

农历抑或公历？数据形式对数理分析结果的影响^{*}

——以清代中后期直隶小麦市场整合分析为例

胡 鹏 李 军

内容提要:清代粮食问题研究中,绝大多数学者所使用的粮价数据是经过农历天数加权平均处理后的公历形式。从理论上讲,这种公历化处理方式存在破坏数据原貌特征的可能,进而会影响相关数理分析。本文选用《清代道光至宣统间粮价表》中的直隶小麦价格农历形式数据,对清代中后期直隶小麦价格进行个案分析:一方面,对农历数据和公历化处理数据的统计特征进行分析;另一方面,借鉴薛华(C. H. Shiue)和凯勒(W. Keller)使用的协整分析法,对小麦市场整合情况进行分析。我们发现,使用农历天数加权平均处理后的公历化数据与原始农历数据的数理分析结果差异较大,公历化数据处理方式会造成计量结论的非系统性偏误。

关键词:清代 直隶 小麦价格 数据形式 市场整合

粮食价格是清代经济史的中心议题之一,作为研究该问题基础的粮价数据,已受到学者的高度关注。目前,国内较为系统的清代粮价数据资料主要有两种:其一,王业键主持整理编制的“清代粮价资料库”(以下简称“粮价库”);^①其二,中国社会科学院经济研究所编写的《清代道光至宣统间粮价表》(以下简称“粮价表”)。^②从研究现状看,学者对“粮价库”和“粮价表”的选用存在明显偏向性,自两套粮价数据资料问世以来,除少数学者外,绝大多数学者选择“粮价库”。^③“粮价库”和“粮价表”在形式上最大的区别是,前者的粮价数据为公历形式,后者为农历形式。“粮价库”中的公历粮价数据是以农历天数加权平均的方式,对原始农历数据进行公历化处理得出。^④多数学者或直接使用“粮价库”数据,或对搜集到的原始农历数据进行农历天数加权平均处理,^⑤普遍认可这种公历化的处理方法。然而,就笔者所见,尚未有文献讨论这种数据处理方法是否会造成数理分析的偏误。从

[作者简介] 胡鹏,中国农业大学经济管理学院博士研究生,北京,100083,邮箱:drhupeng@126.com;李军,中国农业大学经济管理学院教授,北京,100083,邮箱:sirljun@cau.edu.cn。

* 本文为国家社科基金项目“历史时期气候变化对华北地区粮食生产与价格波动的影响”(批准号:13CJL008)和北京市社科基金项目“清代北京地区粮价波动与社会应对”(批准号:12LSC014)阶段性成果之一。

① 网址详见:<http://mhdb.mh.sinica.edu.tw/foodprice/>。

② 桂林:广西师范大学出版社2009年版。

③ 使用“粮价表”数据研究的如马国英、陈永福、李军《晚清山西粮食价格波动、市场成因及政府行为》(《中国经济史研究》2012年第3期)和马国英《清代华北地区粮价变动研究:1736—1911——以山西省为例》(博士学位论文,中国农业大学,2013年)等。虽然学者在其文章中并未直言选择“粮价库”的缘故,但从其研究内容和粮价使用方式上看,原因主要有三点:第一,“粮价库”覆盖的时间范围比“粮价表”更广;第二,“粮价库”中的粮价数据是公历形式的,避免了由于闰月引起的计量不便;第三,“粮价库”已实现电子化,使用时较“粮价表”便捷。

④ 参见“粮价库”的“资料库使用说明”,<http://mhdb.mh.sinica.edu.tw/foodprice/readme.php>。

⑤ 代表学者为李明珠(L. M. Li),参见L. M. Li, *Fighting Famine in North China: State, Market and Environmental Decline, 1690s - 1990s*, California: Stanford University Press, 2007。

理论上讲,农历天数加权平均的方式可能会改变原始数据的特征,进而影响相关计量结论。^①因此,有必要对这一问题深入讨论。

本文以清代中后期直隶小麦市场整合分析为载体,^②分别用农历价格数据以及经过农历天数加权平均的公历化处理数据进行计量分析,考察公历化处理数据可能引起的计量偏误。首先,设定“粮价表”中的农历粮价数据为对照组,以每月最高价和最低价的均值代表该月粮价;对“粮价表”中的农历粮价数据进行农历天数加权平均的公历化处理,并设定公历化处理数据为实验组,以每月最高价和最低价的均值代表该月粮价。其次,对比农历数据和公历化处理数据的统计特征。最后,比较两组数据市场整合计量结论的差异。在市场整合计量结论差异分析部分,本文分别采用月度、季度和年度数据,按照道光至宣统朝(1821年2月—1911年7月,以下简称“时期 I”)、道光朝(1821年2月—1851年1月,以下简称“时期 II”)、咸丰至同治朝(1851年2月—1875年1月,以下简称“时期 III”)、光绪至宣统朝(1875年2月—1911年7月,以下简称“时期 IV”)等4个时间段,从市场整合规模和质量两个层面,分析清代中后期直隶小麦市场整合情况,探讨公历化处理数据引起的计量结论偏误。

一、清代粮价数据概述

在传统重农政治理念的影响下,清代统治者十分重视对粮价的调控。清代的粮价奏报制度,肇始于康熙朝后期,到乾隆朝初期形成定制。乾隆元年(1736)五月谕旨:“各省巡抚具折奏事时,可将该省米粮时价开单,就便奏闻”,^③在全国范围内推行粮价奏报制度。乾隆三年二月,以湖广总督德沛奏报的乾隆二年“湖北、湖南十二月米麦时价清单”为蓝本,清廷统一了粮价奏报单的格式,并一直沿用至清末。在内容构成方面,粮价单中的粮价以府级行政单位为单元,以区间形式记录粮价的月度数据。根据王业键和余开亮的研究,价格为产地价格或批发价格,而非零售价格;府级单元的粮价数据来源于县级粮价数据,其中区间的上限是各县中价格最高者,区间的下限是各县中价格最低者;价格的时间点不明确,没有指出是月初、月中或月末的价格,或是对全月价格的概括。由于粮价单的这些特点,利用其数据的最佳方式是以平均值代表府级单位的价格。^④

可靠性是使用清代粮价数据时无法回避的问题。全汉昇、王业键和余开亮等多数学者均认为清代粮价单所载粮价数据的可信度较高,主要理由是:首先,地方官员造假动机低。地方官员上报的粮价细册与地方实际利益没有直接关系。其次,粮价单编制简便易行。县、府级地方官员很容易从市集获取粮

① 根据农历天数加权平均的原理,月度价格变动与农历—公历月对应日期是影响数据特征和计量结论的主要因素。举例言之,1个公历月往往对应2个农历月的日期,也有个别公历月对应1个或3个农历月日期,假设相邻3个月原始价格数据的基数分别为 P_1 、 P_2 和 P_3 ,则有:农历月价格数据的权数均为1,对应3个农历月价格分别为 P_1 、 P_2 和 P_3 ,即 $P_1 \times 1$ 、 $P_2 \times 1$ 和 $P_3 \times 1$;公历月价格数据的权数分别为 α 、 β 和 γ (满足条件: $\alpha + \beta + \gamma = 1$),对应1个公历月价格为 $\alpha P_1 + \beta P_2 + \gamma P_3$ 。一方面,价格数据随时间和季节不同而有变动;另一方面,从道光元年正月(1821年2月)到宣统三年六月(1911年7月),在农历1119个月、公历1086个月中,公历月价格数据的权数 α 、 β 和 γ 共有108种不同的组合方式。也即是,在使用农历天数加权平均对价格数据进行公历化处理时,基数和权数都发生了变化,且根据笔者分析,这些变化均存在不同程度的随机性。

② 本文所用直隶的“粮价表”中的粮种名称是“麦”。清代直隶种植的麦类主要有小麦、大麦和荞麦等,其中小麦又分为红麦和白麦,价格有一定差异。根据李明珠的考察,1746年以后的粮价单中的“麦”特指白麦。所以道光至宣统间直隶粮价单中的“麦”应指地位较为重要的白麦,参见L. M. Li, *Fighting Famine in North China: State, Market and Environmental Decline, 1690s - 1990s*, California: Stanford University Press, 2007。本文以“小麦”之称代之。在区域选择上,本文使用的直隶粮价数据包括顺天府、保定府、永平府、河间府、正定府、顺德府、广平府、大名府、冀州、赵州、深州、定州、天津府、易州、遵化州和宣化府等16个府(州),不包括口北三厅和承德府。原因为:(1)“粮价表”中没有口北三厅的粮价数据。(2)对于承德府的粮价,在清代粮价奏报系统中,道光元年至道光八年三月承德府以“直隶承德府”的形式上报,包含了最高价和最低价,此后以“热河承德府”的形式上报,且只有一个价格,前后有明显差异。

③ 《乾隆元年湖南巡抚钟保奏报粮价折》,中国第一历史档案馆藏,转引自陈金陵《清朝的粮价奏报与其盛衰》,《中国社会经济史研究》1985年第3期。

④ 王业键:《清代的粮价陈报制度及其评价》,氏著:《清代经济史论文集》,台北:稻乡出版社2003年版,第1—36页;余开亮:《粮价细册制度与清代粮价研究》,《清史研究》2014年第4期。

价。最后,粮价上报时,会进行层层核实,皇帝也对粮价比较重视。^①需要说明的是,粮价单中价格连续数月不变的现象往往被学者视作数据质量降低的表现,^②但笔者并不赞同这种评价方法。正如余开亮所言,粮价单中的府级单元粮价来自于县级单位,每月最高价和最低价的样本地点并不固定,所以,最高价格和最低价格不变并不代表该府内各县的粮价没有发生变化。^③当然,尽管粮价数据整体质量较高,但随时间推移,清代后期中央政府控制力逐渐减弱,地方政府的报告质量确有可能下降。^④

对清代粮价单的整理工作始于民国年间。20世纪30年代,柳诒徵搜集整理了江苏国学图书馆、档案馆所藏光绪年间江苏各府州县逐月呈报的米价细册。以汤象龙为首的一批学者对清代档案中的社会经济资料进行大量抄录,以表格形式整理了道光至光绪年间全国26000多份粮价单,并按省装订成册。20世纪70年代以来,彭信威从《清实录》《东华录》《清史稿》中辑录出约900件米价资料,进而得出清代267年间每10年的平均数,并编制成表;全汉昇和克劳斯(R. A. Kraus)从故宫档案、《李煦奏折》和《雍正朱批谕旨》等档案中辑录出长江中下游地区的米价数据;王道瑞整理编辑了中国第一历史档案馆所藏的3万多件不同年号、地区的粮价清单件数表;蒋建平用《李煦奏折》和《朱批谕旨》等资料制作了康熙至雍正年间苏州和江浙地区米价的统计表及相应趋势图;李明珠从《军机档》《朱批奏折》和《上谕档》等档案中辑录了直隶粮价数据。^⑤具有里程碑意义的是,中国社会科学院经济所历经5年,对该所收藏的汤象龙等学者抄录的粮价表进行整理、核对和增补,于2009年出版《清代道光至宣统间粮价表》;同时,中国台湾地区“中研院”院士王业键教授主持整理编制的“清代粮价资料库”也宣告建成。此后,在两套粮价数据的基础上,清代粮价问题研究快速发展,近5年有关清代粮价问题的论文数量已超过以往论文的总数。

二、统计特征差异分析

本部分从样本观测值、数据集中趋势特征、离散趋势特征和正态分布特征等4个方面,考察相对于农历数据,月度、季度和年度频率的公历化处理数据造成的统计特征变化。^⑥

第一,样本观测值方面。数据时间跨度为道光元年正月(1821年2月)到宣统三年六月(1911年7月),共计91年,其中,农历1119个月、公历1086个月。农历月度、季度和年度数据有效观测值分

① H. Chuan and R. A. Kraus, *Mid-Ch'ing Rice Markets and Trade: An Essay in Price History*, Cambridge: Harvard University Press, 1975;王业键:《清代的粮价陈报制度及其评价》,氏著:《清代经济史论文集》,第1—36页;余开亮:《粮价细册制度与清代粮价研究》,《清史研究》2014年第4期。

② P. C. Perdue, "The Qing State and the Gansu Grain Market, 1739 - 1864", in T. G. Rawski and L. M. Li, *Chinese History in Economic Perspective*, Oakland: University of California Press, 1992, pp. 101 - 126; R. B. Marks, "Rice Prices, Food Supply, And Market Structure In Eighteenth-Century South China", *Late Imperial China*, Vol. 12, No. 2, 1991, pp. 64 - 116; L. M. Li, *Fighting Famine in North China: State, Market and Environmental Decline, 1690s - 1990s*, California: Stanford University Press, 2007;王业键等:《清代粮价资料之可靠性检定》,《清代经济史论文集(二)》,第289—315页;谢美娥:《十九世纪后半期台湾米价的可靠性评估——以清代粮价资料库所辑米价数据为主》,“纪念梁方仲教授诞辰100周年中国社会经济史研究国际学术讨论会”论文,广东广州,2008年;王玉茹、罗畅:《清代粮价数据质量研究——以长江流域为中心》,《清史研究》2013年第1期。

③ 余开亮:《粮价细册制度与清代粮价研究》,《清史研究》2014年第4期。

④ 对于如何明确衡量粮价质量,笔者尚未找到比较合适的方法。故而,只能一方面假设粮价整体质量较高,另一方面通过分时段分析,尽量降低由于可能的粮价质量变化引起的计量偏误。

⑤ 柳诒徵:《江苏各地千六百年间之米价》,《柳诒徵史学论文集续集》,上海古籍出版社1999年版,第461—496页;彭信威:《中国货币史》,上海:群联出版社1954年版,第531、542、549、560、562页;H. Chuan and R. A. Kraus, *Mid-Ch'ing Rice Markets and Trade: An Essay in Price History*, Cambridge: Harvard University Press, 1975, Table C, pp. 100 - 116;王道瑞:《清代粮价奏报制度的确立及其作用》,《历史档案》1987年第4期;蒋建平:《清代前期米谷贸易研究》,北京大学出版社1992年版,第176—179页;L. M. Li, *Fighting Famine in North China: State, Market and Environmental Decline, 1690s - 1990s*, California: Stanford University Press, 2007, Appendix 2, pp. 448 - 494.

⑥ 处理季度数据时,以立春、立夏、立秋和立冬等节气所在月价格代表该季度价格;处理年度数据时,以每年立春和立秋节气所在月代表该年度价格。处理方法参考了颜色、刘丛、徐萌等学者的方式和思路(参见《18世纪中国南北方市场整合程度比较分析——利用清代粮价数据的研究》,《经济研究》2011年第12期;《晚清铁路建设与市场发展》,《经济学(季刊)》2015年第2期)。

别为 17 040 个、5 534 个和 2 768 个,公历化处理月度、季度和年度数据有效观测值分别为 16 087 个、5 321 个和 2 594 个。无论是全省还是各府(州),公历化处理数据均会造成月度、季度和年度数据的观测值减少,其中季度数据观测值的下降幅度最小。具体差异见表 1。

表 1 公历化处理数据有效观测值变化情况

	全省		各府(州)		
	方向	幅度	变化府(州)量占比	方向	幅度区间
月度数据	↓	5.59%	100%	↓	5.44%—5.83%
季度数据	↓	3.85%	100%	↓	3.47%—4.06%
年度数据	↓	6.29%	100%	↓	5.78%—6.94%

说明:表中“↑”表示统计量增加,“↓”表示统计量减少,“↔”表示统计量不变。

第二,数据集中趋势特征方面。公历化处理数据会造成全省月度、季度和年度数据的算数均数和中位数小幅提高。对于各府(州),除顺德府和宣化府(年度数据只有宣化府)外,公历化处理数据会造成绝大多数府(州)月度、季度和年度数据的算数均数小幅提高;对各府(州)中位数的影响各有不同,且个别府(州)的中位数未发生变化(永平府、河间府、顺德府、冀州和定州的月度数据,顺天府、顺德府和定州的季度数据,冀州的年度数据)。具体差异见表 2。

表 2 公历化处理数据集中趋势特征变化情况

	全省				各府(州)					
	算数均数		中位数		算数均数			中位数		
	方向	幅度	方向	幅度	变化府(州)量占比		幅度区间	变化府(州)量占比		幅度区间
					正向	负向		正向	负向	
月度数据	↑	0.11%	↑	0.23%	87.5%	12.5%	0.01%—0.24%	43.8%	25.0%	0.12%—0.74%
季度数据	↑	0.29%	↑	0.57%	87.5%	12.5%	0.01%—0.68%	56.3%	25.0%	0.02%—1.00%
年度数据	↑	0.44%	↑	0.45%	93.8%	6.3%	0.07%—0.84%	68.8%	18.8%	0.11%—1.52%

说明:表中“↑”表示统计量增加,“↓”表示统计量减少,“↔”表示统计量不变。

第三,数据离散趋势特征方面。公历化处理数据对全省月度、季度和年度数据的全距和标准差的影响各不相同,且相对而言,对全省季度数据的影响较小。对于各府(州),公历化处理数据对月度、季度和年度数据的全距和标准差的影响均有所不同,且个别府(州)的全距未发生变化(河间府、顺德府、深州和定州的月度数据,深州和遵化州的季度数据,永平府、冀州、深州和遵化州的年度数据)。具体差异见表 3。

表 3 公历化处理数据离散趋势特征变化情况

	全省				各府(州)					
	全距		标准差		全距			标准差		
	方向	幅度	方向	幅度	变化府(州)量占比		幅度区间	变化府(州)量占比		幅度区间
					正向	负向		正向	负向	
月度数据	↓	0.67%	↓	0.17%	0	68.8%	0.39%—3.77%	12.5%	87.5%	0.06%—0.76%
季度数据	↔	0	↑	0.09%	18.8%	68.8%	0.02%—5.17%	43.8%	56.2%	0.08%—1.62%
年度数据	↑	0.31%	↑	0.43%	12.5%	62.5%	0.02%—4.58%	62.5%	37.5%	0.04%—1.77%

说明:表中“↑”表示统计量增加,“↓”表示统计量减少,“↔”表示统计量不变。

第四,正态分布特征方面。公历化处理数据均会小幅降低全省月度、季度和年度数据的偏度系数和峰度系数(未造成偏态方向的改变)。对于各府(州),公历化处理数据对月度、季度和年度数据的偏度系数和峰度系数的影响各不相同,且对偏度系数的影响大于峰度系数。具体差异见表 4。

表 4

公历化处理数据正态分布特征变化情况

	全省				各府(州)						
	偏度系数		峰度系数		偏度系数			峰度系数			
	方向	幅度	方向	幅度	变化府(州)量占比		幅度	变化府(州)量占比		幅度	
					正向	负向		正向	负向		
月度数据	↓	3.05%	↓	0.63%	6.3%	93.8%	0.16%—10.66%	12.5%	87.5%	0.18%—2.47%	
季度数据	↓	4.26%	↓	0.67%	18.8%	81.3%	0.22%—19.33%	12.5%	87.5%	0.34%—3.54%	
年度数据	↓	3.61%	↓	0.17%	37.5%	62.5%	0.54%—20.64%	37.5%	62.5%	0.01%—3.70%	

说明:表中“↑”表示统计量增加,“↓”表示统计量减少,“↔”表示统计量不变。

通过以上四方面的比较可知,农历天数加权平均的公历化处理方式,改变了粮价数据的统计特征,具有影响范围广、影响方向和程度不定等特点。

三、市场整合计量结论差异分析

本部分以清代中后期直隶小麦价格为基础,借鉴薛华和凯勒的协整分析法,^①从市场整合的规模和质量两个层面,^②考察相对于农历数据,公历化处理数据引起的市场整合计量结论变化。

同阶单整是协整分析的前提条件,本文选取 ADF 检验确定价格序列的单整阶数。^③ 检验结果显示,公历化处理数据方法不但会影响 ADF 统计量,甚至会造成序列平稳性的改变。如表 5 所示,除“时期 I”的季度数据外,公历化处理数据会造成部分府(州)价格序列平稳性的改变。

表 5

公历化处理数据造成平稳性变化的府(州)数量

	时期 I	时期 II	时期 III	时期 IV
月度数据	1	5	1	1
季度数据	0	5	2	2
年度数据	1	—	—	—

① 清代市场整合研究中使用的计量方法主要有相关系数法和协整分析法。其中,相关系数法由于其存在诸多缺陷,在国际学界的相关研究中已基本被协整分析法所代替(P. L. Fackler and B. K. Goodwin, “Spatial Price Analysis”, in B. L. Gardner and G. C. Rausser, *Handbook of Agricultural Economics (Book 1B)*, North Holland: Amsterdam, 2001, pp. 971—1024; G. Federico, “How Much do We Know about Market Integration in Europe?” *The Economic History Review*, Vol. 65, No. 2, 2012, pp. 470—497)。薛华和凯勒最早将协整分析法应用于清代中国市场整合研究,该方法以 Engle-Granger 双变量协整检验为基础,对残差进行 ADF 检验,并以残差的 t 统计量作为判断协整程度的标准(C. H. Shiue and W. Keller, “Markets in China and Europe on the Eve of the Industrial Revolution”, *The American Economic Review*, Vol. 97, No. 4, 2007, pp. 1189—1216)。近年来,这种方法也被国内学者在历史时期市场整合研究中广泛使用(参见 L. Yu, *Two Essays on Price Movement across China's Regions*, Hong Kong: University of Hong Kong, 2010; 赵留彦、赵岩、窦志强《“裁厘改统”对国内粮食市场整合的效应》,《经济研究》2011 年第 8 期;冯颖杰《“裁厘改统”与民国时期市场整合——基于上海、芜湖、天津三地粮价的探讨》,《经济学(季刊)》2012 年第 1 期;朱琳《乾嘉道时期淮河流域粮价研究(1736—1850)》,博士学位论文,南开大学,2014 年;赵伟洪《清乾隆朝湖南省米谷流通与市场整合》,《中国经济史研究》2015 年第 1 期)。因此,本文亦选取基于协整分析法的市场整合计量结果进行论述。但是,正如颜色和刘丛、费德里科(Federico)指出,这种方法并不严格,存在一定问题。如由于该方法通过 OLS(普通最小二乘法)实现回归,存在自变量外生性假设,造成结果对变量顺序比较敏感,对于一组双变量,调整自变量和因变量的顺序不但会得到不同的 t 统计量,甚至在判断二者是否存在协整关系时出现相反的结论(参见颜色、刘丛《18 世纪中国南北方市场整合程度比较分析——利用清代粮价数据的研究》,《经济研究》2011 年第 12 期;G. Federico, “How Much do We Know about Market Integration in Europe?”, *The Economic History Review*, Vol. 65, No. 2, 2012, pp. 470—497)。有鉴于此,本文将相关回归结果分别归于双变量各自的统计中,不进行统一判定。此外,对于协整方程中确定成分的设定,根据一价定律和“冰山成本理论”,应当包含截距项;清代中后期处于中国社会从传统向现代转变的过渡期,技术进步、制度变迁等因素逐渐发挥作用,所以假设存在由技术进步等因素带来的趋势项。

② 以存在协整关系的组数作为各府(州)市场整合规模的指标,以存在协整关系各组的 t 统计量均值作为各府(州)市场整合质量的指标。

③ 为保持数据原貌、减少其他因素干扰,本文未对缺失数据进行估算和填补;计量分析所用数据取自然对数形式;由于分析对象以及所用计量方法对季节性因素未有特殊要求,故不对数据进行季节效应控制;在单位根检验和协整分析中涉及判断最优滞后后期时,均选用 AIC 信息准则加以确定;由于数据缺失,有效年度数据量有限,年度数据计量结论仅按照“时期 I”一个时间段进行比较分析。

根据 ADF 检验结果,协整分析的对象包括:“时期 I”的 240 组府(州)的月度、季度和年度数据,“时期 II”的 132 组府(州)的月度数据和 240 组府(州)的季度数据,“时期 III”的 240 组府(州)的月度数据和 182 组府(州)的季度数据,“时期 IV”的 240 组府(州)的月度和年度数据。^①

对于清代中后期直隶全省小麦市场整合情况,如表 6 所示,在市场整合规模方面,公历化处理数据均会低估市场整合的规模;但低估幅度有所不同,且低估幅度没有特定规律。在市场整合质量方面,公历化处理数据对评估市场整合质量均存在影响,且影响方向(高估或低估)有所不同、没有特定规律;影响幅度(高估或低估幅度)也有所不同,但会随显著水平标准的提高而降低。

表 6 公历化处理数据对全省市场整合规模和质量计量结论的影响统计

		1% 显著水平				5% 显著水平				10% 显著水平			
		规模		质量		规模		质量		规模		质量	
		方向	幅度	方向	幅度	方向	幅度	方向	幅度	方向	幅度	方向	幅度
时期 I	月度数据	↓	30.9%	↑	2.99%	↓	6.0%	↓	1.10%	↓	8.2%	↓	0.57%
	季度数据	↓	21.4%	↓	5.35%	↓	7.5%	↓	5.59%	↓	5.6%	↓	5.05%
	年度数据	↓	20.0%	↓	4.04%	↓	12.8%	↓	4.48%	↓	14.6%	↓	3.30%
时期 II	月度数据	↓	30.0%	↓	3.61%	↓	8.0%	↓	2.42%	↓	7.5%	↓	1.62%
	季度数据	↓	30.0%	↑	6.31%	↓	31.0%	↑	2.94%	↓	25.9%	↑	1.16%
时期 III	月度数据	↓	20.0%	↑	4.40%	↓	21.1%	↑	2.83%	↓	13.5%	↑	1.32%
	季度数据	↓	66.7%	↑	11.77%	↓	73.3%	↑	3.75%	↓	50.0%	↓	2.17%
时期 IV	月度数据	↓	29.6%	↓	0.12%	↓	9.7%	↓	0.35%	↓	9.4%	↓	0.10%
	季度数据	↓	17.6%	↓	2.24%	↓	6.3%	↓	1.68%	↓	19.5%	↑	0.72%

说明:表中“↑”表示统计量增加,“↓”表示统计量减少,“↔”表示统计量不变。

对于清代中后期直隶各府(州)小麦市场整合规模情况,如表 7 所示,不同数据频率、时间段和显著水平下,公历化处理数据会影响部分府(州)的计量结论,但影响的府(州)范围并不确定。同时,公历化处理数据对各府(州)计量结论的影响方向(高估或低估)也有所不同,且多数府(州)会被低估。

表 7 公历化处理数据对府(州)市场整合规模计量结论的影响范围统计

		1% 显著水平			5% 显著水平			10% 显著水平		
		一致	高估	低估	一致	高估	低估	一致	高估	低估
时期 I	月度数据	25.0%	6.3%	68.8%	31.3%	12.5%	56.3%	31.3%	12.5%	56.3%
	季度数据	18.8%	18.8%	62.5%	31.3%	18.8%	50.0%	37.5%	18.8%	43.8%
	年度数据	18.8%	18.8%	62.5%	12.5%	31.3%	56.3%	18.8%	18.8%	62.5%
时期 II	月度数据	58.3%	8.3%	33.3%	33.3%	25.0%	41.7%	25.0%	25.0%	50.0%
	季度数据	31.3%	25.0%	43.8%	25.0%	12.5%	62.5%	18.8%	12.5%	68.8%
时期 III	月度数据	81.3%	6.3%	12.5%	50.0%	18.8%	31.3%	56.3%	12.5%	31.3%
	季度数据	85.7%	0.0%	14.3%	42.9%	7.1%	50.0%	42.9%	7.1%	50.0%
时期 IV	月度数据	43.8%	12.5%	43.8%	31.3%	25.0%	43.8%	25.0%	25.0%	50.0%
	季度数据	68.8%	12.5%	18.8%	43.8%	18.8%	37.5%	18.8%	12.5%	68.8%

① 为便于比较分析,在使用 ADF 检验判断数据序列的平稳性时,采用 1% 显著水平下的平稳性判断结论,并取农历和公历化处理数据序列平稳性一致的府(州)进行协整分析。举例言之,在“时期 I”的月度数据中,ADF 检验结果显示,顺德府农历数据的水平序列非平稳、一阶差分序列在 1% 显著水平下平稳,公历化处理数据的水平序列在 5% 显著水平下平稳、一阶差分序列在 1% 显著水平下平稳,此时采用 1% 显著水平下的平稳性判断结论,即认为顺德府的农历和公历化处理数据均为 I(1)序列,可以进行协整分析。由此,可以进行协整分析的府(州)数量为:“时期 I”的月度、季度和年度数据为 16 个,“时期 II”的月度和季度数据分别为 12 个和 16 个,“时期 III”的月度和季度数据分别为 16 个和 14 个,“时期 IV”的月度和季度数据均为 16 个。

对比表 7 和表 8 可以发现,公历化处理数据和农历数据测算各府(州)市场整合规模结论在某些条件下呈现一致性较强的现象,很大程度上可归因于部分府(州)间不存在协整关系。统计数据显示,计量结论一致的府(州)范围超过 40% 的统计中,均不存在协整关系府(州)量的贡献。

表 8 各府(州)间不存在协整关系对农历和公历化处理数据测算市场整合规模结论一致的贡献率

	1% 显著水平				5% 显著水平				10% 显著水平			
	时期 I	时期 II	时期 III	时期 IV	时期 I	时期 II	时期 III	时期 IV	时期 I	时期 II	时期 III	时期 IV
月度数据	25.0%	57.1%	76.9%	42.9%	20.0%	0	37.5%	0	0	0	22.2%	0
季度数据	100.0%	60.0%	91.7%	45.5%	20.0%	25.0%	50.0%	14.3%	16.7%	0	50.0%	0
年度数据	66.7%	—	—	—	0	—	—	—	0	—	—	—

对于清代中后期直隶各府(州)小麦市场整合质量情况,如表 9 所示,与对市场整合规模的计量结论相似,公历化处理数据会影响部分府(州)有关市场整合质量的计量结论,但影响的府(州)范围并不确定。同时,公历化处理数据对各府(州)市场整合质量的计量结论的影响方向(高估或低估)也有所不同,且会低估多数府(州)的市场整合质量。必须指出的是,公历化处理数据对市场整合规模和质量计量结论的影响虽然整体相似,但具体情况并不相同,且不存在明显的相关性。

表 9 公历化处理数据对府(州)市场整合质量计量结论的影响范围统计

		1% 显著水平		5% 显著水平		10% 显著水平	
		高估	低估	高估	低估	高估	低估
		时期 I	月度数据	50.0%	43.8%	6.3%	87.5%
	季度数据	6.3%	75.0%	18.8%	75.0%	18.8%	75.0%
	年度数据	12.5%	75.0%	6.3%	93.8%	18.8%	81.3%
时期 II	月度数据	41.7%	25.0%	41.7%	58.3%	50.0%	50.0%
	季度数据	50.0%	31.3%	43.8%	50.0%	50.0%	50.0%
时期 III	月度数据	18.8%	18.8%	43.8%	37.5%	50.0%	37.5%
	季度数据	0	21.4%	7.1%	71.4%	14.3%	64.3%
时期 IV	月度数据	37.5%	43.8%	62.5%	37.5%	56.3%	43.8%
	季度数据	25.0%	43.8%	56.3%	37.5%	56.3%	43.8%

对于清代中后期直隶各府(州)小麦市场整合规模和质量的影响程度情况,如表 10 所示,公历化处理数据对被影响府(州)的市场整合规模和质量计量结论的影响程度均有所不同,且没有特定规律。但整体而言,对市场整合规模计量结论的影响程度大于市场整合质量计量结论。

表 10 公历化处理数据对府(州)市场整合规模和质量计量结论的影响程度统计

		1% 显著水平		5% 显著水平		10% 显著水平	
		规模	质量	规模	质量	规模	质量
		时期 I	月度数据	28.39%	3.99%	23.85%	8.91%
	季度数据	33.04%	5.44%	33.59%	5.69%	20.07%	5.53%
	年度数据	42.48%	4.14%	30.33%	4.37%	32.72%	4.87%
时期 II	月度数据	5.63%	5.63%	6.94%	6.94%	5.31%	5.31%
	季度数据	83.33%	10.80%	48.31%	9.42%	32.75%	5.46%
时期 III	月度数据	—	4.00%	66.67%	4.68%	42.78%	7.15%
	季度数据	—	6.28%	—	4.63%	46.67%	7.33%
时期 IV	月度数据	50.00%	3.13%	41.08%	3.53%	28.47%	4.08%
	季度数据	37.50%	3.98%	46.96%	4.18%	33.00%	3.85%

说明:表中“规模”项不包括以下三类变化:一是农历和公历化处理数据协整关系数量相同(变化幅度 0);二是农历数据结果显示存在协整关系,且公历化处理数据结果显示不存在协整关系(变化幅度 100%);三是农历数据结果显示不存在协整关系,且公历化处理数据结果显示存在协整关系(变化幅度 ∞)。表中“质量”项不包括以下三类变化:一是农历和公历化处理数据结果均显示不存在协整关系;二是农历数据结果显示存在协整关系,且公历化处理数据结果显示不存在协整关系(变化幅度 100%);三是农历数据结果显示不存在协整关系,且公历化处理数据结果显示存在协整关系(变化幅度 ∞)。

通过以上表10统计信息及相关分析可知,就清代中后期直隶小麦市场而言,公历化处理数据对市场整合情况计量结论有较大影响,且具体影响范围和程度比较复杂、没有明显的规律性。因此,可以较明确地判断:农历月加权平均的公历化数据处理方式会造成相关计量结论的非系统性偏误。

四、结论与思考

本文以清代中后期直隶小麦市场整合分析为例,借鉴薛华和凯勒的协整分析法,考察相对于农历原始数据,公历化处理数据对市场整合计量结果的影响。分析发现,公历化处理数据不但会改变数据自身的统计特征、平稳性,也会影响对直隶全省和各府(州)市场整合情况的计量结论。同时,公历化处理数据对计量结论的影响范围和程度均有所不同,且没有特定的规律。^①

因此,在使用粮价数据研究粮食问题时,需要充分考虑到价格数据的形式,目前以公历形式粮价数据为基础的清代粮食问题研究结论的准确性尚需进一步商讨。当然,本文仅以清代中后期直隶小麦市场为中介进行个例分析,未对相关问题作系统的理论推导,可能影响了所得结论的普遍适用性。但就公历化处理数据会引起数理分析偏误的问题本身而言,值得研究者予以重视。

基于本文分析观点,笔者对涉及清代粮价数据的研究有以下四点建议:第一,在研究清代粮价问题时,应尽量使用原始的农历数据形式。第二,对于“数据库”中公历形式数据的使用,可做相应逆向计算得到原始农历形式数据后进行相关分析。^② 第三,对于农历闰月问题,可采取闰月和本月取均值合并的处理方法,避免由闰月引起的计量不便。^③ 第四,对于相关比较研究(尤其是中西比较),可以月度单位为基础,对整体相当的时间段进行分析。^④

The Influence of Data Form on the Result of Mathematical Analysis: The Integration of Wheat Market in Zhili Province(1821 - 1911)

Hu Peng Li Jun

Abstract: Most of researchers use the data of Gregorian Calendar form in price study in Qing Dynasty. However, the original data is Chinese Lunar Calendar, and theoretically, change data to the Gregorian Calendar form can make the result of mathematical analysis different. Using the data in *Food Price Lists of Daoguang to Xuantong in Qing Dynasty*, this paper analyzes the wheat price in Zhili Province(1821 - 1911) on aspects of statistical characterization and market integration which uses Shiue and Keller's method. Comparing the differences of mathematical results based on the data form of Chinese Lunar Calendar and Gregorian Calendar, we find that change data to the Gregorian Calendar form can make non-systematic error, and therefore, price studies based on the data of the Gregorian Calendar form in Qing Dynasty should be reanalyzed.

Key Words: Qing Dynasty; Zhili Province; Wheat Price; Data Form; Market Integration

(责任编辑:丰若非)

① 因为本文的核心目标是分析数据形式对数理分析结果的影响,并非在于判断清代中后期直隶小麦市场是否整合、整合程度如何,抑或是比较各种市场整合计量方法的优劣,故而文中观点仅对相关计量结论本身负责。受篇幅所限,本文未附相关计量结果,如有需要可通过电子邮箱向作者索取。

② 从理论上讲可以实现这种转换,罗畅的相关比较分析也反映了其可行性,正常情况下不会有太大出入。参见罗畅《两套清代粮价数据的比较与使用》,《近代史研究》2012年第5期。

③ 道光至宣统朝共有1119个农历月,其中闰月33个,合并后仅有不到3%的数据发生变化,这是目前笔者所见对数据原貌破坏程度最小的方法。

④ 这种方式对独立主体间的比较研究影响较小,但对主体间进行相关性比较研究的影响较大。如对中国和西欧各自市场情况比较研究结果不会有较大影响,但若考察中国与东南亚地区市场间联系问题,由于历法规则差异可能会在时间维度上影响相关研究结果。