

编号：B640

数据要素对山西省城镇相对贫困影响的测度研究

目 录

| | |
|-----------------------|----|
| 摘 要..... | 1 |
| ABSTRACT | 3 |
| 一、 引言 | 5 |
| (一) 研究背景 | 5 |
| (二) 研究意义 | 6 |
| 1. 理论意义 | 6 |
| 2. 实践意义 | 6 |
| (三) 研究内容 | 7 |
| (四) 创新点 | 7 |
| 二、 文献综述 | 8 |
| (一) 数据要素市场研究 | 8 |
| (二) 城镇贫困研究 | 9 |
| 1. 国外城镇贫困的认识与测度 | 9 |
| 2. 国内城镇贫困的认识与测度 | 10 |
| (三) 文献述评 | 12 |
| 三、 基础理论介绍 | 12 |
| (一) 城镇贫困 | 12 |
| 1. 城镇贫困人口的构成 | 12 |
| 2. 城镇贫困发展状况 | 13 |
| (二) 数据要素 | 14 |

| | |
|--------------------------------|----|
| 1. 数据要素推进政府数据开放共享 | 14 |
| 2. 数据要素提升数据资源价值 | 15 |
| 3. 数据要素加强资源整合和安全保护 | 16 |
| (三) 数据要素、数字经济与城镇贫困的关系 | 16 |
| (四) 研究方法 | 17 |
| 1. FGT 指数 | 17 |
| 2. 计量模型 | 22 |
| 四、 山西省城镇相对贫困测度 | 22 |
| (一) 数据说明 | 23 |
| 1. 人口与收入数据 | 23 |
| 2. 人均消费与贫困线数据 | 24 |
| (二) 山西省城镇相对贫困测度 | 24 |
| 1. 洛伦兹曲线拟合 | 24 |
| 2. FGT 指数的变化趋势 | 25 |
| 五、 山西省城镇相对贫困与数字经济关系的实证研究 | 26 |
| (一) 数据说明 | 26 |
| 1. 被解释变量数据 | 27 |
| 2. 核心解释变量数据 | 27 |
| 3. 控制变量数据 | 28 |
| (二) 山西省城镇相对贫困与数字经济关系研究 | 28 |
| 1. 仅有核心解释变量的计量模型 | 28 |
| 2. 加入控制变量的计量模型 | 29 |

| | |
|-----------------------------|----|
| 六、 结论与建议 | 31 |
| （一） 研究结论 | 31 |
| 1. 山西省城镇相对贫困变化趋势 | 31 |
| 2. 数字经济的发展加剧山西省城镇相对贫困 | 31 |
| （二） 建议 | 32 |
| 1. 稳步推进扶贫政策，消除相对贫困 | 32 |
| 2. 减轻数字贫困，加强数字教育 | 32 |
| 3. 保障义务教育，消除贫困文化 | 32 |
| 参考文献 | 34 |
| 附 录 | 37 |
| 致 谢 | 55 |

图目录

| | | |
|-----|--|----|
| 图 1 | 数据要素与城镇贫困的关系图 | 17 |
| 图 2 | 洛伦兹曲线 | 19 |
| 图 3 | 2003 年至 2019 年山西省城镇居民 FGT 指数变化趋势图..... | 26 |

表目录

| | | |
|-----|-------------------------|----|
| 表 1 | 2019 年人口与收入 | 23 |
| 表 2 | 人均消费与贫困线 | 24 |
| 表 3 | GQ 洛伦兹曲线拟合 | 24 |
| 表 4 | 变量描述性统计 | 27 |
| 表 5 | 标准化后的数字经济指标相关系数 | 27 |
| 表 6 | 主成分方差贡献率 | 28 |
| 表 7 | 仅有核心解释变量的计量模型回归结果 | 28 |
| 表 8 | 加入控制变量的计量模型回归结果 | 29 |

数据要素对山西省城镇相对贫困影响的测度研究

摘 要

数据要素是当今世界最具时代特色、最有现实意义、最有发展前景的生产要素,数据要素的核心价值就是实现数字经济的繁荣发展,目前数字经济已成为中国经济发展的新形式。数字经济的活跃会导致数字贫困的加剧,贫困群体与信息时代脱节,加剧了相对贫困,如今,随着城镇化进程的发展,越来越多的农村人口流向城市,城市人口构成复杂、城市结构变革,导致城镇相对贫困问题日益突出,因此我们要足够重视数字经济发展对城镇相对贫困变动的影响。

本文将城镇相对贫困置于数字要素发展背景下,以山西省为例,首先,从政府数据信息共享、发展数字经济及数据安全保护 3 方面系统阐述了数据要素与城镇相对贫困之间的关系,之后采用《山西统计年鉴》及国家统计局 2003-2019 年城镇数据,运用 GQ、Beta 洛伦兹曲线,对山西省的城镇相对贫困指标进行量化和测度,具体测算城镇相对贫困的 FGT 指数即贫困发生率、贫困缺口率和平方贫困距指数的变动,从相对贫困发生的广度、深度和强度来全面反映山西省城镇相对贫困的变动趋势。其次,本文采用计量模型,研究山西省数字经济对城镇相对贫困的影响。研究发现:2003 年至 2013 年,山西省城镇相对贫困不断加剧,且 2013 年相对贫困程度达到最大值;2013 年之后,山西省城镇相对贫困程度递减,说明扶贫政策发挥了作用,可以预见山西省城镇相对贫困程度将会继续减弱。另外,数字经济对山西省城镇相对贫困的广度、深度和强度都有十分显著的正向影响,表明山西省数字经济的发展会加剧城镇相对贫困的程度,大量相对贫困群体与数字经济时代脱节,导致城镇相对贫困人口增加。对山西省城镇相对贫困的测度,可以直观地反映我国扶贫政策在减缓贫困上的显著作用。数字经济发展水

平对山西省城镇相对贫困的显著正向影响,反映出数字贫困的强大影响力,因此,消除数字贫困对山西省减贫事业至关重要。

关键词：数据要素；数字经济；城镇相对贫困；FGT 指数

Abstract

Data element is the most characteristic, realistic and promising production element in the world. The central value of data element is to realize the prosperous development of digital economy. At present, digital economy has become a new form of China's economic development. The activity of digital economy will lead to the aggravation of digital poverty. The poor groups are out of touch with the information age, which aggravates the relative poverty. Nowadays, with the development of urbanization, more and more rural people flow into cities, the composition of urban population is complex, and the urban structure changes, which leads to the increasingly prominent problem of urban relative poverty. Therefore, we should pay enough attention to the impact of digital economic development on the change of urban relative poverty.

This paper puts relative urban poverty in the background of data element. Taking Shanxi Province as an example, firstly, the relationship between data element and relative urban poverty is systematically expounded from three aspects of government data information sharing, development of digital economy and data security protection. Then, the urban data of 2003-2019 results from the Shanxi Statistical Yearbook and National Bureau of Statistics. This paper uses GQ and Beta Lorenz curves to quantify and measure the relative urban poverty index of Shanxi Province. Specifically we will measure the change of FGT index of relative urban poverty, namely, the incidence rate of poverty, the rate of poverty gap and the index of square poverty distance. It comprehensively reflects the trend of relative urban poverty change in Shanxi Province from the breadth, depth and intensity of poverty. Secondly, this paper

uses the method of econometric regression model to study the influence mechanism of digital economy and relative urban poverty. The results show that from 2003 to 2013, the relative urban poverty in Shanxi Province is increasing, and the poverty level reaches the maximum in 2013. After 2013, the relative urban poverty level in Shanxi Province has declined, indicating that the poverty alleviation policy plays a role, and it can be predicted that the relative urban poverty level in Shanxi Province will continue to weaken. In addition, the digital economy has a significant positive impact on the breadth, depth and intensity of relative urban poverty in Shanxi Province. It shows that the development of digital economy in Shanxi Province will aggravate the level of relative urban poverty. A large number of poor groups are disconnected from the digital economic era, which leads to the increase of relative poverty in cities and towns. The measurement of relative urban poverty in Shanxi Province can directly reflect the significant role of poverty alleviation policy in China. The development level of digital economy has significantly positive impact on the relative urban poverty in Shanxi Province, reflecting the strong influence of digital poverty. Therefore, it is very important to eliminate the digital poverty in Shanxi Province.

Key words: Data element; Digital economy; Relative urban poverty; FGT index

一、引言

（一）研究背景

当今世界是大数据世界，我们写的文字、说的声音、画的图像、录的视频都是数据，数据遍布日常生活的各个角落。一般来讲，数据量越多价值越大，2019年10月28-31日，党的十九届四中全会审议通过《关于坚持和完善中国特色社会主义制度推进国家治理体系和治理能力现代化若干重大问题的决定》，将“数据”与传统六大生产要素并列，这充分凸显出数据在社会经济和人民日常生活中的巨大价值。为了促使数据要素成为助力经济更快、更高、更强发展的新动能，2021年3月11日，十三届全国人大四次会议表决通过了《关于国民经济和社会发展第十四个五年规划和2035年远景目标纲要的决议》，提出要提升培育数字创新能力，激活数据要素潜能，用数据助力经济、用数据优化治理体系、用数据普惠大众。建设数据要素市场的关键在于大力发展数字经济，网上购物、移动支付、共享单车、在线辅导教育、云端存储、智能物流分拣等数字经济商业形态逐步渗透到我们的日常生活中，数字经济正成为中国经济这艘大船的新引擎。

贫困一直是世界经济发展道路上的“拦路虎”，是世界人民最为广泛关注的问题。自2015年起，300多万名第一书记和驻村干部、25.5万个驻村工作队一道奋战在中国脱贫一线。到2020年底，我国已消除绝对贫困，但相对贫困仍然存在。如今，中国已进入数字经济时代，数字技术创新与日俱增，在加速发展数字经济的进程中，城镇相对贫困问题越来越突出，截至2018年底，我国城市低保人数为1008万人，其中特困人数为27.7万人；另外全国还有2000多万低收入群体，虽然他们收入高于低保标准，但是在住房、教育、医疗方面依然存在较多阻碍。如今，全国城镇人口数量已经占据五分之三，在当前形势下，城乡人口流动带来了城镇拥挤、城乡发展不协调等问题，进而对城镇经济带来一定的阻力。

对贫困进行测度是了解贫困发展状况的基础,如今,在现有的我国城镇相对贫困问题研究中,大多数学者主要是从收入层面入手来对城镇相对贫困问题进行研究,多维贫困测量指标大部分都是采用收入、健康、教育、生活等指标进行评价分析,但是,随着我国经济社会的蓬勃发展与数字化技术的日益进步,中国城镇相对贫困问题受收入以外的其他因素对贫困的影响越来越大,尤其是反映数字经济的因素。

(二) 研究意义

当今世界的数字创新技术展现出得天独厚的优势且正在为社会创造更多的价值,我们应重视数字经济发展程度,同时,我们需要特别关注数字经济发展进程中衍生出的城镇相对贫困问题。

1. 理论意义

在众多城镇相对贫困的研究中,我国学术界在研究城镇相对贫困问题时都没有把它纳入数字化进程中。大多数学者往往从收入分配、基尼系数等角度来测度城镇相对贫困,本文主要研究中部地区贫困发生率较高的山西省,采用 FGT 贫困指数、GQ、Beta 洛伦茨曲线来度量城镇相对贫困发生的广度、深度和强度,并且在以往影响城镇相对贫困因素的基础上,结合反映数字经济的因素来共同探究其对城镇相对贫困的影响,这是对城市贫困问题研究的补充和完善,同时也可以为日后研究数字经济与贫困之间的理论关系奠定基础。

2. 实践意义

我国正处于经济建设转型的重要战略时期,大多数人可享受数字经济福利,数字经济的发展水平与城镇相对贫困问题密切相关。对城镇相对贫困问题进行测度可为减弱城镇相对贫困提供建议,进而为国家的建设添砖加瓦。

（三）研究内容

本文以山西省为例，主要利用 FGT 指数测度城镇相对贫困，详细描述山西省城镇相对贫困的广度、强度和深度，并且在数字经济发展背景下将反映数字经济的因素引入计量回归模型来深入探究数字经济与城镇相对贫困的影响机制，这为山西省减缓城镇相对贫困提供一定的理论支撑，且可以根据研究结论寻找有效的减贫措施，进而促进山西省经济健康发展。

第一章是引言，介绍论文的研究背景及意义，总结本文的研究内容与创新点。

第二章是文献综述，对数据要素市场研究、城镇相对贫困内涵与测度文献进行综述。

第三章是基础理论介绍，对数据要素市场化背景下城镇相对贫困问题的相关理论进行阐述。首先分别对数据要素市场和城镇相对贫困含义进行阐述，然后系统阐述了数据要素与城镇相对贫困之间的关系。

第四章是城镇相对贫困测度。使用 2003 年至 2019 年的山西省统计年鉴城镇宏观数据，采用 FGT 指数对山西省城镇相对贫困进行测度。

第五章是基于计量模型研究山西省数字经济的发展水平对城镇相对贫困的影响，以 3 个 FGT 指数为被解释变量，以数字经济的发展水平为核心解释变量进行模型研究。

第六章为结论与建议，主要是基于回归实证分析结论，在数据要素市场建设以及数字经济发展背景下给出合理的减缓城镇相对贫困的建议。

（四）创新点

本文的创新点首先在于研究视角上的创新，在众多关于城镇相对贫困测度的研究中，大多数学者往往是单纯从收入分配角度来分析，而本文将城镇相对贫困置于数字经济发展的背景下，具有现实意义；其次，本文结合山西省城镇相对贫

困实际发展状况,并采用最新统计年鉴数据来实证分析数字经济的发展水平对城镇相对贫困的影响,着重强调二者之间的关系,为山西省城镇相对贫困研究注入新鲜血液。

二、 文献综述

(一) 数据要素市场研究

现代数字技术与信息科学的发展将人类带进数字经济新时代,数据要素是数字经济时代的“支架”,它主要作为生产性资源投入来创造有经济价值的信息、数据和知识,作为国民经济的重要生产要素,数据被赋予了全新的含义。何玉长(2021)指出数据在数字经济体系中占据着重要地位,经过云计算处理形成的数据资产,是对信息进行收集和处理后的数据聚合^[1]。我国网信办指出,数字经济是以数据要素为根基,以信息网络为宿体,以高效利用信息通信技术为催化剂来提升效率及优化经济结构的新兴经济活动^[2]。王芳(2020)认为数据资源主要产生于互联网活动所产生的记录,具有复杂、多样、海量、灵活多变等特点^[3]。

在市场经济条件下,数据要素的供求依托于市场交易,数据要素价格是在价值的基础上通过市场来形成。何玉长(2021)指出数据要素市场化的要件主要包括数据要素市场配置、定价、制度、竞争、交易^[1]。杨锐(2020)提出数据要素市场的构建和完善有助于降低行为主体借助市场实现数据交易的成本,而培育要素市场的突破点在于促进数据供给的市场化进程,这一过程的关键在于造就多元化,特别是那些有能力在数据价值上发挥重要作用的数据供给主体;刺激和聚焦行业对数据使用的需求;加速数据价值链环节的市场化进程;设置国家数据管理局,统筹构建多元化培育体系;降低数据要素市场的交易成本^[4]。郭琨(2021)提出数据要素具有独特的技术特征、经济特征及市场属性。在技术上以网络设施

为关键媒介，以“大数据”为存在形态，以拓展聚合为利用方式；在市场属性上表现为以数字平台为主要组织形态，展现交叉网络效应，基于网络效应形成市场正反馈以及在价值形成过程中的多层次性^[5]。唐要家（2020）指出数据要素与经济增长息息相关，数据要素需求侧与供给侧相互协同的自强化机制使得数据要素具有显著增长倍增效应与递增规模效应，数据要素可通过多要素合成效率提升、增强市场有效性、实现较好的公共治理及驱动高效率创新来助力经济增长^[6]。

（二）城镇贫困研究

贫困分为社会物质生活贫困和精神生活贫困，它具有历史性、社会性和地域性。根据贫困的程度，将贫困分为绝对贫困及相对贫困。城镇贫困的涵义与演变依赖于贫困，当前，国内外学众多者都对城镇贫困问题进行了大量的研究。

1. 国外城镇贫困的认识与测度

国外学术界的学者普遍认为城镇贫困最初由英国的朗特里学者和布什学者提出，发展至今已有一百多年的历史。特别是近些年的城市化发展背景下，在城镇生活的人越来越多，城镇成为主要的社会经济活动场所，由此暴露了城镇贫困问题，因此解决好城镇贫困问题对社会发展至关重要。

在城镇贫困的认识方面，国外学者认为城镇贫困不仅是简单的经济问题，它涉及到社会的各个方面。Mingione（1993）提出在过去的20年里，贫困人群的社会生活状况随着工业化的进程变得愈加严重，大城市的街道上处处可见乞丐与流浪汉，很多社会底层群体面临无工作、无收入、无社会保障的窘境，街头犯罪情况愈加频繁。该学者对这种状况进行分析，认为影响贫困的因素是多方面的，例如空间区域分割、种族、社会阶层分化等因素都会影响城镇贫困^[7]。

在城镇贫困的定义方面，学术界众说纷纭，其中较具代表性的是 Silver（1993）

的解释,他认为在经济与就业制度的改变和社会福利制度的更新过程会导致部分人失业和享有低收入,加剧城镇贫困问题^[8]。

在城镇贫困的测度方面,Bourguignon 和 Chakravarty (2003)借助 FGT 指数进行多维贫困测度^[9]。Alkire 和 Foster (2011)改进了 FGT 指数,提出了基于 FGT 指数调整的多维贫困测度方法,并被众多学者所采用^[10]。Knight (2004)基于 1988 年和 1995 年的我国家庭调查数据,将城镇家庭十等分组,以最低收入组的非食品支出与食品支出的比值为依据测算城市贫困发生率,研究发现城镇贫困的加剧源于失业和社会福利保障制度的不完善^[11]。Beccaria (2013)以拉丁美洲的劳动力为调查对象,发现该国家存在过度城市化问题,且认为劳动力流动和社会福利保障体系的不健全是造成城镇贫困最主要的原因^[12]。Meng (2005)用马丁法测算 1986 年至 2000 年的贫困线上下限,并计算出贫困发生率,并且得出在该时间段城镇贫困率上升是因为物价上涨和生活基本保障支出增加等^[13]。

2. 国内城镇贫困的认识与测度

我国学术界于 20 世纪 80 年代开始研究城市贫困问题,主要是结合国外的相关贫困理论并依托于中国的实际贫困问题展开研究。

国内学术界对城镇贫困的定义又要有两种观点:部分学者认为城镇贫困是发生在城市里的贫困,这其中包括居住在城市内但未在城市落户的贫困人口。还有部分学者聚焦于研究贫困理论,他们基于不定期的人口流动与信息的不对称性的背景,将城镇贫困特指有城市户籍的贫困群体。此后,“新城镇贫困”新定义出现,在定义下的新城市贫困群体主要包括长期无正当工作或稳定工作甚至是失业的成年人、被社会所孤立的老人、缺乏教育背景的年轻人、被当地文化所排斥的移民等。随后黄颖(2012)从“城市新贫困”的视角,阐明我国城镇贫困群体中新贫困人口的发展规模及趋势,并对我国城镇返贫过程中存在的潜在问题提出相

应的建议^[14]。随着城市的快速发展，城镇贫困问题日益突出。

在城镇贫困的认识方面，李东坤（2021）指出，随着城镇化推进与农村转移性贫困的出现，中国的城镇贫困问题日渐凸显^[15]。与农村贫困不一致，梁汉媚（2011）认为城镇贫困人口土地保障底线不足，他们与农村人口相比更容易面临社会排斥和边缘化的风险^[16]。

在城镇贫困的测度方面，近些年，国内外相关学者对我国问题进行了较多的研究。他们出于不同的角度，分析时使采用不同的方法来测算城镇贫困的程度，并在此基础上探讨致贫的原因及特征。李冬（2016）研究城市化背景下的城镇贫困，应用 FGT 指数对湖北省 1992-2012 年的城镇贫困进行测度，并从经济增长、人口城镇化和产业结构三方面阐释城市化对城镇贫困的影响机制，得出湖北省经济增长可以缓解城镇贫困^[17]。单德朋（2014）基于 FGT 指数对我国民族地区的城镇贫困情况进行测算得出民族地区的相对贫困问题日益加剧^[18]。张建华（2006）先是确定贫困线，接着对总量贫困进行测量，得出经济转型期间城镇贫困的广度、深度与强度呈现上升趋势，出现这一现象是由于社会经济结构转型、制度变化及对外开放程度等多种因素共同作用导致^[19]。张冰子（2019）用马丁法测算各省各年城镇地区的贫困率，研究结果发现 1988-2013 年间城镇地区的贫困率显著下降，贫困群体多数为选择性贫困，社会人员消费不足是导致我国城镇人口贫困的主要原因^[20]。陈宗胜（2017）构建阶梯式城镇贫困线并对其测度发现，中国城镇绝对贫困下降，相对贫困尤其是中西部地区的城镇贫困仍较为突出^[21]。刘娜（2019）利用城市经济贫困指标来测度民族地区的城市经济发展水平，得出民族地区贫困城市数量多，贫困人口占比大，城市发展水平整体偏低，民族地区特色和资源开发有助于经济发展，但是关键在于交通基础设施水平的提高^[22]。陈书（2015）采用 ELES 模型确定城镇贫困线，随后利用城镇贫困线求出贫困发生率，并以此得

出我国城镇贫困问题愈加严重的结论^[23]。侯卉(2012)用多维贫困测度方法,从社会排斥、参与度不足、能力贫困和收入贫困4方面7个维度对城镇贫困进行深入评估,为制定相关政策提供依据^[24]。

(三) 文献述评

总的来说,当前学者对数据要素的研究多是描述数据要素的定义,讨论数据要素市场的特征,数据要素市场与经济增长的关系;城镇贫困研究从城镇贫困的内涵与认识出发,研究城镇贫困的特征、城镇贫困的影响因素、城镇贫困测度及减贫建议等方面,尚未将数据要素与城镇贫困联系起来进行理论和实证分析。基于此,本文试图结合当前精准扶贫与数字经济化背景,尝试对山西省城镇贫困进行识别,并研究数据要素对山西省城镇贫困影响的影响,为解决山西省城镇贫困问题提供借鉴。

三、 基础理论介绍

(一) 城镇贫困

城镇贫困是指由于经济体制改变、产业结构调整、社会结构转型、企业改革创新等原因所导致的贫困。如今,城镇贫困正在逐步侵蚀社会经济发展,失业下岗、退休人员、无补助农民工、进城务工及外地务工者,形成了城镇贫困人群。

1. 城镇贫困人口的构成

改革开放之前,城市贫困人口主要是指“三无人员”,他们是相对于其他城镇贫困人口最弱势的群体,是扶贫首要对象。然而随着改革开放的持续深入,社会经济结构改变,产业结构也跟着经济结构进行调整,突如其来的社会变革使得大批社会人口陷入贫困,他们适应能力弱、应对危机能力低,不能很好地寻找就

业契机，使之成为新时代生活困难的特殊群体。该特殊群体主要由下岗职工、危机企业员工、退休人员以及由农村进城务工人员（该人员多打短工以获取经济收益，所以工作稳定性差、收入波动较大、劳动权益易受侵害，且由于是农村户籍而不能享受城镇福利保障）组成，因此，城镇贫困人口不仅包括“三无”社会人士，也应该包括有劳动能力但是没有就业机会的贫困人员。

2. 城镇贫困发展状况

了解城镇贫困发展状况是解决贫困的基础，根据《2019 年度山西省人力资源和社会保障事业发展统计公报》可知，2019 年外出农民工 271.4 万人，比上年增加 12.2 万人，增长 4.7%，年末城镇登记失业人员 21.26 万人，城镇登记失业率为 2.71%，虽然失业率较上年有所减缓，但是随着农村人民大量涌入城市，城镇失业率还有待观察，必须控制在合理的范围内，才能保证经济良好发展。在物质生活条件上，城镇贫困群体较非贫困群体来说，他们的可支配收入与消费支出都低于非贫困群体，且他们的财富积累速度也低于非贫困群体；在精神、身体健康方面，他们的健康状况也是较差的，贫困不只反映在物质层面，精神健康层面也很重要，对于城镇贫困群体来说，他们的工作、家庭压力都大于非贫困群体，这些无形的压力对他们的精神及健康都是非常不利的，良好的身体是革命的本钱，只有在健康的思想与强壮的体魄基础之上才能以最好的状态处理工作及家庭的各种事件，如果身体健康出现异常，对于本就收入不高的城镇贫困人口来说，高昂的医疗支出无疑是雪上加霜；在教育方面，许多贫困家庭因各种原因而无法获得基础性教育，他们的文化程度低于正常水平，高等教育费用对他们来说更是“心有余而力不足”，导致了贫困群体中父代与子代间周而复始的弱势地位。综上所述，城镇贫困人口的经济、教育、思想发展依然受阻，城镇贫困问题日益凸显。

（二）数据要素

数据作为国家战略性及基础性资源，已成为新时代社会发展的生产要素。随着区块链、AI、5G 技术、云计算、量子计算等数字创新技术的广泛应用，我国数据体量迅猛增长。吴志刚（2020）分别从知识演化、人类历史、技术演进、发展趋势、基础资源以及数据科学六个维度来理解数据要素，在知识演化维度上，利用观察、度量等手段获取原始数据，之后将有价值的数据加工成信息，最后对各类信息进行整合归纳形成知识；在人类历史维度上，3500 年前，人类利用图形符号对数据信息进行记载，2000 年前造纸术、印刷术的发明为数据找到了“宿体”，19 世纪中期以后，数据信息传播介质由传统物质转化为电；1940 年底到现在，信息通信技术推动着人类快速进入信息社会；在技术演进维度上，1946 年第一台计算机诞生，1969 年互联网出现，20 世纪 90 年代开始全球信息化建设，数据逐步走向信息化；在发展趋势维度上，数字经济正在为世界经济发展注入新动能；在基础资源维度上，数据是定义世界的基础，是驱动数字经济引擎的燃料；在数据科学维度上，世界正处于新一轮科技革命和数字化转型的转折点，数据作为一种新型生产要素正推动产业升级和经济转型^[28]。

1. 数据要素推进政府数据开放共享

当今世界，数据遍布生活的各个领域，政府作为公共信息服务的主要持有者，掌握着大量社会基础数据，占比高达 80%。一般来说，数据量越大价值越高，这些数据具有全覆盖、专业性、权威性特征，有着极大的利用价值。政府数据开放共享最重要的目的就是便民，政府数据开放共享的宗旨是提供可供公众使用的原始数据资源，同时对分散于政府各部门的数据进行有效整合；政府数据开放共享最主要的步骤就是利用大数据技术搭建数据共享平台，通过平台搜集数据、整合数据进而以最完整、最方便、最有效的方式来向大众共享数据。搭建数据共享

平台最主要的有三方面：数据格式、数据定价以及权责界定；

首先对于数据格式来说，数据的种类多种多样，不同的数据有不同的格式，不能张冠李戴，否则数据会失去其本身价值，如图片也属于一种数据，它的对应格式就是 jpg；对于数据定价，数据主要分为一手数据和二手数据，一般来说，一手数据是调查员亲自走访、搜集、整理、分析所得出的数据，数据搜集过程复杂、搜集成本高、数据价值相对于二手数据较高，最重要的是该类数据一般都属于内部数据，一般不对外公开，部分对外公开的需要收取一定的费用。数据价格的确定应该依据供需双方的交易机制，如若供需双方数量较多且交易较频繁，可以由市场供求自行确定价格；当供求交易双方数量相对较少时，应该建议数据持有者率先坦白保留价格，然后结合数据特征制定合理价格；对于权责界定，如果数据信息一旦被公开、甚至被交易，那么数据就会有残留，被称为“缓存信息”，这些信息会被相关人士非法利用，私自对数据进行修改、删减，造成的结果就是相关部门会对其进行调查，但是由于权责主体不清晰，数据交易双方就会借此疏漏推诿责任。所以，数据权属应以新颁布的民法典为依据，确定数据间的权属关系，以此来保证交易合法合规、数据产权清晰。

2. 数据要素提升数据资源价值

如今数字经济的空前活跃离不开丰富数据资源的支持，数据要素已成为国家经济建设的关键。云计算、区块链、AI、5G 技术、量子计算等数字创新技术的蓬勃发展与数据要素息息相关，数据要素大大提升了数据资源的价值，使之成为经济活动中的重要资源。利用大数据技术，可以实现资源的快速优化配置。数字技术与传统技术的交叉融合，以及数字经济与实体经济的高度关联，有助于打造一个极富竞争力的数字强国。

3. 数据要素加强资源整合和安全保护

当今世界最不缺的是什么？是数据。数据量大的同时也面临着众多数据安全问题，个人、企业、政府、国家等都是数据的持有者，数据安全与否直接关系到其利益，所以数据是数字经济发展道路上不可忽视的方面。数据资源整合与安全保护需要建立统一规范的数据管理制度，制定数据隐私保护制度和审查制度。数据安全环节涉及多个方面：收集、存储、传输、使用等，传统的数据安全管理制度及手段已经不再适应当今数字经济时代。因此，要加强核心数据保护，最重要的就是提高核心关键技术安全监管系数，完善核心数据生命周期保护和管理。

（三）数据要素、数字经济与城镇贫困的关系

将数字经济比喻为人的身体，那么数据要素就是人体中的血液。较传统生产要素而言，数据要素具有可共享、可复制、可无限供给等特点。土地、资本、劳动这些传统生产要素并不能无限供给，所以对经济增长有所制约，这也就凸显出数据要素对经济增长具有倍增效应。

数据要素主要通过政府数据开放共享、提升数据资源价值及加强数据安全保护来助力经济发展，其中在提升数据资源价值方面，大力发展数字经济对经济增长最为关键。数字经济是当今世界各国首当其冲的经济形态，随着云计算、区块链、AI、5G 技术、量子计算等数字创新技术的发展，数字技术的创新发展促进了国家经济增长，所谓国富家才能富，家富人才富，人富起来，贫困就能有所减缓。

但一个硬币有正反两面，数字化的创新发展也会对经济有所制约，在城镇人口分布中，大部分人口是由中老年人与农村进城务工人员组成，他们对数字化技术了解不深、接触不多，接收数字化信息能力相对较差，这就导致了大部分城镇

人口遭遇了“数字贫困”。

数据要素与城镇贫困的关系见图 1。

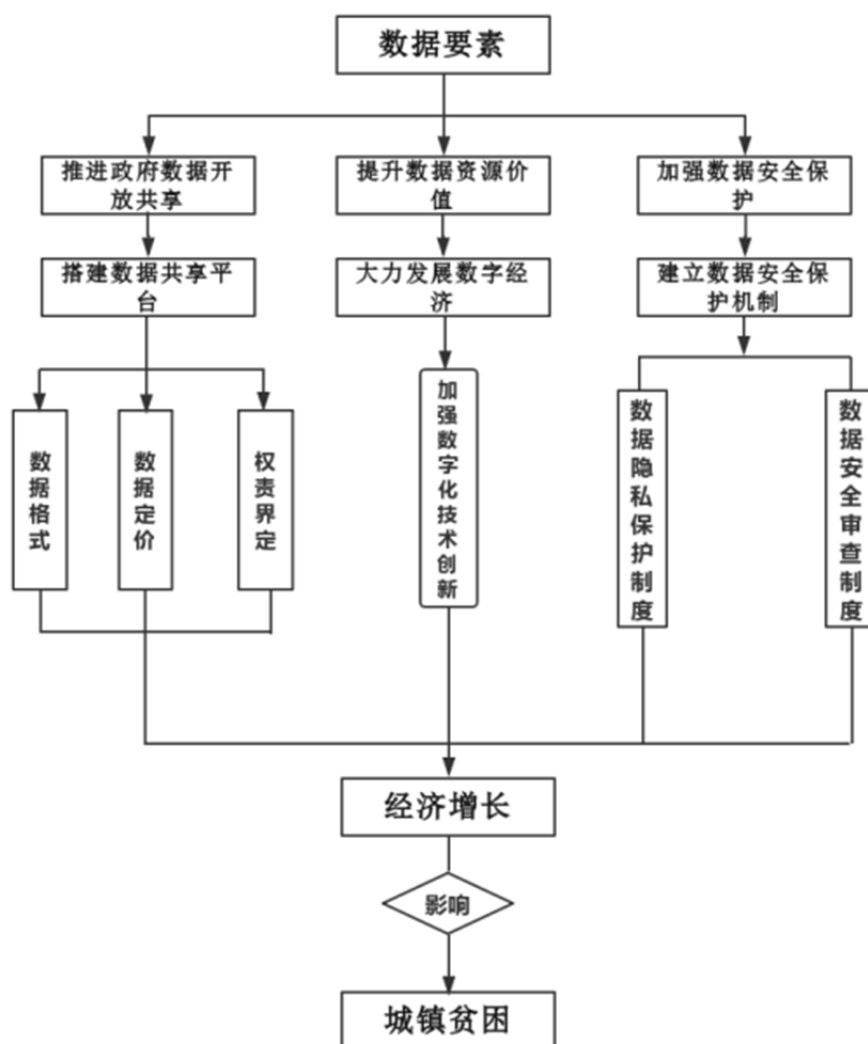


图 1 数据要素与城镇贫困的关系图

（四）研究方法

1. FGT 指数

度量贫困或是对贫困测度就是对贫困指数进行定义和计算，本文采用 FGT 指数，它由 Foster、Greer 和 Thorbecke (1984) 三位学者提出，用于计算社会内部的贫困广度、深度和强度^[25]。本文选用 FGT 指数测度我国城镇贫困是因为 FGT

指数可在确定的贫困线下精准计算我国城镇贫困的广度和深度。FGT 指数因社会贫困厌恶系数 α 取不同值而结果不同, 具体而言, H、PG、SPG 分别代表城镇贫困人口占总人口的比重、城镇贫困者收入与贫困线的差距、城镇贫困人口之间收入的不平等状况。

与基尼系数相比, 使用 FGT 指数测度山西省城镇贫困的优势在于: 基尼系数主要衡量收入不平等的程度, 反映不出城镇贫困者收入不平等的变动情况及贫困者收入与贫困标准的差距, 而 FGT 指数中的 PG 指标恰好能准确的计算出城镇贫困者的收入水平与收入贫困线的差距, SPG 指标能精准的计算出这些城镇贫困群体之间的收入差距。

FGT 指数有可分解性的优势, 与其他测度贫困的指数不同, 它将总人口数按比例分组, 并在此基础上计算各组的收入差距, 因此被世界各国广泛使用。

(1) 贫困线的确定

贫困线, 又称贫困标准, 是贫困测度的前提, 贫困线分为两种, 即相对贫困线及绝对贫困线。确定贫困线的方法有主观法、相对法、绝对法。采用主观法确定贫困线时, 被调查者的主观感受对贫困线结果影响很大, 因此本文未采用这一方法。相对法是将社会成员的平均收入或社会中等收入者收入的比例作为贫困线。OECD 将贫困线定为社会成员平均收入或中等收入的 50%, 美国和加拿大将贫困线定为社会成员平均收入或中等收入的 40%。绝对法基于人们对生活必需品支付能力。世界银行于 2015 将国际贫困线标准从之前的每人每天 1.25 美元上调至 1.9 美元。

(2) 洛伦兹曲线

贫困测度的 FGT 指数的基础是洛伦兹曲线, 1905 年, 美国统计学家洛伦兹提出洛伦兹曲线, 它可体现一国收入分配的差异。它先将一国的人口按照收入划

分为不同组别，然后再计算每组居民收入在国民收入中的比重，接着用二维图将人口累计百分比和收入累计百分比的关系展示出来，如图 2 所示。在图 2 中，纵轴 OA 代表收入累计百分比；横轴 OB 表示人口（按收入由低到高分组）累计百分比；45 度的 OC 线代表收入分配绝对平等线，该线上每组居民的收入相等；折线 OBC 代表所有的社会财富被高收入居民占有的极端情况，洛伦兹曲线（D 曲线）介于 OBC 和 OC 曲线之间。根据洛伦兹曲线的曲折度可反映每个收入阶层的差别。

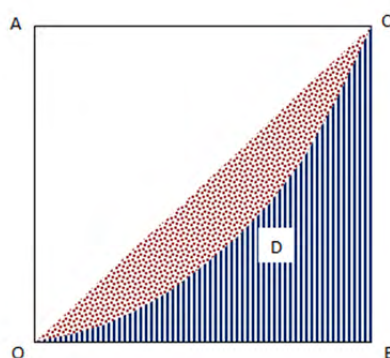


图 2 洛伦兹曲线

洛伦兹曲线的表达式为：

$$L = L(P, \Pi) \quad (1)$$

公式中，L 代表累计收入占总收入的百分比，P 代表各组人口占总人口的百分比， Π 代表待估向量。L 代表各组人口比重与收入比重间的函数关系。

洛伦兹曲线有多种函数形式，Villasenor(1989)提出 General Quadratic(GQ) Lorenz 曲线，Kakwani(1980)提出 Beta(B)Lorenz 曲线^[26,27]。这两种曲线的度量结果较好，两种模型的公式如下：

曲线 GQ 模型：

$$L(1-L)=a(P^2-L)+bL(P-1)+c(P-L) \quad (2)$$

Beta 模型：

$$L(P)=P-\theta P^{\gamma}(1-P)^{\delta} \quad (3)$$

(3) FGT 指数

FGT 指数的连续形式为：

$$P_{\alpha}=\int_0^z\left(\frac{z-x}{z}\right)^{\alpha}f(x)dx \quad (4)$$

式中, x 为家庭收入或消费支出, $f(x)$ 代表收入分布密度, z 为贫困线, α 反映贫困深度或强度的敏感程度, 当 $\alpha=1$ 时, FGT 指数代表贫困发生率指数 H , 当 $\alpha=2$ 时, FGT 指数代表贫困距指数 PG , 当 $\alpha=3$ 时, FGT 指数代表平方贫困距指数 SPG 。

贫困发生率

贫困发生率 H 是测度贫困程度最基本的 1 个指标, 应用广泛。该指标测算的是贫困人口占总人口的比例, 反映贫困的广度, 其计算公式为：

$$H=\frac{q}{n} \quad (5)$$

其中, n 代表总人口数, q 代表贫困人口数。该指标的确定贫困线息息相关, 即贫困线设置得较高, 则测算出的贫困发生率越高, 且其计算公式简单直观、实操性强。

贫困缺口率

贫困缺口率, 又称贫困距指数, 测量的是贫困者收入与贫困线的差距, 体现贫困线以下人口收入的变动, 以衡量贫困的深度。其计算公式：

$$PG = \sum_{i=1}^q \frac{z - y_i}{qz} \quad (6)$$

其中, q 是贫困人口数, z 是贫困线, y_i 是贫困人口的收入, $z - y_i (i = 1, 2, \dots, q)$ 表示贫困人口 i 与贫困线 z 之间的收入差距。 $0 \leq PG \leq 1$, PG 越大, 所反映的贫困越严重。

平方贫困距指数

平方贫困距指数反映贫困人口之间的收入不平等状况, 是对贫困强度的测度。计算公式如下:

$$SPG = \sum_{i=1}^q \frac{[(z - y_i)/z]^2}{q} \quad (7)$$

其中, z 为贫困线, y_i 为贫困人口的收入水平。平方贫困距指数赋予权重的方法是越穷权数越大, 在 H 、 PG 不变时, SPG 的值越大, 贫困群体内部的收入差距越大, 扶贫的难度也越大。

FGT 的计算需要依靠两种参数化的洛伦兹曲线, 即 GQ 洛伦兹模型、Beta 洛伦兹模型。依托 GQ 模型测算的 FGT 指数公式为:

$$L(P) = -\frac{1}{2} \left(bP + e + \sqrt{mp^2 + nP + e^2} \right) \quad (8)$$

$$H = -\frac{1}{2m} \left(n + r \frac{(b + 2z/u)}{\sqrt{(b + 2z/u)^2 - m}} \right) \quad (9)$$

$$PG = H - \frac{u}{z} L(H) \quad (10)$$

$$SPG = 2PG - H - \left(\frac{u}{z} \right)^2 \left[aH + bL(H) - \frac{r}{16} \ln \left(\frac{1 - H/S_1}{1 - H/S_2} \right) \right] \quad (11)$$

其中, L 代表累计收入百分比, P 代表累计人口百分比, z 为贫困线, u 为人均消费。

其中, $e = -(a + b + c + 1)$; $m = b^2 - 4a$; $n = 2be - 4c$; $r = \sqrt{(n^2 - 4me^2)}$;

$$S_1 = \frac{(r-n)}{(2m)}; S_2 = -\frac{(r+n)}{(2m)}。$$

依托 Bata 模型测算的 FGT 指数公式为：

$$\theta H^\gamma (1-H)^\delta \left[\frac{\gamma}{H} - \frac{\delta}{1-H} \right] = 1 - \frac{z}{u} \quad (12)$$

$$PG = H - (u/z)L(H) \quad (13)$$

$$SPG = \left(1 - \frac{u}{z}\right) \left[2PG - \left(1 - \frac{u}{z}\right)H \right] + \theta^2 \left(\frac{u}{z}\right)^2 \left[\begin{array}{l} \gamma^2 B(H, 2\gamma-1, 2\delta+1) \\ - 2\gamma\delta B(H, 2\gamma, 2\delta) \\ + \delta^2 B(H, 2\gamma+1, 2\delta-1) \end{array} \right] \quad (14)$$

其中， $B(k, r, s) = \int_0^k P^{r-1} (1-P)^{s-1} dP$ 。

2. 计量模型

构建计量模型，如公式（15）。

$$fgt = \alpha + \beta dig + \pi X + u \quad (15)$$

其中， fgt 为被解释变量，反映贫困程度，内含三个变量，分别为 h 、 pg 、 spg ，分别表示贫困的广度、深度、强度。 dig 为核心解释变量，表示数字经济的发展水平。 X 为控制变量，有 5 个控制变量，可提高解释变量对模型的解释程度。 α 为回归常数， β 为核心解释变量的回归系数， π 为控制变量的回归系数向量， u 为误差项。

四、 山西省城镇相对贫困测度

以往贫困测度研究多使用 FGT 指数，通过贫困的广度、深度及强度三方面来测度贫困。本研究也使用 FGT 指数，从历年《山西统计年鉴》获取 2003 年至 2019 年山西省城镇数据，使用 H 、 PG 和 SPG 指数测度山西省城镇相对贫困。

（一）数据说明

在对山西省城镇相对贫困进行测度时，数据来源于 2003 年至 2019 年《山西统计年鉴》。计算 FGT 指数要用到人口与收入数据、人均消费与贫困线数据，利用世界银行的 PovcalNet 在线分析得到 FGT 指数值。

1. 人口与收入数据

人口与收入数据用于拟合洛伦兹曲线，以 2019 年为例，人口与收入数据如表 1。

表 1 2019 年人口与收入

| 收入区间 | 占调查户比重 (%) | 各组均值 (元) | 人口累计比 P (%) | 收入比 (%) | 收入累计比 L (%) |
|-------|---------------|-------------|----------------|------------|----------------|
| 低收入户 | 20 | 13869.79 | 20 | 8.03 | 8.03 |
| 中低收入户 | 20 | 23152.53 | 40 | 13.40 | 21.43 |
| 中等收入户 | 20 | 31313.89 | 60 | 18.12 | 39.55 |
| 中高收入户 | 20 | 41311.46 | 80 | 23.91 | 63.46 |
| 高收入户 | 20 | 63121.19 | 100 | 36.54 | 100 |
| 合计 | 100 | 172768.86 | — | 100 | — |

表 1 第一列为 5 个收入区间的名称，2019 年山西省城镇居民按可支配收入分组，分为低收入户、中低收入户、中等收入户、中高收入户和高收入户 5 组。第二列为各组占调查户的比重，2019 年山西省城镇居民按照等距五分法分组，每组占调查户的比重均为 20%。2003 年至 2019 年人口与收入数据的分组来源于《山西统计年鉴》，各年的分组方法并非均为等距五分法，为方便后续分析，统一分为 5 组，部分年份的分组数据对《山西统计年鉴》中的数据处理后得到。第三列为 2019 年山西省城镇居民各组可支配收入的均值。第四列为各组占调查户比重的累积值，称为人口累计比（P）。第五列为各组可支配收入占总可支配收入的比重，称为收入比。第六列为各组可支配收入占总可支配收入比重的累积值，称为收入累积比（L）。

2. 人均消费与贫困线数据

人均消费与贫困线数据用于计算 FGT 指数。李冬（2016）表示，比起收入数据，消费数据与家庭福利更为相关，因此选用生活性消费数据测度相对贫困^[17]。因此，本文的贫困线设定为山西省城镇居民人均消费支出的 50%。2003 年至 2019 年的人均消费与贫困线数据如表 2。

表 2 人均消费与贫困线

| 年份 | CPI 平减后的人均消费 u (元) | CPI 平减后的贫困线 z (元) |
|------|----------------------|---------------------|
| 2003 | 5210.00 | 2605.00 |
| 2004 | 5588.80 | 2794.40 |
| 2005 | 6186.33 | 3093.16 |
| 2006 | 6894.60 | 3447.30 |
| 2007 | 7502.51 | 3751.25 |
| 2008 | 7647.22 | 3823.61 |
| 2009 | 8232.95 | 4116.47 |
| 2010 | 8386.03 | 4193.01 |
| 2011 | 9252.04 | 4626.02 |
| 2012 | 9717.27 | 4858.63 |
| 2013 | 10171.83 | 5085.92 |
| 2014 | 10626.50 | 5313.25 |
| 2015 | 11416.14 | 5708.07 |
| 2016 | 12129.96 | 6064.98 |
| 2017 | 12955.78 | 6477.89 |
| 2018 | 13685.14 | 6842.57 |
| 2019 | 14261.04 | 7130.52 |

表 2 中的人均消费 (u) 与贫困线 (z) 数据皆为 CPI 平减后的数据，人均消费数据来自 2003 年至 2019 年《山西统计年鉴》，贫困线设定为人均消费的 50%。

(二) 山西省城镇相对贫困测度

1. 洛伦兹曲线拟合

根据人口累计比 (P) 与收入累积比 (L) 数据，利用世界银行的 PovcalNet 在线分析得到洛伦兹曲线拟合结果，拟合结果如表 3。

表 3 GQ 洛伦兹曲线拟合

| 年份 | a | b | c | R-squared | 基尼系数(%) |
|------|------|-------|------|-----------|---------|
| 2003 | 1.04 | -1.53 | 0.17 | 0.99 | 28.40 |
| 2004 | 1.01 | -1.55 | 0.14 | 0.99 | 28.93 |
| 2005 | 1.09 | -1.48 | 0.19 | 0.99 | 28.98 |
| 2006 | 1.07 | -1.54 | 0.16 | 0.99 | 28.62 |
| 2007 | 0.98 | -1.55 | 0.12 | 0.99 | 29.58 |
| 2008 | 1.03 | -1.47 | 0.17 | 0.99 | 29.96 |
| 2009 | 0.99 | -1.43 | 0.20 | 0.99 | 29.82 |
| 2010 | 1.03 | -1.52 | 0.15 | 0.99 | 29.30 |
| 2011 | 1.06 | -1.53 | 0.13 | 1.00 | 29.96 |
| 2012 | 1.10 | -1.49 | 0.16 | 0.99 | 30.01 |
| 2013 | 1.08 | -1.40 | 0.16 | 0.99 | 31.87 |
| 2014 | 1.15 | -1.40 | 0.19 | 0.99 | 30.88 |
| 2015 | 1.16 | -1.44 | 0.19 | 0.99 | 30.35 |
| 2016 | 1.16 | -1.47 | 0.18 | 0.99 | 30.03 |
| 2017 | 1.16 | -1.49 | 0.18 | 0.99 | 29.36 |
| 2018 | 1.17 | -1.50 | 0.19 | 0.99 | 29.02 |
| 2019 | 1.17 | -1.52 | 0.19 | 0.99 | 28.55 |

由表 3 可知，分析可得到 2003 年至 2019 年洛伦兹曲线的参数值 a,b,c，拟合 GQ 洛伦兹曲线。拟合的 R-squared 接近 1，曲线拟合结果极佳。基尼系数在 28%至 32%之间，洛伦兹曲线弧度较大，说明山西省城镇居民收入差距较大，收入分配失衡，推测山西省城镇相对贫困程度较高。

2. FGT 指数的变化趋势

世界银行的 PovcalNet 在线分析得到 2013 年至 2019 年的 FGT 指数，H、PG 和 SPG 值的变化趋势如图 3。

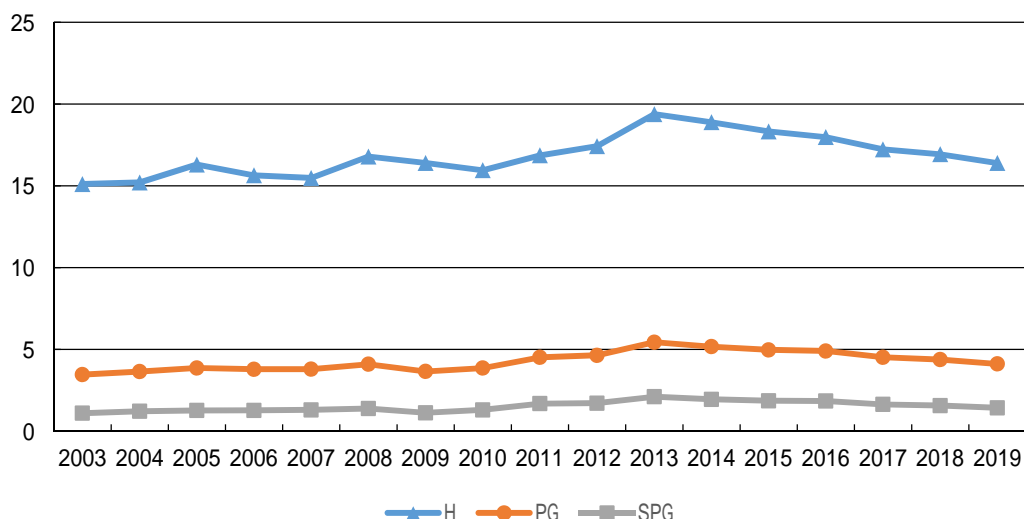


图 3 2003 年至 2019 年山西省城镇居民 FGT 指数变化趋势图

由图 3 可知，3 个 FGT 指数的变化趋势基本一致，2003 年至 2013 年，指数值起伏不定，大致呈上升趋势，2013 年达到最高，期间山西省城镇相对贫困不断加剧。习近平总书记于 2013 年首次提出“精准扶贫”的概念，政府同步推进各种扶贫政策。2013 年之后，FGT 指数值开始下降，说明山西省城镇相对贫困程度递减，扶贫政策发挥了作用。2013 年至 2019 年，3 个 FGT 指数逐年下降，随着减贫工作的继续开展，可以预见山西省城镇相对贫困程度将会继续减弱。

五、 山西省城镇相对贫困与数字经济关系的实证研究

从历年《山西统计年鉴》和国家统计局官方网站获取 2003 年至 2019 年山西省城镇数据，构建计量模型，探究影响山西省城镇相对贫困的诸多因素，尤其是数字经济的发展对其的影响。

（一） 数据说明

构建计量模型时用到的数据来源于 2003 年至 2019 年《山西统计年鉴》和国家统计局官方网站。

计量模型中的被解释变量、核心解释变量、控制变量如表 4。

表 4 变量描述性统计

| 变量类型 | 变量 | 指标 | 样本量 | 平均值 | 标准差 | 最小值 | 最大值 |
|--------|-----|----------------------|-----|--------------|------|--------|--------|
| 被解释变量 | h | H 指数 | 17 | 16.83 | 1.2 | 15.10 | 19.38 |
| | pg | PG 指数 | 17 | 4.27 | 0.59 | 3.45 | 5.43 |
| | spg | SPG 指数 | 17 | 1.51 | 0.30 | 1.09 | 2.1 |
| 核心解释变量 | dig | 数字经济 | 17 | 9.87 e-18 | 1.41 | -2.39 | 2.06 |
| 控制变量 | gdp | 人均 GDP 增长率 | 17 | 12.00 | 9.77 | -2.59 | 25.22 |
| | fin | 一般公共预算支出占 GDP 比重 | 17 | 21.82 | 3.72 | 16 | 27.67 |
| | une | 城镇登记失业率 | 17 | 3.29 | 0.27 | 2.70 | 3.90 |
| | gen | 性别比 | 17 | 104.61 | 1.67 | 102.27 | 108.34 |
| | pat | 发明专利申请受理量占专利申请受理量的比重 | 17 | 34.42 | 4.14 | 26.56 | 40.97 |

1. 被解释变量数据

本研究是对山西省城镇相对贫困的研究，故被解释变量为测度贫困的 FGT 指数。3 个变量为 h，pg 和 spg，使用 H 指数、PG 指数和 SPG 指数值作为被解释变量的数据。3 个被解释变量的样本量均为 17，即数据时间跨度为 2003 年至 2019 年，共 17 年。

2. 核心解释变量数据

计量模型的核心解释变量是 dig，代表数字经济的发展水平。从国家统计局官方网站收集到关于数字经济发展的 3 个指标的数据，3 个指标分别为移动电话交换机容量（万户）、长途光缆线路长度（万千米）、技术市场成交额（亿元）。

对这 3 个指标进行主成分分析，首先利用 R 软件得到 3 个指标的标准化值，之后进行相关分析，得到表 5。

表 5 标准化后的数字经济指标相关系数

| 指标 | sx1 | sx2 | sx3 |
|-----|-----------|-----------|-----|
| sx1 | 1 | | |
| sx2 | 0.8921474 | 1 | |
| sx3 | 0.8049749 | 0.6520111 | 1 |

由表 5 可知，这 3 个指标的相关性很高，适合提取主成分。

R 软件输出的各个主成分的贡献率如表 6。

表 6 主成分方差贡献率

| 主成分 | 标准差 | 方差贡献率 (%) | 累积方差贡献率 (%) |
|-----|-----------|-------------|---------------|
| 1 | 1.603177 | 85.67255 | 85.67255 |
| 2 | 0.5967561 | 11.87059 | 97.54315 |
| 3 | 0.2714877 | 2.456852 | 100 |

由表 6 可知，前两个主成分的累积方差贡献率达到 97.54315%，故提取前两个主成分，并以两个主成分的方差贡献率为权重，计算第一、二主成分的加权平均值，即为数字经济指标值。核心解释变量的样本量为 17，即数据时间跨度为 2003 年至 2019 年，共 17 年。

3. 控制变量数据

计量模型中包括 5 个控制变量，分别为 gdp ， fin ， une ， gen ， pat ，其对应的指标分别为人均 GDP 增长率、一般公共预算支出占 GDP 比重、城镇登记失业率、性别比、发明专利申请受理量占专利申请受理量的比重，这 5 个控制变量与城镇相对贫困有极大相关度。控制变量的样本量均为 17，即数据时间跨度为 2003 年至 2019 年，共 17 年。

(二) 山西省城镇相对贫困与数字经济关系研究

1. 仅有核心解释变量的计量模型

模型 1、2、3 分别是 h 、 pg 、 spg 作被解释变量的模型，解释变量仅为代表数字经济发展水平的 dig ，3 个模型的回归结果如表 7。

表 7 仅有核心解释变量的计量模型回归结果

| 模型 | 1 | 2 | 3 |
|-----------|-------------------------|---------------------------|---------------------------|
| FGT 指数 | h | pg | spg |
| dig | 0.6137*** (0.1643) | 0.29619*** (0.07605) | 0.15023*** (0.04013) |
| Intercept | 16.8375*** (0.226) | 4.27766*** (0.10459) | 1.51047*** (0.05519) |

| | | | |
|--------------------|--------|--------|--------|
| Adjusted R-squared | 0.4473 | 0.4697 | 0.4486 |
|--------------------|--------|--------|--------|

注：表中括号内的数据为标准误差，***，**和*分别代表 1%，5%和 10%的显著性水平

由表 7 可知，回归模型中 dig 的系数显著为正，这说明数字经济的发展水平对贫困的广度、深度和强度都有十分显著的正向影响，山西省城镇数字经济的发展会加剧城镇相对贫困。数字经济的活跃导致了数字贫困的加剧，大量贫困人群跟不上时代的发展，被时代的浪潮甩在远方，这必然会导致贫困的加剧。对贫困人群进行数字教育，消除数字贫困，是减贫的重要举措。

由于解释变量单一，3 个模型调整的 R 方均不足 0.5，这表示单一解释变量 dig 对模型的解释效果有限，需在模型中引入控制变量。

2. 加入控制变量的计量模型

为解决模型解释度低的问题，在模型中加入控制变量，模型 4、5、6 分别是 h、pg、spg 作被解释变量的模型。加入控制变量后的模型回归结果如表 8。

表 8 加入控制变量的计量模型回归结果

| 模型 FGT 指数 | 4 h | 5 pg | 6 spg |
|--------------------|----------------------------|----------------------------|----------------------------|
| dig | 1.00549*** (0.28652) | 0.38669** (0.16602) | 0.180675* (0.095173) |
| gdp | -0.12329*** (0.02392) | -0.04495*** (0.01386) | -0.02054** (0.007946) |
| fin | -0.42702** (0.13615) | -0.14355* (0.07889) | -0.06343 (0.045224) |
| une | -2.97172*** (0.90385) | -1.6069** (0.52374) | -0.84793** (0.300235) |
| gen | 0.27211** (0.09932) | 0.14362** (0.05755) | 0.077041** (0.03299) |
| pat | 0.21758*** (0.05626) | 0.10822*** (0.0326) | 0.055777** (0.018689) |
| Intercept | 1.47075 (9.22129) | -5.50566 (5.34332) | -4.04495 (3.063058) |
| Adjusted R-squared | 0.8119 | 0.717 | 0.6527 |

注：表中括号内的数据为标准误差，***，**和*分别代表 1%，5%和 10%的显著性水平

由表 8 可知，数字经济对贫困的广度、深度和强度都有显著的影响，且系

数显著为正,说明山西省城镇数字经济的发展会提高城镇相对贫困程度,这是由于数字贫困会使贫困人群与时代脱节,非但不能享受时代红利,反而愈加贫困,所以需加强对贫困人群的数字教育,使之摆脱数字贫困。

模型 4 和模型 5 中的 5 个控制变量均显著,说明山西省城镇人均 GDP 增长率、一般公共预算支出占 GDP 比重、城镇登记失业率、性别比和发明专利申请受理量占专利申请受理量的比重均对城镇相对贫困的广度和深度有显著影响。模型 6 中的控制变量仅有 fin 不显著,说明山西省城镇一般公共预算支出占 GDP 比重对城镇相对贫困的强度没有显著影响。

控制变量 gdp, fin 和 une 的系数均为负值。山西省城镇人均 GDP 增长过快会显著提升城镇居民的生活水平,降低贫困的广度、深度和强度。一般公共预算支出占 GDP 比重增加会改善居民的生活水平,降低贫困的广度和深度。将失业率这一变量引入模型回归时使用的指标为城镇登记失业率,这一指标与实际失业率并不完全相同,城镇登记失业率系数显著为负,并不代表单纯地提高山西省城镇失业率就可解决城镇相对贫困问题。

控制变量 gen 和 pat 的系数均为正值。山西省城镇性别比失衡会导致一系列经济和社会问题,进而加剧城镇相对贫困,控制性别比例势在必行。山西省城镇发明专利申请受理量占专利申请受理量的比重代表科技创新的水平,科技创新水平越高,贫困人群越与时代脱节,进而造成社会财富分配不均,加剧城镇相对贫困,对贫困人群的科技教育必须重视起来。

加入控制变量后 3 个模型调整的 R 方均高于 0.6,解释变量对模型的解释效果较强,模型拟合较为合理。

六、 结论与建议

从历年《山西统计年鉴》和国家统计局官网获取 2003 年至 2019 年山西省城镇数据，通过构建 FGT 指数测度山西省城镇相对贫困程度，建立计量模型探究数字经济的发展与减贫的关系。

（一） 研究结论

1. 山西省城镇相对贫困变化趋势

从历年《山西统计年鉴》获取人口、收入、消费数据，数据跨度为 2003 年至 2019 年，利用世界银行的 PovcalNet 在线分析得到 3 个 FGT 指数 H、PG 和 SPG 的值，分别测度贫困的广度、深度和强度。

从历年 FGT 指数值可知，2003 年至 2013 年，指数值起伏不定，大致呈上升趋势，2013 年达到最高，期间山西省城镇相对贫困不断加剧。2013 年之后，指数值开始下降，说明山西省城镇相对贫困程度递减，扶贫政策发挥了作用。2013 年至 2019 年，3 个 FGT 指数逐年下降，减贫工作持续稳定进行，可以预见山西省城镇相对贫困程度将会继续减弱。

2. 数字经济的发展加剧山西省城镇相对贫困

从历年《山西统计年鉴》和国家统计局官方网站获取解释变量的数据，数据跨度为 2003 年至 2019 年。

数字经济对城镇相对贫困的广度、深度和强度都有十分显著的正向影响，山西省数字经济的发展会提高城镇相对贫困程度。当今世界风云变化，我国数字经济发展态势较好，数字经济的活跃导致数字贫困的加剧，大量贫困人群跟不上时代的发展，被时代的浪潮甩在远方，这必然会加剧贫困。

（二）建议

1. 稳步推进扶贫政策，消除相对贫困

习近平总书记发表 2020 年新年贺词时强调，2020 年是具有里程碑意义的一年，是决战脱贫攻坚的关键一年，至 2020 年底，我国已消除绝对贫困，全面建成小康社会，实现第一个百年奋斗目标^[29]。接下来便是与相对贫困作战的漫长时期，第七次人口普查数据显示，我国人口数仍在持续增长，作为一个人口大国，减轻相对贫困无疑是一大挑战。实证研究表明我国的扶贫政策是正确且有效的，依据当前国情稳步推进扶贫政策消除相对贫困，实现共同富裕，是我国的一个长期目标。

2. 减轻数字贫困，加强数字教育

身处大数据时代，我国数字经济空前活跃，数据要素作为新型生产要素，为经济发展注入新动能，提高了我国综合国力。但数字经济的极速发展也会带来其他问题，时代在发展，但贫困人群却被时代甩在身后，他们不能享受数字化时代的便利，反而要承担与时代脱节的痛苦。且我国人口老龄化趋势日渐严重，老年人口比重不断提高，老年人作为社会弱势群体，他们难以适应社会的巨变。作为社会大比重群体，老年群体的生活质量并没有随着数字经济的发展而迅猛提高，他们为国家贡献了一生，在人生的最后一段路上却没能享受时代红利。尤其对城镇居民而言，收入分配愈加失衡，底层人群与老年人群被飞速发展的时代抛弃，数字贫困问题日趋严重。想要消除城镇相对贫困，消除数字贫困，对贫困人群进行数字教育势在必行。

3. 保障义务教育，消除贫困文化

2006 年 7 月 1 日，九年义务教育正式施行，至今我国已推行义务教育政策

将近 15 年。义务教育政策帮助了大批的贫困学子完成基础教育，为他们进入社会奠定了坚实的基础。但是教育对于贫困家庭而言是个奢侈品，贫困家庭的父母或因为家庭经济窘况让子女放弃义务教育，或因为贫困文化的操纵迫使孩子提早进入社会。尽管义务教育已经普及多年且深入人心，但客观与主观的因素仍会使适龄的在学儿童离开教室，扛起生活的重担。教育的缺失导致了很多社会问题，例如贫困群体中父代与子代间周而复始的弱势地位，金字塔型收入结构，人民生活水平不能持续提升。这些问题都会影响中国的经济建设与社会建设，从而加剧贫困问题，尤其是对贫富分化较为严重的城镇而言。故而，保障义务教育，消除贫困文化，开阔贫困群体的眼界，对减轻城镇相对贫困有重大作用。

参考文献

- [1] 何玉长,王伟.数据要素市场化的理论阐释[J].当代经济研究,2021(04):33-44.
- [2] 中共中央网络安全和信息化领导小组办公室 . 二十国集团数字经济发展与合作倡议 [EB/OL] . (2016 - 09 - 29) [2020 - 06 - 20] . [http://www . cac . gov . cn/2016 - 09 /29 /c_1119648520 . htm](http://www.cac.gov.cn/2016-09/29/c_1119648520.htm) .
- [3] 王芳.关于数据要素市场化配置的十个问题[J].图书与情报,2020(03):9-13.
- [4]杨锐. 培育数据要素市场的关键:数据供给的市场化[J]. 图书与情报, 2020, No.193(03):33-34.
- [5] 郭琨,王磊.科学认识数据要素的技术经济特征及市场属性[J].中国物价,2021(05):12-14+26.
- [6]唐要家,唐春晖. 数据要素经济增长倍增机制及治理体系[J]. 人文杂志, 2020, No.295(11):88-97.
- [7] Mingione E. The New Relative urban poverty and the Underclass: Introduction [J] , International Journal of Urban andRegion Research.1993,(07).
- [8] HILARY, SILVER. National Conceptions of the New Relative urban poverty: Social Structural Change in Britain, France and the United States[J]. International Journal of Urban and Regional Research, 1993, 17(3):336-354.
- [9] Bourguignon F . Journal of Economic Inequality 1: 25--49, 2003. 2003.
- [10] Foster A J . Counting and multidimensional poverty measurement[J]. Journal of Public Economics, 2011.
- [11] J Knight, L Yueh. Urban Insiders versus Rural Outsiders: Complementarity or Competition in China`s Urban Labour Market?[C]// University of Oxford,

Department of Economics. University of Oxford, Department of Economics, 2004.

- [12] Beccaria L , Maurizio R , Fernandez A L , et al. Relative urban poverty and labor market dynamics in five Latin American countries: 2003–2008[J]. Journal of Economic Inequality, 2013, 11(4):555-580.
- [13] Xin M , Gregory R , Wang Y . Poverty, Inequality, and Growth in Urban China, 1986-2000[J]. Journal of Comparative Economics, 2005, 33(4):710-729.
- [14] 黄颖.试论中国城镇反贫困政策的实践及改进[J].社科纵横（新理论版），2012，27（04）:62-63.
- [15] 李东坤,郑浩生,张晓玲.交通基础设施建设能减缓城镇相对贫困吗?[J].河海大学学报(哲学社会科学版),2021,23(01):89-97+108.
- [16] 梁汉媚,方创琳.中国城市贫困人口动态变化与空间分异特征探讨[J].经济地理,2011,31(10):1610-1617.
- [17] 李冬. 基于 FGT 指数的城市化背景下城镇相对贫困问题研究[D].云南财经大学,2016.
- [18] 单德朋, 郑长德. 民族地区城镇相对贫困动态演化与经济增长益贫性研究[J]. 西北人口, 2014(01):79-85.
- [19] 张建华. 中国经济转型发展中的城镇相对贫困问题[C]// 中华外国经济学说研究会学术讨论会. 2006.
- [20] 张冰子, 贾琬, 申广军.城镇相对贫困的特征演变[J].统计研究, 2019, 36（02）:11-22.
- [21] 陈宗胜, 于涛. 中国城镇相对贫困线、贫困率及存在的问题[J]. 经济社会体制比较, 2017, No.194(06):40-53.

- [22] 孙娜, 张梅青, 陶克涛. 交通基础设施对民族地区经济增长的影响——兼论民族地区高铁建设[J]. 中央民族大学学报(哲学社会科学版), 2019(1).
- [23] 陈书. 中国城镇相对贫困估计及原因探索[J]. 贵州财经大学学报, 2015 (06) :99-108.
- [24] 侯卉, 王娜, 王丹青. 中国城镇多维贫困的测度[J]. 城市发展研究, 2012, 19 (12) :123-128.
- [25] Foster J, Greer J, Thorbecke E. A Class of Decomposable Poverty Measures[J]. Econometrica, 1984, 52(3).
- [26] Villaseñor José A., Arnold Barry C.. Elliptical Lorenz curves[J]. North-Holland, 1989, 40(2).
- [27] Kakwani N. On a Class of Poverty Measures[J]. Econometrica, 1980, 48(2).
- [28] 吴志刚. 从多个维度理解数据要素[N]. 中国电子报, 2020-06-05(004).
- [29] 习近平总书记的重要论述[J]. 当代江西, 2020(01):30.

附 录

附录 1 研究数据

附表 1 核心解释变量原始指标数据

| 年份 | 移动电话交换机容量 (万户) | 长途光缆线路长度 (万千米) | 技术市场成交额 (亿元) |
|------|-------------------|-------------------|-----------------|
| year | x1 | x2 | x3 |
| 2003 | 698.2 | 1.66 | 3.23 |
| 2004 | 942.4 | 2.01 | 6 |
| 2005 | 1088.2 | 2.13 | 4.8 |
| 2006 | 1457.3 | 2.25 | 5.92 |
| 2007 | 1867.4 | 2.29 | 8.27 |
| 2008 | 2363.9 | 2.61 | 12.84 |
| 2009 | 4223.5 | 2.61 | 16.21 |
| 2010 | 3772.5 | 2.77 | 18.49 |
| 2011 | 4477.2 | 2.76 | 22.48 |
| 2012 | 4648.9 | 2.78 | 30.61 |
| 2013 | 5255.8 | 3.61 | 52.77 |
| 2014 | 5021.16 | 2.98 | 48.46 |
| 2015 | 5021.16 | 3.09 | 51.2 |
| 2016 | 5096.49 | 3.16 | 42.56 |
| 2017 | 5562.5 | 3.24 | 94.15 |
| 2018 | 6090.83 | 3.01 | 150.76 |
| 2019 | 7359.2 | 3.11 | 109.52 |

附表 2 控制变量指标数据

| 年份 | 人均 GDP 增长率 (%) | 一般公共预算支出 占 GDP 比重(%) | 城镇登 记失业 率(%) | 性别比 | 发明专利申请 受理量占 专利申请受理 量的比重(%) |
|------|----------------------|-------------------------|--------------------|---------|-------------------------------------|
| year | gdp | fin | une | gen | pat |
| 2003 | 21.98531488 | 16.9212163152324 | 3 | 105.01 | 3500.333333 |
| 2004 | 21.71547633 | 17.0607711017618 | 3.1 | 105.44 | 3401.290323 |
| 2005 | 15.97717546 | 16.0007369302548 | 3 | 102.97 | 3432.333333 |
| 2006 | 14.86674867 | 19.2650142030510 | 3.2 | 102.27 | 3195.9375 |
| 2007 | 25.22844089 | 18.3123940419297 | 3.2 | 103.87 | 3245.9375 |
| 2008 | 21.04663094 | 18.6373976019728 | 3.3 | 102.93 | 3119.090909 |
| 2009 | -1.544692474 | 21.2237160757240 | 3.9 | 104.22 | 2672.307692 |
| 2010 | 21.65885392 | 20.9910345727048 | 3.6 | 104.575 | 2904.861111 |
| 2011 | 19.52504522 | 21.0351640919769 | 3.5 | 104.93 | 2998 |
| 2012 | 6.694078947 | 22.7813404002378 | 3.3 | 104.46 | 3165.454545 |
| 2013 | 2.084168337 | 24.0442145022298 | 3.1 | 105.52 | 3403.870968 |
| 2014 | 0.38053819 | 24.1766250048975 | 3.4 | 103.3 | 3038.235294 |
| 2015 | -2.593495201 | 26.8123604746798 | 3.5 | 107.49 | 3071.142857 |
| 2016 | 0.466409266 | 26.4449106137496 | 3.5 | 106.04 | 3029.714286 |
| 2017 | 20.61735227 | 24.1905154426728 | 3.4 | 108.34 | 3186.470588 |
| 2018 | 9.629893964 | 25.4719617554896 | 3.3 | 104.73 | 3173.636364 |
| 2019 | 6.310160428 | 27.6671345592511 | 2.7 | 102.32 | 3789.62963 |

附表 3 被解释变量指标数据

| 年份 | H 指数 | PG 指数 | SPG 指数 |
|------|---------|---------|---------|
| year | h | pg | spg |
| 2003 | 15.1016 | 3.45226 | 1.09391 |
| 2004 | 15.2019 | 3.63853 | 1.2157 |
| 2005 | 16.2955 | 3.8619 | 1.26138 |
| 2006 | 15.6376 | 3.78197 | 1.26926 |
| 2007 | 15.4747 | 3.78645 | 1.29812 |
| 2008 | 16.7859 | 4.08783 | 1.3776 |
| 2009 | 16.3929 | 3.65085 | 1.11934 |
| 2010 | 15.951 | 3.85607 | 1.29508 |
| 2011 | 16.8614 | 4.50999 | 1.68359 |
| 2012 | 17.4234 | 4.63202 | 1.70678 |
| 2013 | 19.3803 | 5.43414 | 2.10424 |
| 2014 | 18.8868 | 5.16512 | 1.94232 |
| 2015 | 18.323 | 4.96522 | 1.85313 |
| 2016 | 17.975 | 4.89663 | 1.84118 |
| 2017 | 17.228 | 4.51625 | 1.63428 |
| 2018 | 16.927 | 4.37302 | 1.55888 |
| 2019 | 16.3918 | 4.11194 | 1.42323 |

附录 2 R 代码

```
> #数字经济主成分分析

> #载入数字经济指标数据

> x1<-c(698.2,942.4,1088.2,1457.3,1867.4,2363.9,4223.5,3772.5,4477.2,4648.9,
5255.8,5021.16,5021.16,5096.49,5562.5,6090.83,7359.2)

> x2<-c(1.66,2.01,2.13,2.25,2.29,2.61,2.61,2.77,2.76,2.78,3.61,2.98,3.09,3.16,3.2
4,3.01,3.11)

> x3<-c(3.23,6,4.8,5.92,8.27,12.84,16.21,18.49,22.48,30.61,52.77,48.46,51.2,42.
56,94.15,150.76,109.52)

> #数字经济指标标准化

> sx1<-scale(x1,center=T,scale=T)

> sx2<-scale(x2,center=T,scale=T)

> sx3<-scale(x3,center=T,scale=T)

> #数据格式从变量转化为数据框

> data1<-data.frame(sx1,sx2,sx3)

> data1
```

| | sx1 | sx2 | sx3 |
|---|-------------|-------------|-------------|
| 1 | -1.54213213 | -2.07206016 | -0.86917795 |
| 2 | -1.42151532 | -1.38137344 | -0.80351832 |
| 3 | -1.34950086 | -1.14456656 | -0.83196292 |
| 4 | -1.16719266 | -0.90775969 | -0.80541462 |
| 5 | -0.96463347 | -0.82882406 | -0.74971060 |
| 6 | -0.71939905 | -0.19733906 | -0.64138406 |

```

7   0.19910635 -0.19733906 -0.56150212
8   -0.02365442  0.11840344 -0.50745737
9    0.32441546  0.09866953 -0.41287905
10   0.40922261  0.13813734 -0.22016684
11   0.70898650  1.77605156  0.30511024
12   0.59309163  0.53281547  0.20294669
13   0.59309163  0.74988844  0.26789521
14   0.63029910  0.88802578  0.06309404
15   0.86047370  1.04589703  1.28597510
16   1.12142980  0.59201719  2.62784944
17   1.74791114  0.78935625  1.65030311

```

```
> #数字经济指标相关分析
```

```
> cor_data1<-cor(data1)
```

```
> cor_data1
```

```

              sx1      sx2      sx3
sx1 1.0000000 0.8921474 0.8049749
sx2 0.8921474 1.0000000 0.6520111
sx3 0.8049749 0.6520111 1.0000000

```

```
> #数据经济指标主成分分析
```

```
> pr<-princomp(data1,cor=T)
```

```
> #数字经济指标主成分分析结果
```

```
> summary(pr,loadings=T)
```

```
Importance of components:
```

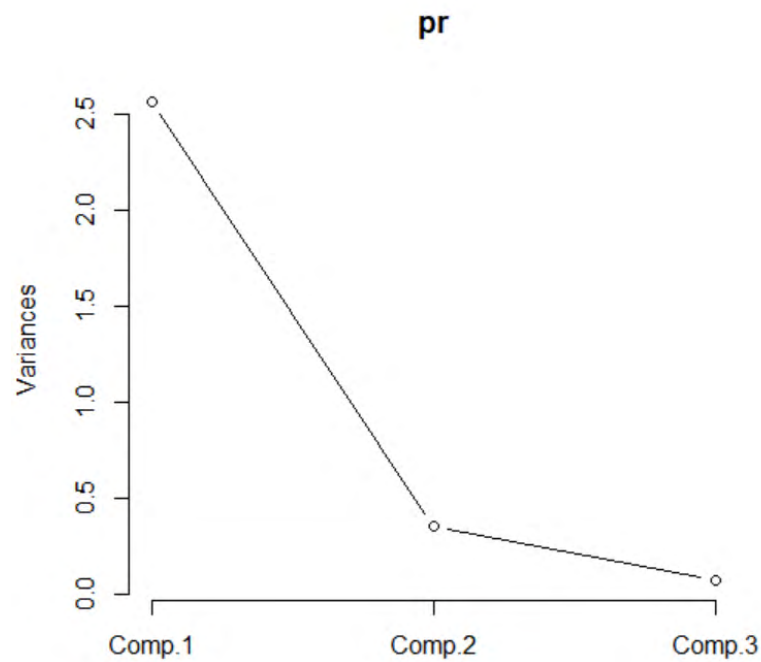
| | Comp.1 | Comp.2 | Comp.3 |
|------------------------|-----------|-----------|------------|
| Standard deviation | 1.6031770 | 0.5967561 | 0.27148770 |
| Proportion of Variance | 0.8567255 | 0.1187059 | 0.02456852 |
| Cumulative Proportion | 0.8567255 | 0.9754315 | 1.00000000 |

Loadings:

| | Comp.1 | Comp.2 | Comp.3 |
|-----|--------|--------|--------|
| sx1 | 0.608 | 0.130 | 0.784 |
| sx2 | 0.573 | 0.611 | -0.546 |
| sx3 | 0.550 | -0.781 | -0.297 |

> #绘制数字经济指标主成分碎石图

> screeplot(pr,type="lines")



> #数字经济指标主成分预测值

> pca_data1<-predict(pr)

```
> pca_data1
```

| | Comp.1 | Comp.2 | Comp.3 |
|-------|-------------|-------------|--------------|
| [1,] | -2.68291156 | -0.81156469 | 0.186052704 |
| [2,] | -2.16193166 | -0.41331120 | -0.125196528 |
| [3,] | -1.99297573 | -0.23164606 | -0.191558295 |
| [4,] | -1.72379305 | -0.07949279 | -0.185657429 |
| [5,] | -1.51872430 | -0.04752180 | -0.083498677 |
| [6,] | -0.93052233 | 0.29577522 | -0.273837745 |
| [7,] | -0.31004193 | 0.35436020 | 0.443610681 |
| [8,] | -0.23230591 | 0.47989539 | 0.069512653 |
| [9,] | 0.02759264 | 0.43790275 | 0.332822234 |
| [10,] | 0.21320498 | 0.31897497 | 0.320165220 |
| [11,] | 1.66663312 | 0.96774465 | -0.519903533 |
| [12,] | 0.80133440 | 0.25153131 | 0.117195865 |
| [13,] | 0.96643375 | 0.33595405 | -0.024799480 |
| [14,] | 0.95536374 | 0.59278738 | -0.009814492 |
| [15,] | 1.88559480 | -0.26139179 | -0.286805020 |
| [16,] | 2.54092462 | -1.59249980 | -0.231156680 |
| [17,] | 2.49612443 | -0.59749779 | 0.462868521 |

```
> #输出数字经济变量
```

```
> dig<-pca_data1[,1]*0.8567255+pca_data1[,2]*0.1187059
```

```
> dig
```

| | | | | | |
|-----|-------------|-------------|-------------|-------------|-------------|
| [1] | -2.39485627 | -1.90124446 | -1.73493088 | -1.48625372 | -1.30677096 |
|-----|-------------|-------------|-------------|-------------|-------------|

-0.76209195

[7] -0.22355618 -0.14205598 0.07562096 0.22052235 1.54272409

0.71638187

[13] 0.86784816 0.88885183 1.58440840 1.98783579 2.06756694

>

> #载入被解释变量数据

>

h<-c(15.1016,15.2019,16.2955,15.6376,15.4747,16.7859,16.3929,15.951,16.8614,17.
4234,19.3803,18.8868,18.323,17.975,17.228,16.927,16.3918)

>

pg<-c(3.45226,3.63853,3.8619,3.78197,3.78645,4.08783,3.65085,3.85607,4.50999,4.
63202,5.43414,5.16512,4.96522,4.89663,4.51625,4.37302,4.11194)

>

spg<-c(1.09391,1.2157,1.26138,1.26926,1.29812,1.3776,1.11934,1.29508,1.68359,1.
70678,2.10424,1.94232,1.85313,1.84118,1.63428,1.55888,1.42323)

>

> #载入控制变量数据

>

gdp<-c(21.98531488,21.71547633,15.97717546,14.86674867,25.22844089,21.04663
094,-1.544692474,21.65885392,19.52504522,6.694078947,2.084168337,0.38053819,
-2.593495201,0.466409266,20.61735227,9.629893964,6.310160428)

>

fin<-c(16.92,17.06,16.00,19.27,18.31,18.64,21.22,20.99,21.04,22.78,24.04,24.18,26.8


```

1,26.44,24.19,25.47,27.67)

> une<-c(3,3.1,3,3.2,3.2,3.3,3.9,3.6,3.5,3.3,3.1,3.4,3.5,3.5,3.4,3.3,2.7)

>

gen<-c(105.01,105.44,102.97,102.27,103.87,102.93,104.22,104.575,104.93,104.46,10
5.52,103.3,107.49,106.04,108.34,104.73,102.32)

>

pat<-c(27.53872633,29.29707542,30.73047859,34.1713881,36.36363636,38.1173412
6,35.50278511,38.42563391,36.04041037,32.27094007,31.94761122,38.93032447,3
7.99839443,40.97648645,35.65251003,34.66022283,26.56994165)

>

> #被解释变量描述性统计

> length(h)

[1] 17

> length(pg)

[1] 17

> length(spg)

[1] 17

> mean(h)

[1] 16.83752

> mean(pg)

[1] 4.277658

> mean(spg)

[1] 1.510472

```

```
> sd(h)
[1] 1.25336

> sd(pg)
[1] 0.5921472

> sd(spg)
[1] 0.3064306

> min(h)
[1] 15.1016

> min(pg)
[1] 3.45226

> min(spg)
[1] 1.09391

> max(h)
[1] 19.3803

> max(pg)
[1] 5.43414

> max(spg)
[1] 2.10424

>

> #解释变量描述性统计

> length(dig)
[1] 17

> mean(dig)
```

```
[1] 9.872617e-18  
  
> sd(dig)  
  
[1] 1.417635  
  
> min(dig)  
  
[1] -2.394856  
  
> max(dig)  
  
[1] 2.067567  
  
>  
  
> #控制变量描述性统计  
  
> length(gdp)  
  
[1] 17  
  
> length(fin)  
  
[1] 17  
  
> length(une)  
  
[1] 17  
  
> length(gen)  
  
[1] 17  
  
> length(pat)  
  
[1] 17  
  
> mean(gdp)  
  
[1] 12.00283  
  
> mean(fin)  
  
[1] 21.82529
```

```
> mean(une)
[1] 3.294118

> mean(gen)
[1] 104.6126

> mean(pat)
[1] 34.42317

> sd(gdp)
[1] 9.772657

> sd(fin)
[1] 3.724562

> sd(une)
[1] 0.2771971

> sd(gen)
[1] 1.671771

> sd(pat)
[1] 4.142981

> min(gdp)
[1] -2.593495

> min(fin)
[1] 16

> min(une)
[1] 2.7

> min(gen)
```

```
[1] 102.27  
  
> min(pat)  
  
[1] 26.56994  
  
> max(gdp)  
  
[1] 25.22844  
  
> max(fin)  
  
[1] 27.67  
  
> max(une)  
  
[1] 3.9  
  
> max(gen)  
  
[1] 108.34  
  
> max(pat)  
  
[1] 40.97649  
  
>  
  
> #计量模型回归  
  
> #H 指数为被解释变量的回归  
  
> test1<-lm(h~dig)  
  
> summary(test1)
```

Call:

```
lm(formula = h ~ dig)
```

Residuals:

| Min | 1Q | Median | 3Q | Max |
|---------|---------|---------|--------|--------|
| -1.7146 | -0.5609 | -0.2662 | 0.5227 | 1.6096 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|----------|------------|---------|-------------|
| (Intercept) | 16.8375 | 0.2260 | 74.503 | < 2e-16 *** |
| dig | 0.6137 | 0.1643 | 3.735 | 0.00199 ** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9318 on 15 degrees of freedom

Multiple R-squared: 0.4818, Adjusted R-squared: 0.4473

F-statistic: 13.95 on 1 and 15 DF, p-value: 0.001993

```
> test2<-lm(h~dig+gdp+fin+une+gen+pat)
```

```
> summary(test2)
```

Call:

```
lm(formula = h ~ dig + gdp + fin + une + gen + pat)
```

Residuals:

| Min | 1Q | Median | 3Q | Max |
|----------|----------|----------|---------|---------|
| -0.71149 | -0.41040 | -0.03104 | 0.32389 | 0.71348 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|----------|------------|---------|--------------|
| (Intercept) | 1.47075 | 9.22129 | 0.159 | 0.876455 |
| dig | 1.00549 | 0.28652 | 3.509 | 0.005638 ** |
| gdp | -0.12329 | 0.02392 | -5.154 | 0.000429 *** |
| fin | -0.42702 | 0.13615 | -3.136 | 0.010574 * |
| une | -2.97172 | 0.90385 | -3.288 | 0.008180 ** |
| gen | 0.27211 | 0.09932 | 2.740 | 0.020841 * |
| pat | 0.21758 | 0.05626 | 3.867 | 0.003124 ** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5437 on 10 degrees of freedom

Multiple R-squared: 0.8824, Adjusted R-squared: 0.8119

F-statistic: 12.51 on 6 and 10 DF, p-value: 0.0003843

> #PG 指数为被解释变量的回归

> test3<-lm(pg~dig)

> summary(test3)

Call:

lm(formula = pg ~ dig)

Residuals:

| Min | 1Q | Median | 3Q | Max |
|----------|----------|----------|---------|---------|
| -0.77810 | -0.23069 | -0.05548 | 0.28905 | 0.69955 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|----------|------------|---------|-------------|
| (Intercept) | 4.27766 | 0.10459 | 40.900 | < 2e-16 *** |
| dig | 0.29619 | 0.07605 | 3.895 | 0.00144 ** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4312 on 15 degrees of freedom

Multiple R-squared: 0.5028, Adjusted R-squared: 0.4697

F-statistic: 15.17 on 1 and 15 DF, p-value: 0.001436

```
> test4<-lm(pg~dig+gdp+fin+une+gen+pat)
```

```
> summary(test4)
```

Call:

```
lm(formula = pg ~ dig + gdp + fin + une + gen + pat)
```

Residuals:

| Min | 1Q | Median | 3Q | Max |
|----------|----------|----------|---------|---------|
| -0.32305 | -0.15768 | -0.05786 | 0.06055 | 0.53872 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|----------|------------|---------|------------|
| (Intercept) | -5.50566 | 5.34332 | -1.030 | 0.32711 |
| dig | 0.38669 | 0.16602 | 2.329 | 0.04211 * |
| gdp | -0.04495 | 0.01386 | -3.243 | 0.00883 ** |
| fin | -0.14355 | 0.07889 | -1.820 | 0.09884 . |
| une | -1.60690 | 0.52374 | -3.068 | 0.01188 * |
| gen | 0.14362 | 0.05755 | 2.496 | 0.03169 * |
| pat | 0.10822 | 0.03260 | 3.319 | 0.00776 ** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.315 on 10 degrees of freedom

Multiple R-squared: 0.8231, Adjusted R-squared: 0.717

F-statistic: 7.755 on 6 and 10 DF, p-value: 0.002646

> #SPG 指数为被解释变量的回归

> test5<-lm(spg~dig)

> summary(test5)

Call:

```
lm(formula = spg ~ dig)
```

Residuals:

| | Min | 1Q | Median | 3Q | Max |
|--|----------|----------|----------|---------|---------|
| | -0.39785 | -0.11421 | -0.01604 | 0.16318 | 0.36201 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|----------|------------|---------|--------------|
| (Intercept) | 1.51047 | 0.05519 | 27.369 | 3.22e-14 *** |
| dig | 0.15023 | 0.04013 | 3.744 | 0.00196 ** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2276 on 15 degrees of freedom

Multiple R-squared: 0.483, Adjusted R-squared: 0.4486

F-statistic: 14.01 on 1 and 15 DF, p-value: 0.001956

```
> test6<-lm(spg~dig+gdp+fin+une+gen+pat)
```

```
> summary(test6)
```

Call:

```
lm(formula = spg ~ dig + gdp + fin + une + gen + pat)
```

Residuals:

| Min | 1Q | Median | 3Q | Max |
|----------|----------|----------|---------|---------|
| -0.18365 | -0.09612 | -0.04432 | 0.03879 | 0.32409 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|-----------|------------|---------|----------|
| (Intercept) | -4.044948 | 3.063058 | -1.321 | 0.2161 |
| dig | 0.180675 | 0.095173 | 1.898 | 0.0869 . |
| gdp | -0.020542 | 0.007946 | -2.585 | 0.0272 * |
| fin | -0.063429 | 0.045224 | -1.403 | 0.1910 |
| une | -0.847933 | 0.300235 | -2.824 | 0.0180 * |
| gen | 0.077041 | 0.032990 | 2.335 | 0.0417 * |
| pat | 0.055777 | 0.018689 | 2.984 | 0.0137 * |
| --- | | | | |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1806 on 10 degrees of freedom

Multiple R-squared: 0.7829, Adjusted R-squared: 0.6527

F-statistic: 6.012 on 6 and 10 DF, p-value: 0.006799

致 谢

首先，要感谢我们的指导老师，老师在学术上一直兢兢业业、认真钻研，在此次研究中耐心细致地指导我们的论文，一直鼓励我们，帮助我们攻克论文中的难题，正是因为老师的鼓励与教导，我们才能顺利完成论文。老师在日常生活中非常平易近人，让我们能够在很轻松的氛围下交流和学习，千言万语道不尽我们的感谢，谢谢您！其次要感谢本学院所有老师，老师们的课程使我们收获颇深，他们在专业知识上的学术造诣是我们科研道路上的指明灯。再次，要感谢我们的同班同学及本学院学姐学长在科研过程给予我们宝贵性的意见，他们的意见无疑为我们的论文增光添彩。最后，要感谢论文引用文献的学者，他们的研究为我们的论文奠定了扎实的基础、为本研究提供了借鉴，从他们的论文中可以学习到各类知识和统计软件的操作。

非常有幸在论文撰写的过程中遇到各位老师、学长学姐及同学们，祝愿老师们康乐如意、家庭和睦，祝愿同学及学长学姐们学有所成，未来可期！