

大数据经济学与中国经济社会 复杂系统动态CGE模型构建及应用

娄 峰 著

Dashuju Jingjixue Yu Zhongguo jingji Shehui
Fuza Xitong Dongtai CGE Moxing Goujian Ji Yingyong

中国社会科学出版社

中国社会科学院创新工程学术出版资助项目

经济与管理系列

大数据经济学与中国经济社会 复杂系统动态CGE模型构建及应用

娄 峰 著

Dashuju Jingjixue Yu Zhongguo jingji Shehui
Fuzhaxitong Dongtai CGE Moxing Goujian Ji Yingyong

中国社会科学出版社

图书在版编目 (CIP) 数据

大数据经济学与中国经济社会复杂系统动态 CGE 模型构建
及应用/娄峰著. —北京: 中国社会科学出版社, 2016. 10

ISBN 978 - 7 - 5161 - 9223 - 8

I. ①大… II. ①娄… III. ①中国经济—均衡模型—研究
IV. ①F12

中国版本图书馆 CIP 数据核字(2016)第 258456 号

出版人 赵剑英
责任编辑 卢小生
责任校对 周晓东
责任印制 王超

出版 中国社会科学出版社
社址 北京鼓楼西大街甲 158 号
邮编 100720
网址 <http://www.csspw.cn>
发行部 010 - 84083685
门市部 010 - 84029450
经销 新华书店及其他书店

印刷 北京明恒达印务有限公司
装订 廊坊市广阳区广增装订厂
版次 2016 年 10 月第 1 版
印次 2016 年 10 月第 1 次印刷

开本 710 × 1000 1/16
印张 12.25
插页 2
字数 170 千字
定价 55.00 元

凡购买中国社会科学出版社图书, 如有质量问题请与本社营销中心联系调换
电话: 010 - 84083683
版权所有 侵权必究

目 录

第一章 研究意义和文献综述	1
第一节 研究背景及意义	1
第二节 文献综述	3
第二章 大数据经济学的内涵、特征及影响	8
第一节 大数据的定义及其特征	8
一 大数据的定义	8
二 大数据的主要特点	9
三 正确理解大数据分析	12
第二节 大数据经济学的定义及其特征	14
一 大数据经济学的定义	15
二 大数据经济学的特征	16
第三节 大数据经济学与传统经济学的关系	20
一 大数据经济学与信息经济学的关系	20
二 大数据经济学与信息技术及其他相关 学科的关系	21
三 大数据分析 with 计量经济学的关系	23
第四节 大数据对经济学的冲击	24
一 大数据对传统信息经济理论的冲击	24
二 大数据对传统计量经济学的冲击	26
三 大数据对经济学研究方法的冲击	32

第五节	大数据对经济学的影响	35
一	非因果关系研究	36
二	非假设检验研究	36
三	非样本研究	37
四	研究结果更具有时效性	37
五	研究工具多样化	37
六	统计调查方式的转变	38
第三章	大数据在经济学中应用面临的困难	39
第一节	数据资源不足和分散	39
第二节	专业数据挖掘工具和人才缺失	40
第三节	隐私安全与数据安全的管理薄弱	42
第四节	非结构化数据处理不易	43
第四章	大数据在经济学中的应用及其发展前景	44
第一节	结构化大数据在金融模型中的应用	44
一	运用金融模型处理大数据	44
二	大数据在金融领域中应用的典型案例	47
第二节	大数据分析法在生产物流模型中的应用	49
一	大数据分析法	50
二	基本物流模型	50
三	物流模型和大数据分析组合的可能性	51
四	大数据在物流模型中应用的典型案例	51
第三节	大数据在经济预测中应用的发展前景	52
一	搜索数据的应用可提高预测的精度	54
二	媒体数据的应用可增强预测的时效性	55
三	其他可挖掘的在线数据十分丰富	56

第五章 大数据在中国经济社会复杂系统动态 CGE 模型中应用的基本思路	59
第六章 中国经济社会复杂系统动态 CGE 模型构建	62
第一节 经济社会复杂系统动态模型构建的必要性	62
第二节 中国经济—能源—环境—税收—人口动态 CGE 模型的部门划分	64
第三节 中国经济—能源—环境—税收—人口动态 CGE 模型方程	66
一 生产模块	67
二 国际贸易模块	78
三 居民收入和消费模块	81
四 企业模块	85
五 政府模块	85
六 均衡模块	88
七 社会福利模块	90
八 环境污染模块	91
九 动态模块	93
第七章 中国经济社会复杂系统模型社会核算矩阵构建及数据处理	95
第一节 能源和税收数据来源及处理	95
一 石油和天然气开采部门数据	95
二 电力能源部门数据	96
三 部门关税数据的确定	96
四 中国经济—能源—环境—税收—人口宏观 SAM 表	98

第二节	中国经济—能源—环境—税收—人口动态	
	CGE 模型参数的标定	99
一	替代弹性设定	99
二	模型参数标定	103
三	居民消费函数参数	107
四	部门资本存量	110
五	部门劳动力数量	112
六	二氧化碳排放系数	113
第三节	动态参数校准和设定	114
一	历史模拟	114
二	预测模拟 (基准情景)	115
第四节	模型的计算机实现步骤	116
第五节	中国经济—能源—环境—税收—人口 (CGE)	
	模型与程序检验	117
一	静态 CGE 模型检验	117
二	动态 CGE 模型检验	118

第八章 人民币汇率改革政策模拟分析

	——基于中国经济社会复杂系统 CGE 模型	120
第一节	研究背景及意义	120
第二节	相关文献综述	122
一	汇率变动对贸易结构的影响	122
二	汇率变动对产业结构的影响	137
三	汇率变动对收入分配的影响	139
四	汇率变动对经济增长的影响	142
五	应用 CGE 模型研究汇率问题文献	146
第三节	人民币汇率政策静态模拟结果分析	148
第四节	人民币汇率动态模拟结果分析	154
一	对宏观经济的影响	155

二	对各行业出口的影响	156
三	对城乡不同阶层居民收入的影响	164
四	对不同行业不同时期关税的影响	166
第五节	研究结论及政策建议	173
一	顺应市场, 实施人民币汇率适当贬值战略	173
二	加强管理, 完善人民币汇率市场形成机制	174
三	大力推进人民币国际化, 逐步实现人民币 资本自由兑换	175
四	大力发展外汇衍生品市场, 适时推出和发展 人民币外汇期货	175
五	加强我国金融监管, 合理引导外资流动, 有效防范金融风险	176
六	加快转变对外贸易增长方式, 提高自主 研发能力	177
	参考文献	178
	后记	186

第一章 研究意义和文献综述

第一节 研究背景及意义

随着云计算、物联网、社交网络等信息技术的快速发展，每时每刻数以亿计的计算机和移动通信设备正在持续不断地创造出海量的数据信息，并逐渐融入我们的学习、工作和生活中，这意味着一个前所未有的大数据时代已经来临。

数据信息的创造正在以前所未有的指数级速度扩张。全球 IP 流量达到 1EB 所需的时间，2001 年尚需要 1 年，2013 年仅需 1 天，到 2020 年则仅需 1 小时，全球互联网上的数据每年增长率约为 50%，两年就会翻一番，目前，世界上数据总量的 80% 都是最近两年产生的。2012 年，全球知名研究机构 IDC（互联网络数据中心）发布的《数字宇宙 2020》指出，2005—2020 年，全球数据总量规模将增加 300 倍之多，从 130EB 增加到 40ZB，也就是 40 万亿 GB，这个数量约为地球上所有海滩上沙粒数量的 57 倍之多。美国著名高科技思科公司预测 2016 年全世界产生的数据总量将达到 1.3ZB；谷歌前 CEO 施密特指出，从人类文明开始到 2003 年的近万年时间里人类社会总共产生了大约 5EB 数据，而 2010 年人类社会每两天就能产生 5EB 数据。

2011 年，美国麦肯锡全球研究院发布《大数据：下一个创新、竞争和生产力的前沿》大数据分析报告，指出大数据是新世纪的基

础生活资料与市场要素，其重要程度不亚于物质资产和人力资本，大数据将成为企业提高生产力和提高核心竞争力的关键要素和主要途径之一。资产要素、产业垂直整合、泛互联网化是大数据时代的三大发展趋势。随着大数据分析技术的发展和成熟，将为人类深入探索宇宙奥秘提供有力方法和捷径。

大数据开启了一次重大的时代转型，其带来的信息风暴正在改变人类的生活、工作和思维，改变人类对社会的认知、交流、管理和决策的行为方式，从而引发在经济社会、学术科研、国防军事等领域的深刻变革。值得说明的是，大数据不仅仅是指数据规模庞大，更注重的是数据处理技术和应用，是数据对象、技术与应用三者的和谐统一。从数据来源来看，大数据既包括政府部门或企业掌握的各种传统的数字数据，也包括诸如微博、微信、评论等社交网络上虚拟的文本数据；从数据处理技术来看，大数据技术包括数据采集、储存、清洗、管理、分析、挖掘、可视化等技术及其集成，并将理论、实验和计算仿真三种传统研究方法统一起来；从应用来看，大数据应用是根据大数据技术对各种海量数据进行挖掘分析，从中找到海量数据信息背后隐含的规律和有价值信息的过程，越来越多的商业模式、产业格局、生态价值、科技创新、教育理念、政府企业，甚至个人消费者行为等诸多问题都可以通过大数据技术来解决或优化。因此，大数据作为信息技术与各行业领域紧密融合的新事物，需要坚持数据对象、技术和应用三位一体共同发展之路，才能最终形成变革之力，才能拥有旺盛的需求空间和广阔的发展前景。

在经济学领域，大数据应用还处于初步探索阶段，但随着海量经济数据的几何式增长，以及网络技术与计算方法的不断完善和更新，将大数据运用于经济学逐渐成为学术界、商界以及社会广泛关心的热点问题之一。目前，大数据作为一种有用的信息资源，在商业、金融等领域发挥越来越重要的作用，也逐渐成为社会科学的国际前沿应用研究内容之一，然而，在经济学领域，大数据还很少被

用到 [据统计,截至2014年12月,谷歌(Google)学术搜索到的与“大数据”有关的研究论文共3026篇,其中仅有29篇是和经济学相关的]。但因海量经济数据资源的快速增长,计算技术和能力的不断提高,以及方法论的不断发展,将大数据分析技术运用于经济学已成为一个值得探讨的新课题。展望未来,由于经济学是一门理论与实践相结合的学科,将大数据应用于经济学,有可能会开辟一个全新的经济学发展领域。

大数据时代的到来,对我国的经济社会发展既是一个机遇也是一种挑战。从方法论上看,大数据时代所带来的对科研思维方式的转变将是一次重大变革,这将对促进学科融合、推动科技创新、集成学术资源、突破科技“瓶颈”,甚至引发一场新的科技革命等方面发挥重大作用。从经济学的角度,研究大数据经济学及其应用,探索如何把大数据与传统经济学研究方法、国际主流的大型系统模型结合起来,尝试构建中国宏观—中观—微观一体化的经济社会复杂系统模型,用于模拟、分析和评价中国各项政策,从而为我国的宏观政策制定、产业发展规划等提供科学、合理和可行的政策建议,这既是一种挑战和尝试,也是我们的追求和期望。

第二节 文献综述

最近几年,国际学术界对大数据的研究才开始进入一个蓬勃发展阶段。2008年,《自然》(*Nature*)杂志推出“大数据”专题分析报告,系统地分析和探讨大数据在商业贸易、市场营销、互联网金融、生物医药、气候环境等学科和领域的应用,大数据急需技术及其面对的主要困难和挑战。

2011年,《科学》(*Science*)杂志以“应对大数据”专题的形式深入探讨了大数据储存、管理、清洗和计算等具体技术问题,详细分析全球信息化时代下大数据带来的机遇和挑战,并指出由于大

数据技术可以促进科学技术更好地转化成生产力，从而大幅提高科学技术对经济社会的推动作用，因此，各国政府和科研机构应充分重视大数据技术并应该进一步加强对大数据的研发投入和开发。

2012年，被誉为“大数据商业应用第一人”的维克托·迈尔·舍恩伯格教授出版了《大数据时代：生活、工作与思维的大变革》专著，该书前瞻性地指出，大数据带来的信息风暴正在改变着人们的工作、生活和思维方式，大数据开启了人类历史上的一次重大时代变革，该书明确提出大数据时代处理数据理念上的三大转变，即从抽样样本转向全体样本，从绝对精确转向总体效率，从注重因果性转向相关性；接着，舍恩伯格还从数据化和数据交叉使用的巨大价值两个方面战略性地论述了大数据研究的价值和根本动力；另外，作者也冷静地分析了目前大数据发展面临的主要问题和可能导致的挑战，包括产业生态环境、数据国家安全、个人信息隐私、社会发展公平等问题。

2009年，联合国提出了“数据脉动”计划，开展了对大数据技术及其应用的初步研究，并于2012年5月发布了“大数据发展：挑战和机遇”系列报告，该报告应用实时信息数据分析技术，检验和实证了大数据实时分析和共享技术对促进政府决策由“预报”转向“实报”等方面的改进作用和效果，并在此基础上提出了发展大数据分析技术的相关建议和政策指导。

随着国外大数据技术的蓬勃发展，近年来，国内学者对大数据研究也开始迅速发展起来。2012年，我国著名信息管理专家涂子沛率先出版了《大数据：正在到来的数据革命》一书，该书围绕美国政府基于数据的管理历史线条展开，让大家更加容易理解大数据在政府执政和公共事务管理中发挥的作用，它以数据为轴线，描绘了美国走过的开放创新历程以及美国如何利用大数据理念和技术更好地实现政府治理和社会管理，全书以别开生面的经典案例——奥巴马建设“前所未有的开放政府”的雄心、全民医改法案的波澜、统一身份证的百年纠结、街头警察的创新传奇、美国矿难的悲情历

史，以及商务智能的前世今生、数据开放运动的全球兴起等，细解数据创新给公民、政府、社会带来的种种挑战和变革。

郑毅（2012）的《证析：大数据与基于证据的决策》一书中在大数据背景下，结合实例论述了大数据对证据尤其是量化证据对商业与社会领域决策的影响。该书由两大部分组成，第一部分对证析所处的时代背景、证析对传统决策方式的挑战、证析在企业中的应用案例进行详细介绍，并着重分析了实验在指导社会实践中的思想和实例；第二部分主要从分析企业价值，为了发挥证析的价值而在企业组织架构、考核体系、决策流程、组织文化等方面应有的考虑。除了介绍最新的管理思想与企业实践之外，该书还探讨了大数据对科学研究方法的影响和转变。总的来说，该书的重点并不是讲述大数据分析具体技术，而是论述如何利用大数据分析思想来提升决策的质量，如何在统计学、哲学、经济学、管理学、社会学、心理学等学科中，把传统分析技术和现代的大数据技术相结合等问题。

朱志军等（2012）在《大数据（大价值、大机遇、大变革）》一书中介绍了大数据产生的背景、特征和发展趋势，从实证的角度探讨了它对社会和商业智能的影响，并认为大数据正影响着商业模式的转变，将带来新的商业机会，能否对大数据进行处理、分析与整合将成为提升企业核心竞争力的关键，大数据，既是一场大机遇，也将引发一场大变革。

目前，国外关于大数据在经济学领域的分析和应用也处于探索起步阶段。相关文献较少，具有代表性的文献有：瓦桑特（2013）从预测模型视角探讨分析了大数据技术。他认为，利用大数据分析技术在评估新知识是否有助于制定合理政府决策时，更要关注预测模型对未来的预测能力，仅仅对历史数据拟合效果好的预测模型不一定能够精确地预测到未来结果，其预测能力不一定优于拟合效果一般的模型；大数据分析技术能够使机器自主学习能力得到加强，从而使机器有可能发现人们并没有注意到的很有意义的问题或线

索。因此，基于网络数据驱动的机器自主学习和预测模型的两者结合在未来商业中将越来越多地受到重视。

沃夫根冈（2013）从科学学方法角度探讨了大数据挖掘技术的主要特征。他认为，大数据方法显著不同于传统的计算机模拟方法，大数据方法的数据来源更加多元化和多样化，包含数字数据和非结构数据；大数据更加注重相关关系，但是，这并不意味着大数据方法否定了因果关系，两者应互为验证和补充，仅仅根据大数据的相关性计算特征来反驳事物间的因果关系是不正确的；仅仅根据数据信息，利用大数据分析技术得到的预测结构缺乏经济学上的合理解释，往往难以令人信服，因此，大数据分析结果相对缺乏解释力；大数据模型的特征是“层状型的”，而传统经济计量模型所得出的“水平型的”结构，从水平型模型到层状型模型的转变是需要数据多元性、参数时变性以及关系空间性。

里卡德（2013）基于大数据降维技术，提出了一种把大数据应用于经济预测的新方法，即文本直接算法分析（DATA）法，并与信号测试分析法、信息标准分析法和因子分析法三种时间序列分析方法进行预测精度对比，研究结果显示，文本直接算法分析（DATA）法具有良好的预测一致性和稳健性。

泰勒（2014）探讨如何通过大数据分析技术把社会调查数据和劳动市场微观数据应用于经济学分析的新途径，并分析了大数据在经济学研究中面临的主要问题和未来潜在贡献，还有如何根据大数据改进经济计量模型以及大数据如何在经济学和其他学科之间进行联结和相互验证等问题。

瓦里安（2014）论述了五种大数据分析和使用方法。认为这些方法可以应用于经济学研究，并建议计算科学系为研究生增设一门机器学习仿真课程，并提出计算机学者与统计学学者或者计量经济学学者相互交流和配合可能是未来大数据在经济学研究领域中很有前景的发展方式。

国内关于大数据的研究文献主要集中在探讨大数据特征及其应

用案例介绍（主要集中在商业、信息领域），关于大数据在经济学领域的介绍和分析还寥寥无几。^①从上面的文献综述中可以看到，大数据分析技术具有巨大的发展潜力和远大的发展前景，大数据技术已经引起世界各国政府的关注，激发了国际学术界的研究热情，然而，从研究深度上讲，国内外关于大数据的研究总体还处于基础起步阶段，诸多文献基本都是定性地或者案例式地阐述大数据的发展、特征及作用等；从专业领域来看，目前大数据主要应用于商业和管理领域，对于大数据在经济学中的应用寥寥无几，尤其是大数据如何与经济社会复杂系统模型的结合几乎是空白。

本书在系统学习和实地考察的基础上（2014年10月7日至11月8日期间，笔者参加了中国社会科学院组织的“大数据与社会科学创新”美国学术培训班，与中国社会科学院相关研究所以及院属单位共23人一起到美国学习，在美国加利福尼亚州立大学听取了美国众多大数据分析及应用方面的知名学者、大学教授、软件开发科技人员的研究报告和课程培训，并且在此期间，我们还利用周六、周日的业余时间实地参观考察了思科公司、英特尔、谷歌、IBM、惠普、甲骨文以及波音等美国国际一流高新技术企业，倾听这些企业在大数据方面的应用和实践经验。通过这三十天的交流和学习，笔者对大数据在美国商业领域、社会领域和IT领域，尤其是经济学等领域最新进展、用途及未来发展方向有了深入的了解和认识，感触颇深），本书结合笔者自己多年的大型宏观经济系统模型建模经验，试图从一个新的角度或方向探索大数据与经济社会系统模型结合及其应用。

^① 根据知网，仅找到三篇相关文献。

第二章 大数据经济学的内涵、特征及影响

第一节 大数据的定义及其特征

一 大数据的定义

二三十年前，关于经济活动的相关数据还是极其稀缺的。但在短时间内，这一现象却发生了戏剧性的改变。其中一个很重要的原因是网络的飞速发展。现实中的一切都能被互联网所记录。当你在 Google 或 Bing 上搜索时，你的查询和随后的点击结果就会被记录下来。当你在 Amazon 或 eBay 上购物时，每一笔交易和点击也被记录下来。当你在网上读报、看视频或查询个人金融账户余额时，同样也会被记录下来。个体行为的数据记录并不仅仅局限于互联网。短信、电话、定位、扫描数据、就业记录与电子健康记录都会留下记录痕迹。

我们通过一个具体的例子对此加以说明。不妨以零售商收集的数据为例。几十年前，商店会收集每日的日常销售数据，若这些数据以商品名或商品种类加以分类，就会被认为是高质量数据。如今，数据扫描的出现使数据收集工作更加细化。如追踪个体的购买与销售记录及其发生的具体时间，运用电子交易数据将物品在货架的摆放位置和物品的库存相联系。网络零售商不仅观察这些信息数据，他们还关心消费者因价格变化所做出的反应。如消费者最初的搜索查询（关键词、促销活动、产品介绍和卖家评论等查询）。实

际上，这些销售数据还与人口统计资料、广告支出、社交媒体活动、线下支出、信用记录等息息相关。

商业活动与数据发展同步。当企业将其线下经营转移到线上经营时，编译物价、实际招聘与商品的出货量等数据信息便不再困难。因此，就产生了许多工作协同能力、人员评估和生产率度量的电子数据。公众可运用这些线上数据评价分析税收文件、社会保险、政府支出与常规活动的效率。

“大数据”一词最初是由英文单词“Big Data”直接翻译得到的。目前，国际学术界还没有一个“大数据”公认的定义，但基本内涵大致相同，比较具有代表性的定义包括：

著名的IT咨询公司加德纳（Gartner）把大数据定义为：大数据是指需要应用新处理模式才能获得的具有更强决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产。

维基百科关于大数据的定义是：大数据指的是所涉及的资料量规模巨大到超过目前主流软件工具分析能力，并能在合理时间内达到攫取、管理、处理和整理，最终能够帮助企业经营决策目的的资讯。

全球著名咨询公司麦肯锡（McKinsey）关于大数据的定义是：大数据是指在一定时间内无法用传统的数据库软件工具对其内容进行采集、存储、管理和分析的数据集合。

二 大数据的主要特点

从以上的诸多定义中，国际学术界和IT业界较为普遍认可大数据具有“4V”特征，即Volume（规模）、Variety（种类）、Velocity（速度）和Value（价值），这些特征较好地阐释了大数据的主要特点，具体解释分析如下：

Volume是指数据体量的海量性。目前，我们正处于数据信息爆炸时代，由于互联网的高速发展、电子商务平台的普遍覆盖和微博客等新兴社交网站的迅速崛起，手机、平板电脑、PC终端的通话、浏览网页等产生了无穷无尽的各种数据内容。同时，由于传