

中国直辖市社会核算矩阵数据集

张 扬¹, 马天乐^{2*}

1. 湖南财政经济学院, 长沙 410022; 2. 长沙环境保护职业技术学院, 长沙 410004

摘 要: 可计算一般均衡模型 (Computable General Equilibrium, CGE) 是经济学的重要研究方法之一, 近年来被广泛应用在关税、贸易壁垒、资源与环境等领域。本文基于投入产出理论, 将我国 41 个产业部门与土地利用现状分类一级分类中 5 种国有建设用地相对应。根据最新的投入产出表, 结合《中国国土资源年鉴》、《北京市统计年鉴》、《上海市统计年鉴》、《天津市统计年鉴》、《重庆市统计年鉴》等相关数据编制了 2012–2015 年北京、上海、天津、重庆四个直辖市的社会核算矩阵 (SAM), 并且根据直接交叉熵法进行了平衡。本数据集对构建城市用地 CGE 模型, 研究城市用地及城市经济学相关问题能提供数据支持。

关键词: CGE 模型; SAM 表; 城市经济学; 经济地理

DOI: 10.3974/geodp.2020.01.07

1 前言

现代经济学分析产业之间的投入产出联系, 精确地刻画了社会经济中各个部门各个产业是如何关联的。比如汽车产量的增加, 不但引起钢铁产量的增加, 还会引起塑料、橡胶、电子、化工、机械等各个部门产量的增加, 这些部门的发展又会进一步刺激更多行业的发展。因此, 经济学家列昂惕夫 (Wassily Leontief) 提出了投入产出模型, 也即列昂惕夫模型, 运用联立的线性方程组对国民经济各个部门经济与技术的相互联系与相互影响进行精确的描述^[1]。投入产出方法的提出, 对现代经济学的发展产生了巨大的促进作用与深远的影响。CGE 模型全称是可计算一般均衡模型 (Computable General Equilibrium Model), 是投入产出分析的深化与发展。传统的投入产出模型只是从生产的中间产品/中间投入与要素投入、最终产品等环节进行分析, 并没有考虑居民、企业、政府等部门的经济活动, 也没有考虑储蓄、投资、进出口等社会经济账目, 因此其作用具有一定的局限性^[1]。CGE 模型将居民、企业、政府、储蓄、投资、进出口等部门与账目统一设置在模型中进行综合考虑, 根据瓦尔拉斯的一般均衡思想^[2]对其价格变量进行求解, 分析结果更加全面、系统, 因此在国际贸易、政府税收、资源与环境、公共政策等领域发挥了巨大的作用, 得到了广泛的应用, 逐渐成为现代经济学研究的一个重要工具^[1]。

收稿日期: 2019-10-22; 修订日期: 2019-11-24; 出版日期: 2020-03-25

基金项目: 湖南省教育厅 (18C0964)

*通讯作者: 马天乐, 长沙环境保护职业技术学院, 16582397@qq.com

数据引用方式: [1] 张扬, 马天乐. 中国直辖市社会核算矩阵数据集[J]. 全球变化数据学报, 2020, 4(1): 46–51.

DOI: 10.3974/geodp.2020.01.07.

[2] 张扬. 中国四个直辖市社会核算矩阵数据集(2012, 2015) [DB/OL]. 全球变化数据仓储, 2019.

DOI: 10.3974/geodb.2019.05.12.V1.

CGE 模型数据来源于社会核算矩阵 (Social Accounting Matrix, SAM), 而社会核算矩阵以投入产出表为基础, 在投入产出表的基础上增加了政府、储蓄投资、进出口等部门与模块, 反映社会经济各个部门之间资金流动与比例关系, 对社会经济活动进行了更加细致的刻画。同时, 我国一般只有直辖市与省级行政区才编制投入产出表, 而且每逢尾数为 2、7 的年份编表, 编制时间间隔为 5 年, 为了保持数据的更新, 每逢尾数为 0、5 的年份编制投入产出延长表, 投入产出延长表是在投入产出表的基础上进行数据更新, 部门分类与投入产出表完全一致, 数据也具有延长性, 只是数据来源于统计资料, 而不是像投入产出表中的数据那样来自直接调查, 因此投入产出延长表与投入产出表基本一致, 可以一起使用。由于目前国家统计局尚未编制 2017 年直辖市投入产出表, 最新的投入产出表是 2015 年重庆市投入产出延长表。本文以北京市、天津市、上海市与重庆市为例, 根据相关数据编制 2015 年重庆市 SAM 表、2012 年北京、天津、上海、重庆 4 个直辖市的 SAM 表。

2 数据集元数据简介

《中国四个直辖市社会核算矩阵数据集 (2012, 2015) 》^[3]的名称、作者、地理区域、数据年代、数据格式、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表 1。

表 1 《中国四个直辖市社会核算矩阵数据集 (2012, 2015) 》元数据简表

条目	描述
数据集名称	中国四个直辖市社会核算矩阵数据集 (2012, 2015)
数据集短名	SAM_4MunicipalitiesChina
作者信息	张扬 0000-0002-1633-2478, 湖南财政经济学院, 520andylau@sina.com
地理区域	4 个直辖市: 北京、上海、天津、重庆
数据年代	2012–2015
数据格式	.xlsx
数据量	25.2 KB
数据集组成	由 1 个 Excel 文件组成, 包括 6 个工作表, 分别是: (1) 2012 年北京市社会核算矩阵表; (2) 2012 年上海市社会核算矩阵表; (3) 2012 年天津市社会核算矩阵表; (4) 2012 年重庆市社会核算矩阵表; (5) 2015 年重庆市社会核算矩阵表; (6) 社会核算矩阵中的代码、含义与单位
基金项目	湖南省教育厅 (18C0964)
出版与共享服务平台	全球变化科学研究数据出版系统 http://www.geodoi.ac.cn
地址	北京市朝阳区大屯路甲 11 号 100101, 中国科学院地理科学与资源研究所
数据共享政策	全球变化科学研究数据出版系统的“数据”包括元数据 (中英文)、实体数据 (中英文) 和通过《全球变化数据学报》(中英文) 发表的数据论文。其共享政策如下: (1) “数据”以最便利的方式通过互联网系统免费向全社会开放, 用户免费浏览、免费下载; (2) 最终用户使用“数据”需要按照引用格式在参考文献或适当的位置标注数据来源; (3) 增值服务用户或以任何形式散发和传播 (包括通过计算机服务器) “数据”的用户需要与《全球变化数据学报》(中英文) 编辑部签署书面协议, 获得许可; (4) 摘取“数据”中的部分记录创作新数据的作者需要遵循 10% 引用原则, 即从本数据集中摘取的数据记录少于新数据集总记录量的 10%, 同时需要对摘取的数据记录标注数据来源 ^[4]
数据和论文检索系统	DOI, DCI, CSD, WDS/ISC, GEOSS, China GEOSS, Crossref

3 数据研发方法

3.1 要素定义

SAM 表中的 10 种要素包括工矿仓储、交通运输、住宅、商服、公共管理、劳动力、资本、土地、居民、政府。行业用地类型与投入产出部门的对应关系见表 2。

表 2 我国用地类型与投入产出部门对比表

土地利用类型	土地利用类型的含义	对应的投入产出表部门
工矿仓储用地	主要用于工业生产、物资存放场所的土地，包括工业用地、采矿用地与仓储用地	煤炭采选产品 石油和天然气开采产品 金属矿采选产品 非金属矿和其他矿采选产品 食品和烟草 纺织品 纺织服装鞋帽皮革羽绒及其制品 木材加工品和家具 造纸印刷和文教体育用品 石油、炼焦产品和核燃料加工品 化学产品 非金属矿物制品 金属冶炼和压延加工品 金属制品 通用设备 专用设备 交通运输设备 电气机械和器材 通信设备、计算机和其他电子设备 仪器仪表 其他制造产品 废品废料 金属制品、机械和设备修理服务 电力、热力的生产和供应 燃气生产和供应 水的生产和供应 建筑
交通运输用地	用于运输通行的地面线路、场站等的土地。包括民用机场、汽车客货运场站、港口、码头、地面运输管道和各种道路以及轨道交通用地	交通运输、仓储和邮政
住宅用地	主要用于人们生活居住的房基地及其附属设施的土地	房地产
商服用地	主要用于商业、服务业的土地，包括批发零售用地、住宿餐饮用地、商务金融用地和其他商服用地	批发和零售 住宿和餐饮 信息传输、软件和信息技术服务 金融 租赁和商务服务

续表

土地利用类型	土地利用类型的含义	对应的投入产出表部门
公共管理与公共 服务用地	用于机关团体、新闻出版、科教文卫、科学研究和技术服务 风景名胜、公共设施等的土地	水利、环境和公共设施管理
		居民服务、修理和其他服务
		教育
		卫生和社会工作
		文化、体育和娱乐
		公共管理、社会保障和社会组织

与投入产出表类似，SAM 表是一个行与列相等的矩阵，相同的行与列对应国民经济中的某个部门。一般来说，SAM 表包括生产/活动、劳动力、土地与资本等生产要素、以及居民、政府等账户与部门^[1]。其中生产/活动部门反应了产业部门之间的中间需求与中间投入关系，可以分开成生产与活动两个部门，也可以根据需要合成一个部门，为了分析需要进行简化处理，将生产与活动两个部门合并为生产/活动部门，具体见表 3 所示。

表 3 SAM 表的组成部分

	生产/活动	生产要素	居民	政府	汇总
生产/活动	中间投入/中间使用		居民消费	政府购买	
生产要素	要素增值部分				要素收入
居民		居民要素收入			居民总收入
政府			个人所得税		政府总收入
汇总	总投入	要素支出	居民支出	政府支出	

土地是一切生产与消费活动的物质载体，承载着所有产业部门的生产与服务活动，通过分析可以发现，我国投入产出分析的 42 个产业部门基本上与我国土地利用现状分类一级类^[5]中的交通运输用地、工矿仓储用地等 5 种国有建设用地类型相对应（表 2）。具体来看，根据土地分类中的建设用地划分与行业定义，可以将投入产出表中的信息传输、软件和信息技术服务等 5 个行业与商服用地进行对应；将煤炭开采和洗选业、金属冶炼和压延加工品等 27 个行业与工矿仓储用地进行对应；将科学研究和技术服务、水利、环境和公共设施管理等 7 个行业与公共管理与公共服务用地进行对应；将房地产业与住宅用地进行对应。实际上，投入产出表 42 部门存在部分重叠领域，因此无法和土地利用现状分类完全一一对应，比如仓储业的划分问题。由于仓储业与交通运输业在我国投入产出表中被合并为交通运输、仓储和邮政业（代码 30），无法进行拆分，并且在这里仓储业是指专门从事货物运输中转仓储，以及以仓储为主的货物配送活动，其主要的业务与交通运输属于同一门类，因此将交通运输及仓储业与交通运输用地进行对应。国有建设用地中的特殊用地包括军事设施、使领馆等用地，建设用地中的水域及水利设施用地包括水库水面、沿海滩涂等用地，其他用地包括空闲地。以北京市为例，2012 年北京市特殊用地、水域及水利设施用地和其他用地分别为 3.33、6.83 和 0 hm²^[6]，所占比例很小，并且与城市社会经济联系并不紧密，因此没有被纳入研究范围。由于本数据集主要用来研究城市建设用地结构，因此剔除投入

产出表中农林牧渔等部门数据与相应的农用地数据。SAM 表的其他部门借鉴文献[1], 与标准 CGE 模型相匹配, 包括要素、居民、政府与汇总等部门。

3.2 统计方法

根据以上分析, 首先将 2012 年北京市、上海市、天津市与重庆市投入产出表^[7-10]与 2015 年重庆市投入产出延长表^[10]中的原始数据根据表 2 部门分类进行横向与纵向的合并, 然后将 SAM 表的其他部分数据根据各省市统计年鉴补齐。其中 2012、2015 年重庆市与 2012 年北京市、上海市、天津市交通运输用地、工矿仓储用地、公共管理与公共服务用地、住宅用地以及商服用地数据来源于文献[6, 11]: 国有建设用地供应情况——按用地类型和省市分列; SAM 表要素增值部分来源于 2012、2015 年重庆市与 2012 年北京市、上海市、天津市投入产出表要素增值部分^[7-10, 12]; SAM 表的居民要素收入来源于 2013、2016 年重庆市与 2013 年北京市、上海市、天津市统计年鉴人民生活与物价部分^[13-17]; SAM 表的居民消费与政府购买来源于 2012、2015 年重庆市与 2012 年北京市、上海市、天津市投入产出表^[7-10, 12]; SAM 表个人所得税来源于 2013、2016 年重庆市与 2013 年北京市、上海市、天津市统计年鉴财政部分^[13-17]。

与投入产出表相似, SAM 表每列的总数与每行的总数相等, 即实现 SAM 表的“平衡”。但是, 通常意义上初级的 SAM 表并不平衡, 因此需要通过不同的方法使其平衡。交叉熵法 (Cross Entropy) 是现代 CGE 模型采用的较多的平衡方法, 包括直接交叉熵法与系数交叉熵法, 其中系数交叉熵法主要是针对现有 SAM 表数据的更新。因此, 为了保留经济学思想, 又克服其主观性, 本数据集采用以直接交叉熵法为主、手动平衡法为辅的方法对 SAM 表进行平衡。交叉熵法根据信息经济学提出的信息熵来构建模型, 其目标函数为^[1]:

$$\min_{X_{ij}} z = \frac{1}{H} \sum_j \sum_i Q_{ij} \log \frac{Q_{ij}}{\bar{Q}_{ij}} - \log \frac{H}{\bar{H}} \quad (1)$$

$$H = \sum_j \sum_i Q_{ij}, \quad \bar{H} = \sum_j \sum_i \bar{Q}_{ij}, \quad s.t. \sum_i Q_{ik} = \sum_j Q_{kj} \quad (2)$$

式中, \bar{Q}_{ij} 是 SAM 表的原始数据, Q_{ij} 是平衡后的数据, 均为正数。

4 数据集组成

本数据集存储为 1 个.xlsx 格式文件, 由 6 个数据表组成, 数据量为 25.2 KB, 具体参见表 1。

5 讨论和总结

SAM 表是 CGE 模型的基础数据, 也是投入产出表的进一步发展, 在投入产出表的基础上增加了政府、储蓄投资、进出口等部门与模块, 反映社会经济各个部门之间资金流动与比例关系, 对社会经济活动进行了更加细致的刻画。以北京市 SAM 表为例, 通过可计算一般均衡分析, 分析 GDP 增加、劳动力、资金与土地等各种要素价格与总量变化等情况下, 预测北京市经济社会各部门受到的影响, 以此进行更精准的政策评估与调整。例如, 当 GDP 增速为 6.5%、土地价格上涨 20%、劳动力总量增加 20%时, 将 SAM 表数据输入

GAMS（General Algebraic Modeling System）软件，建立相应 CGE 模型后的结果见表 4。

表 4 3 种情景方案下北京市最优建设用地结构变动

	工矿仓储		交通运输		住宅		商服		公共管理与公共服务	
	面积 (km ²)	变动 比例 (%)	面积 (km ²)	变动 比例 (%)	面积 (km ²)	变动 比例 (%)	面积 (km ²)	变动 比例 (%)	面积 (km ²)	变动 比例 (%)
实际值	5.31	—	0.33	—	6.14	—	1.49	—	3.00	—
情景 1	0.03	−99.48	24.42	7,398.97	86.26	1,305.85	0.01	−99.08	28.82	861.16
情景 2	0.48	−90.88	0.66	103.22	5.78	−5.88	0.15	−90.25	0.21	−92.85
情景 3	0.38	−92.87	0.62	89.89	5.02	−18.17	0.12	−92.13	0.18	−94.06

由表 4 可知：总体来看，伴随着经济发展，在 GDP 平稳增加、土地要素价格不断上涨、二胎政策带来劳动力供给显著增加等 3 种情景模拟下，北京市新增建设用地结构发生了显著的变化。当预期 GDP 平稳增长 6.5% 时，与实际值相比，要有效加大交通运输用地、住宅用地、公共管理与公共服务用地的供给；当土地价格增加 20%，由于用地成本的明显增加，除交通运输用地以外，要严格控制工矿仓储用地、住宅用地、商服用地、公共管理与公共服务用地的审批与供给；预期劳动力总量增加 20% 时，模拟结果与情景 2 比较类似。因此，编制北京市、上海市、天津市与重庆市 SAM 表有利于运用 CGE 模型对各直辖市的社会经济问题进行更加全面而深入地分析。

参考文献

[1] 张欣. 可计算一般均衡模型的基本原理与编程[M]. 上海: 上海人民出版社, 2010: 1–272.

[2] 高鸿业. 西方经济学[M]. 北京: 中国人民大学出版社, 2014: 1–685.

[3] 张扬. 中国四个直辖市社会核算矩阵数据集(2012, 2015) [DB/OL]. 全球变化数据仓储, 2019. DOI: 10.3974/geodb.2019.05.12.V1.

[4] 全球变化科学研究数据出版系统. 全球变化科学研究数据共享政策 [OL]. DOI: 10.3974/dp.policy.2014.05 (2017 年更新).

[5] GB/T 21010—2007. 土地利用现状分类[S]. 北京: 中国标准出版社, 2007.

[6] 中华人民共和国国土资源部. 中国国土资源年鉴 2013[M]. 北京: 中国国土资源年鉴编辑部, 2013: 830–831.

[7] 北京市统计局. 2012 年 42 部门投入产出表[Z]. 北京: 北京市统计局.

[8] 上海市统计局. 2012 年 42 部门投入产出表[Z]. 上海: 上海市统计局.

[9] 天津市统计局. 2012 年 42 部门投入产出表[Z]. 天津: 天津市统计局.

[10] 重庆市统计局. 2012 年 42 部门投入产出表[Z]. 重庆: 重庆市统计局.

[11] 中华人民共和国国土资源部. 中国国土资源年鉴 2016[M]. 北京: 中国国土资源年鉴编辑部, 2016: 110–111.

[12] 重庆市统计局、国家统计局重庆调查总队. 重庆市统计年鉴 2017[M]. 北京: 中国统计出版社, 2017.

[13] 北京市统计局, 国家统计局北京调查总队. 北京统计年鉴 2013[M]. 北京: 中国统计出版社, 2013.

[14] 天津市统计局, 国家统计局天津调查总队. 天津统计年鉴 2013[M]. 北京: 中国统计出版社, 2013.

[15] 上海市统计局, 国家统计局上海调查总队. 上海统计年鉴 2013[M]. 北京: 中国统计出版社, 2013.

[16] 重庆市统计局, 国家统计局重庆调查总队. 重庆统计年鉴 2013[M]. 北京: 中国统计出版社, 2013.

[17] 重庆市统计局, 国家统计局重庆调查总队. 重庆统计年鉴 2016[M]. 北京: 中国统计出版社, 2016.