{\psi}(X, y)] = \int \hat{\psi}(x, y) f(x) dx = k - G(y) \quad (22)$$

我们因而通过获得的  $G(x_{it})$  和  $G(x_{it-1})$  运用相似的样本均值估计, 得到下面两个式子:

$$\hat{G}^{(1)}(x_{it}) = \frac{1}{N(T-1)} \sum_{j=1}^{N(T-1)} \hat{\psi}(x_{it}, x_j) \quad (23)$$

$$\hat{G}^{(2)}(x_{it-1}) = -\frac{1}{N(T-1)} \sum_{j=1}^{N(T-1)} \hat{\psi}(x_j, x_{it-1}) \quad (24)$$

最后, 由于组内变换时常数项被消去了, 我们不能直接从局部线性估计中得到  $\hat{G}(x)$ , 但是我们可以用黎曼总和得到  $G(x)$ 。

$$\hat{G}(x) = [\hat{G}^{(1)}(x) + \hat{G}^{(2)}(x)] / 2 \quad (25)$$

## (六) 结果分析

我们运用了非参数估计的方法, 研究碳排放和人均 GDP 之间的回归关系, 并验证其是否符合倒 U 型曲线假说, 以此来给出一些结果分析和合理性建议。

运用 2007—2016 年人均 GDP 和碳排放的面板数据, 并运用修改后的交叉检验法选定最佳窗宽, 非参数估计如下:

表 3 人均 GDP 对 CO2 排放量的预测

GDP	CO2	GDP	CO2
0	869.6	11	2566.7
1	715.6	11.5	2997.7
7	670.7	12	5860.5
8.5	640.3	12.5	1747.6
9	658.8	13	1273.6
10	735.3	13.5	1.2141
10.5	1962.5	14	0.9274

通过上面的表格分析, 我们发现这十年来的省际面板数据都可以反应出, 碳排放量的变化与人均 GDP 的增加之间的关系是复杂的变化趋势, 即人均 GDP 每单位的增长带来的碳排放量的增加值逐渐减少, 且局部趋势呈先增加再减少再增加。为了更清楚的观察它们的变化趋势, 我们运用非参数面板数据的局部线性估计和编程语言, 将我国碳排放量与经济增长的关系, 绘制出曲线, 如下: (其中横坐标代表人均 GDP, 单位为万元; 纵坐标代表碳排放量, 单位为吨)

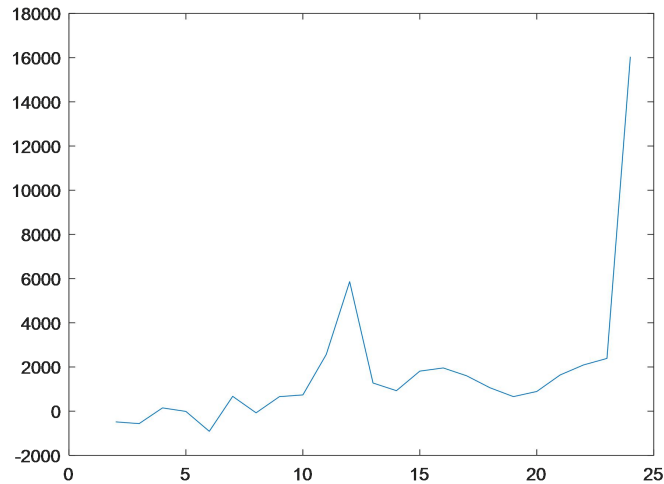


图 7 基于省际面板数据碳排放量的 EKC 曲线

我们将碳排放量的  $G(x)$  的固定效应模型的估计值以平滑曲线连接在一起，形成了碳排放量的环境库茨涅兹（EKC）曲线，通过了拟合度检验，发现碳排放量的 EKC 曲线拟合效果较好。<sup>[2]</sup>同时发现不再是假说中的倒 U 型，而是呈现出 N 型，即随着人均 GDP 的增长，省际碳排放量先增加，随后经过短暂的减少，又进入增加的过程，在 12.5 万元处存在第一个转折点，在 15 万元处存在第二个转折点。

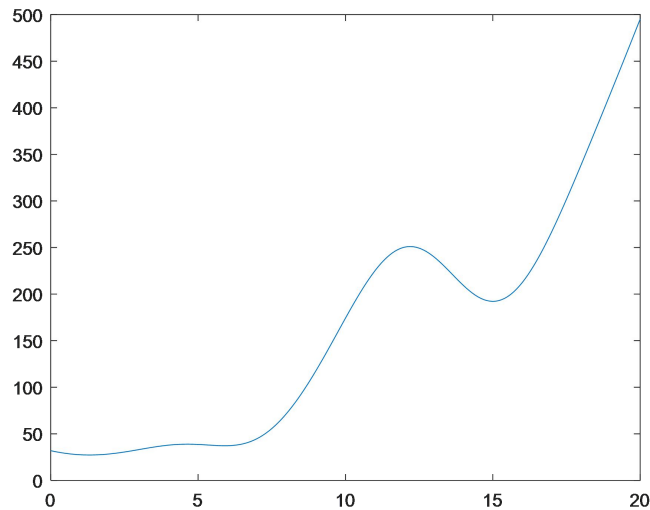


图 8 碳排放量的 EKC 曲线（平滑拟合）

同时由于非参数局部线性估计方法不仅可以估计出  $G(x)$  的值，还可以求出  $G'(x)$ ，即  $G(x)$  的导函数的值，所以我们可以来研究碳排放量的增长率随着人均 GDP 增长产生的变化。通过图 6 可以看出来，碳排放量的增长率先为正后为负再为正，在碳排放量增长的情况下，碳排放量的增加率呈现为先平稳后单调递增再单调递减的趋势。即随着经济的发展，人均 GDP 每增长一单位，碳

排放量的增加量越来越少。开始时，人均 GDP 每增长一个单位，碳排放量增加 5 吨，之后陡增为 20 吨，并不断的攀升，之后单调递减为 5 吨，之后有一段时间的稳定期。

综上实证分析，我们可以看出，人均 GDP 与碳排放量呈倒 N 型曲线反映在一段较长的时间维度上。总体来看，我国各省、市、自治区的环境质量（碳排放量）随着人均收入水平的增长来回波动，即初期环境随着经济发展产生幅度较小的振荡，变化不大。但是，很快环境随着经济发展而极具恶化，表现为碳排放量随着人均 GDP 的增加而陡增。但之后的转折点后有一小段时间里环境随着经济发展而呈现变好趋势，逐步改善。最后，仍然会随着经济发展的增加，碳排放量有一个无法控制的飙升，产生了违背环境库茨涅兹假说的脱钩关系。

## 六、结论与建议

本文利用我国统计局官网和能源统计年鉴的官方数据，对 30 个省、市、自治区 2006—2017 年的面板数据建立个体固定效应的非参数局部线性估计模型，分析碳排放量与经济增长的关系。结果表明，我们省际碳排放与人均 GDP 呈现非倒“U”型曲线关系，而是呈现为“N”型曲线关系。总体上我国存在着环境质量随着经济增长先平缓振荡，之后退化再改善，最后不可控制陡增的趋势。由于 10 年的时间范围比较短，对于未来是否能够达到平稳趋势，我们还难以估计。

我们展示了在参数估计下规范的环境库茨涅兹(EKC)曲线，虽然可以拒绝人均 GDP 的严格外生性，但是其碳排放、非参数化和一阶差分估计都与 EKC 曲线矛盾。我们可以说明由参数估计给出的环境库茨涅兹曲线在非参数估计的规范下是可以被拒绝的。

本文发现的我国碳排放与经济增长的“N”型曲线对处理我国环境问题有着非常重要的借鉴意义。经济发展到一定程度并不意味着环境问题就会迎刃而解，环境不会随着经济增长而自发改善，只有重视生态问题，加大环保投入，才会避免“脱钩飙升”现象。<sup>[5]</sup>同时我们还发现在局部存在着满足倒“U”型曲线的情况，但是拐点的来临也不单单只能靠经济增长所拉动的。所以，我们要平衡经济增长和环境问题，从环保技术、制度政策等方面多管齐下，吸取发达国家的经验教训，坚持可持续发展道路。

我们根据实证分析、文献查阅，依托国家国情和实际情况提出以下几点建议：

### （一）调整产业结构和能源消费结构

我国产业结构复杂，人口众多，很多产业结构都对能源有着较高的依赖度，

所以调整能源消费结构就显得尤为关键。我们要促进节能环保产业的高速发展，从国家层面给予大力的鼓励和政策扶持，鼓励第三产业的发展。与此同时，我们要多使用清洁能源，比如说天然气、风能、太阳能、可燃冰等，加快中国各省、市、自治区以一次能源为主的能源消费结构的优化，多使用可再生能源，加强绿色能源的技术研发，发展绿色经济。<sup>[3]</sup>

## **（二）建立中国绿色 GDP 核算体系**

将环境质量作为估测该省份经济增长的一个不可忽视的重要参数，由政策引导提倡可持续发展。完善自然资源成本的价值估计，要将自然资源的使用、碳排放的数值等归为政府业绩的评价体系。在设置中国绿色 GDP 核算体系的时候，还需要根据该省、市、自治区的发展程度进行聚类分析，给出略有不同的评价标准。还可以征收碳排放税等，运用相关国家部门强有力的看得见的“手”，来有效的在经济发展与环境保护中间给出良好平衡点。

## **（三）加强生态环保标准与科技创新合作**

我们在发展绿色经济的过程中，要进行深层次的改革与探索，国家层面支持和鼓励生态环保与科技相结合，大胆先行。我们要推动相关环保科技科研机构的成熟稳定发展，为绿色经济提供“智力保障”。比如说，可以自主研发低碳技术，研究可燃冰等作为稳定清洁能源的突破性技术难题，为绿色经济的中长期战略发展提供技术上的源源不断的支持。



## 参考文献

### 引文文献

- [1]许广月,宋德勇.中国碳排放环境库兹涅茨曲线的实证研究:基于省域面板数据[J].中国工业经济,2010(5):37-47.
- [2]付森.我国环境库兹涅茨曲线:形态、拐点和影响因素[J].数量经济技术经济研究,2008,(11):40-55.
- [3]王锋,吴丽华,杨超.中国经济发展中碳排放增长的驱动因素研究[J].经济研究,2010,(2):123-136.
- [4]刘华军,闫庆悦,孙曰瑶.中国二氧化碳排放的环境库兹涅茨曲线:基于时间序列与面板数据的经验估计[J].中国科技论坛,2011(4):108-113.
- [5]周茜.中国区域经济增长对环境质量的影响:基于东、中、西部地区环境库兹涅茨曲线的实证研究[J].统计与信息论坛,2011,26(10):45-51.
- [6]Alexander Cotte Poveda,Jorge Enrique Martínez Carvajal,Nicolas Ronderos Pulido. Relations between economic development, violence and corruption: A nonparametric approach with DEA and data panel[J]. Heliyon,2019,5(4).
- [7]Poveda Alexander Cotte,Carvajal Jorge Enrique Martínez,Pulido Nicolas Ronderos. Relations between economic development, violence and corruption: A nonparametric approach with DEA and data panel.[J]. Heliyon,2019,5(4).
- [8]Huadong Zhao,Wanzhu Tu,Zhangsheng Yu. A nonparametric time-varying coefficient model for panel count data[J]. Journal of Nonparametric Statistics,2018,30(3).

### 阅读型文献

- [9]李玉文,徐中民,王勇,焦文献.环境库兹涅茨曲线研究进展[J].中国人口·资源与环境,2005(05):11-18.
- [10]许士春,何正霞.中国经济增长与环境污染关系的实证分析:来自1990—2005年省级面板数据[J].经济体制改革,2007(4):22-26.
- [11]虞依娜,陈丽丽.中国环境库兹涅茨曲线研究进展[J].生态环境学报,2012,21(12):2018-2023.

- [12]杨万平,袁晓玲.环境库兹涅茨曲线假说在中国的经验研究[J].长江流域资源与环境,2009,18(8):704-710.
- [13]吕志鹏.中国CO<sub>2</sub>排放的环境库兹涅茨曲线形态及其地区差异分析[D].东北财经大学,2012.
- [14]肖彦,王金叶,胡新添,等.广西环境库兹涅茨曲线研究[J].西北林学院学报,2006,21(4):9-12,35.
- [15]Théophile Azomahou,François Laisney,Phu Nguyen Van. Economic development and CO<sub>2</sub> emissions: A nonparametric panel approach[J]. Journal of Public Economics,2005,90(6).
- [16]Youquan Pei,Tao Huang,Jinhong You. Nonparametric fixed effects model for panel data with locally stationary regressors[J]. Journal of Econometrics,2018,202(2).
- [17]Yoonseok Lee,Debasri Mukherjee,Aman Ullah. Nonparametric estimation of the marginal effect in fixed-effect panel data models[J]. Journal of Multivariate Analysis,2018.

## 附 录

附件一：最小二乘交叉法检验程序

```
function [y]=LSCV(X,h)
n=length(X);
y1=0;
y2=0;
for s=1:n
for t=1:n
        y1=y1+KK((X(s)-X(t))/h);
end
end
y1=y1/(n^2*h);
for i=1:n
for j=1:n
        y2=y2+K((X(i)-X(j))/h);
end
end
y2=y2/(n*(n-1))-K(0)/(n-1);
y2=2*y2/h;
y=y1-y2;
end
```

附件二：确定 h 值程序

```
x=0:0.01:20
n=length(x);
y=0;
min=1000000;
for i=1:n;
y(i)=LSCV(GDP1,x(i));
if y(i)<min
```

```

        min=y(i);
        hm=x(i);
end
end
plot(x,y)
hm
附件三：处理原始数据程序
Y=CO2(2:10,1)-CO2(1:9,1);
fori=2:30
    y=CO2(2:10,i)-CO2(1:9,i);
    Y=[Y;y];
end
Xstar=[GDP(2:10,1),GDP(1:9,1)];
fori=2:30
    x=[GDP(2:10,i),GDP(1:9,i)];
Xstar=[Xstar;x];
end
[a,b]=size(X);
t=linspace(1,1,a);
X=[t,Xstar];
附件四：FA函数程序
function [y]=Fa(X,Y,Xstar,x,h)
e1=[1,0,0];
l=length(Y);
z=linspace(0,0,l);
year=2:10;
fori=2:30
year=[year,2:10];
end
fori=1:l

```

```

    u=Xstar(i,:)-x;
    j=year(i);
    z(i)=Kh(u,h,j);
end
Z=diag(z);
y=e1/(X'*Z*X)*X'*Z*Y;
end

```

附件五：核函数程序

```

function [y]=Kh(x,h,j)
x1=x(1);
x2=x(2);
h1=h(j);
h2=h(j-1);
y=K(x1/h1)*K(x2/h2)/(h1*h2);
end

```

附件六：G函数程序

```

function [y]=G0(X,Y,Xstar,x,h)
G1=0;
G2=0;
X1=X(:,1);
X2=X(:,2);
N=length(X1);
for i=1:N
    G1=G1+Fa(X,Y,Xstar,[x,X1(i)],h);
end
G1=G1/N;
for i=1:N
    G2=G2+Fa(X,Y,Xstar,[X2(i),x],h);
end
G2=G2/N;

```

```
y=(G1-G2)/2;
```

```
end
```

附件七：绘图程序

```
x=0:1:20;
```

```
len=length(x);
```

```
y=linspace(0,0,len);
```

```
for i=1:len
```

```
    y(i)=G0(X,Y,Xstar,x(i),h);
```

```
end
```

```
plot(x,y)
```

## 致 谢

《中国经济发展与碳排放库茨涅茨曲线的验证研究——基于省域面板数据的非参数回归模型》论文完成在即，这一段刻苦钻研、虚心求教的难忘时光也将就此成为我人生中浓重的一笔。

首先感谢我们的指导老师魏传华老师！他严肃细致的科学态度，严谨的治学精神，深深地感染和激励着我们。在此次论文的选题上，魏老师根据我们的实际情况，结合国家热点话题和自己的擅长方向，为我们小组提供了很多的帮助，并认真耐心地为我们解答很多问题，是指引我们沿着正确方向前进的基石和明灯，使得我们很好的理解了选题背景，在专业能力和科研水平都得到很大的提高。魏老师为人正直诚恳，是我今后生活应该学习的榜样。在完成论文之际，我们小组全体成员，向您表示我最深厚的谢意。魏老师，谢谢您！

其次，感谢中央民族大学理学院所有老师对我们在学习上的帮助,在统计学领域方面，我们的专业素养得以很大的提升。

同时，我们还要感谢在一起愉快研究、鼎力支持的成员，大家一起相互支持、帮助，共同克服一个一个的困难和疑惑。还特别感谢师兄师姐在研究领域给予的方向指导和鼓励，谢谢你们！

最后，我们要感谢含辛茹苦培养我们茁壮长大的父母，在钻研课题遇到困难的时候给予我们恰当的安慰和宽解，并且在精神和物质上给予我稳固的支持。在这里，我们向我的父母和家人表示最诚挚的谢意！