

DOI: 10.13718/j.cnki.xdzk.2022.10.012

超越对数成本函数的模型设定研究

——来自我国肉制品企业面板数据的实证分析

杨琴¹, 王鹏¹, 徐成波², 郭毅¹

1. 四川旅游学院 旅游文化产业学院, 成都 610100; 2. 四川师范大学 经济与管理学院, 成都 610101

摘要: 超越对数成本函数在食品安全生产实证研究中起着重要作用。在拓展之后的框架下, 文章结合面板数据对扰动项进行稳健标准差检验, 并对模型形式的选择进行了检验; 同时, 为保留无法观测的食品安全变量, 对固定效应模型作出展开式的一种新表达; 运用幂阶梯转换同时处理严重多重共线性和模型识别问题。在采用面板似不相关迭代回归和最小距离估计之后, 得到系统简化模型的结果, 发现食品安全内生时的模型更优越; 由模型模拟出食品安全水平的数据, 得出 2004—2007 年中国肉制品行业的食品安全平均水平分别为 1.206 2、1.314 5、1.191 8 和 1.232 7, 且数据具有良好的分布。本文的研究对超越对数成本函数的模型设定可提供重要借鉴。

关键词: 超越对数成本函数模型设定; 模型形式选择检验;

严重多重共线性处理; 模型识别处理; 食品安全

中图分类号: F064.1

文献标志码: A

开放科学(资源服务)标识码(OSID):



文章编号: 1673-9868(2022)10-0101-13

Study on the Model Setting of the Translog Cost Function

——An Empirical Analysis of Panel Data from Meat Enterprises in China

YANG Qin¹, WANG Peng¹, XU Chengbo², GUO Yi¹

1. School of Tourism and Culture Industry, Sichuan Tourism University, Chengdu 610100, China;

2. School of Economics and Management, Sichuan Normal University, Chengdu 610101, China

Abstract: The translog cost function plays an important role in the empirical study of safe food production. Under the expanded framework, this study tested the robust standard deviation of the perturbation term and the choice of the model form with combined panel data. At the same time, in order to retain the unobservable food safety variables, a new expression of the expansion type for fixed effect model was made.

收稿日期: 2021-10-26

基金项目: 四川旅游学院校级项目(2021SCTUSK29); 四川省科技厅软科学研究项目(2018ZR0087)。

作者简介: 杨琴, 硕士, 讲师, 主要从事食品安全监管等研究。

通信作者: 王鹏, 副研究员。

The power ladder transformation was used to simultaneously deal with severe multicollinearity and model identification problems. After adopting panel seemingly unrelated iterative regression and minimum distance estimation, the systematic simplified model was obtained. The study found that the endogenous model of food safety is superior. The data of the food safety level was simulated with the model, and the average food safety level of the meat industry in China from 2004 to 2007 was 1.206 2, 1.314 5, 1.191 8 and 1.232 7, respectively, and the data has a good distribution. The study in this paper may provide an important reference for setting the model of the translog cost function.

Key words: model setting of the translog cost function; model form selection testing; severe multicollinearity treatment; model identification treatment; food safety

目前,对企业食品安全生产实证研究比较典型的做法是用学术界提出的质量调整成本模型(Quality-adjusted Cost Model)来刻画,并将其设定为超越对数成本函数(Translog Cost Function)形式.近年来,这一函数在国外得到了广泛应用:Kitenge 运用该方法分析食品和农产品进口对美国农业部门的影响^[1],Shin 等分析了近海渔业的规模经济^[2],Bhattacharya 对美国烟草业技术创新进行了分析^[3],Modrego 等研究了“新冠”疫情对智利就业的影响^[4],Bagadeem 通过研究指出沙特电信行业规模报酬递增^[5],Mazi-otis 等研究了 2010—2017 年智利水和排污行业规模经济和范围经济的存在性^[6].国内学者也运用该函数展开相应的研究,陈林等对上市零售企业的规模经济效应进行测度^[7],章玉等分析了财政补贴对公交企业成本函数和生产率的变动效果和影响程度^[8],查冬兰等对我国工业部门能源与非能源要素之间的替代关系进行了分析^[9],林善浪等检验了劳动力转移与农业机械化之间的关系^[10],陈林对混合所有制改革进行了分析^[11].笔者之一曾对这一模型进行了拓展,发展出该模型的一般形式,即超越对数成本面板方程,对文献中关于该模型设定的各种形式作出了有力解释^[12];并对模型的特征进行了分析,包括:超越对数成本方程为二阶泰勒展开式、超越对数成本方程相关参数约束条件、系统结构模型扰动项方差—协方差矩阵为奇异矩阵及其处理等^[13].目前,学术界对该模型的设定还未作分析,为此,本文将从模型形式选择检验、严重多重共线性处理和模型可识别处理等方面作进一步研究,以期对超越对数成本函数的模型设定提供相应借鉴.

1 数据来源、变量说明和模型设定

1.1 数据来源和变量说明

本文选取的样本数据来源于中国工业企业数据库,考虑到数据的可获得性,最终选取了 2004—2007 年度样本量最大的肉制品及副产品加工企业(小类编码为 1 352),它隶属于屠宰及肉类加工(中类编码为 135),归口于农副食品加工业(大类编码为 13),属于制造业(C 门类).定义相关指标:销售收入为 POR ,利润总额为 TP ,应交所得税为 ITP ,本年工资为 WP ,福利费为 WEL , K 为资产,直接材料为 DM ,两种生产要素分别为劳动 L 、原材料 M ,两种要素投入所占成本份额分别表示为 S_l 和 S_m .相应地,笔者之一的前期成果对相关变量的界定进行了详细讨论^[12],总成本 $C = POR - TP + ITP$,工资 $W_l = WP + WEL$,原材料投入量 $W_m = DM$,产品质量 $Q = TP/K$,需求变动因素 Z 由企业所在地区的城镇居民人均可支配收入表示, $S_l = W_l/C$ 、 $S_m = W_m/C$.本文采用 Stata 11 软件处理数据、编写相关程序、得出回归结果(限于篇幅,本文省略了相关数学证明、运行程序和部分回归结果).

1.2 模型设定

在前期研究中,笔者在 Braeutigam 等^[14]、Gertler 等^[15]提出的框架下,结合选取的样本数据和我国企业的特殊情况对该框架作了拓展,在此继续沿用这一成果:质量调整成本函数为 $C = C(Y, S, Q, W, K)$,

均衡食品安全方程为 $S = S(Q, W, K, P, Z)$, 其中, C 为总成本, Y 为总产量, S 为食品安全, Q 为产品质量, W 为要素价格, K 为资产, P 为产品价格, Z 为需求变动因素. 在研究中需解决质量调整成本函数的模型形式设定. 一方面, 从目前国内外的研究来看, 各种文献采取了超越对数成本函数形式, 该函数是经验研究中最频繁使用的灵活函数形式, 在众多奇异的函数形式中成为最可靠与最受欢迎的函数^[16]. 其特征为: 包含每个解释变量的一次项、二次项和变量之间的交互项. 另一方面, 该函数的难点在于产品质量和食品安全均为无法观测到的变量. 其中, 学术界已解决了产品质量的量化问题^[17], 却未分析食品安全的量化处理. 为此, 我们试着将涉及食品安全变量的各项均纳入一个统一的“方框”变量 α_i , 这正是面板数据模型的特征. 考虑 $h = 1, \dots, H$ 种生产要素的情况, 最终将质量调整成本函数拓展为超越对数成本面板方程:

$$\begin{aligned} \ln C_{it} = & \alpha_0 + \alpha_y \ln Y_{it} + 1/2 \alpha_{yy} (\ln Y_{it})^2 + \alpha_{yk} \ln Y_{it} \ln K_{it} + \sum_{h=1}^H \alpha_{yw \cdot h} \ln Y_{it} \ln W_{it \cdot h} + \alpha_{yq} \ln Y_{it} \ln Q_{it} + \\ & \beta_k \ln K_{it} + 1/2 \beta_{kk} (\ln K_{it})^2 + \sum_{h=1}^H \beta_{kw \cdot h} \ln K_{it} \ln W_{it \cdot h} + \beta_{kq} \ln K_{it} \ln Q_{it} + \\ & \sum_{h=1}^H \gamma_{w \cdot h} \ln W_{it \cdot h} + 1/2 \sum_{h=1}^H \sum_{j=1}^H \gamma_{ww \cdot hj} \ln W_{it \cdot h} \ln W_{it \cdot j} + \sum_{h=1}^H \gamma_{wq \cdot h} \ln Q_{it} \ln W_{it \cdot h} + \\ & \theta_q \ln Q_{it} + 1/2 \theta_{qq} (\ln Q_{it})^2 + \alpha_i + \sum_{h=1}^H \epsilon_h \ln W_{it \cdot h} + v_{it} \end{aligned} \quad (1)$$

式(1)中, i 为观测值个体, $i = 1, \dots, N$; t 为时间, $t = 2004, \dots, 2007$; $h = 1, \dots, H$ 种要素; v_{it} 为随机扰动项; α_i 为所有涉及食品安全变量的项, 等于 $\alpha_{ys} \ln Y_{it} \ln S_{it} + \beta_{ks} \ln K_{it} \ln S_{it} + \sum_{h=1}^H \gamma_{ws \cdot h} \ln S_{it} \ln W_{it \cdot h} + \theta_{qs} \ln Q_{it} \ln S_{it} + \eta_s \ln S_{it} + 1/2 \eta_{ss} (\ln S_{it})^2$. 由式(1)可导出劳动的成本份额方程:

$$S_l = \frac{\partial \ln C_{it}}{\partial \ln W_{it \cdot l}} = \gamma_{w \cdot l} + \gamma_{ww \cdot ll} \ln W_{it \cdot l} + \gamma_{ww \cdot lm} \ln W_{it \cdot m} + \alpha_{yw \cdot l} \ln Y_{it} + \gamma_{wq \cdot l} \ln Q_{it} + \beta_{kw \cdot l} \ln K_{it} + \gamma_{ws \cdot l} \ln S_{it} + \epsilon_l \quad (2)$$

式(1)和(2)组成系统结构模型, 采用似不相关估计提高估计的效率. 其中, 食品安全 S 是无法观测到的, 需要把均衡食品安全方程代入其中, 将系统结构模型转化为系统简化模型. 经过下文的模型识别和严重多重共线性处理后, 均衡食品安全方程设定为:

$$\ln S_{it} = \tau_q \cdot Q_{it} + \tau_l \cdot \ln W_{it \cdot l} + \tau_m \cdot \sqrt[3]{W_{it \cdot m}} + \tau_k \cdot \sqrt{K_{it}} + \tau_z \cdot \ln Z_{it} + \zeta_{it} \quad (3)$$

式(1)为面板方程, 那么首先面临着形式选择问题: 混合模型 (POOL)、固定效应模型 (FE) 和随机效应模型 (RE), 下面对此进行检验.

2 模型形式选择检验

结合样本数据, 在式(1)中检验的基本思路为对 3 种形式进行两两对比. 在对模型形式选择问题分析之前, 首先需要对扰动项是采用普通标准差还是稳健标准差进行检验.

2.1 稳健标准差检验

普通标准差计算方法是在假定扰动项为独立同分布 (Independently Identically Distribution, 简记 IID) 的情况下进行的, 但在现实中, 每个企业不同年份之间的扰动项一般存在自相关 (即组内自相关), 因此, 在这种情况下普通标准差的估计并不准确. 为此, 需要检验式(1)的扰动项是否存在自相关. 对于这个问题, Wooldridge 提出了一个对组内自相关的检验方法^[18], 其原假设表达为 H_0 : 式(1)不存在一阶自相关, 备择假设表达为 H_1 : 式(1)存在一阶自相关. 根据运行结果, $p = 0.0028 < 0.01$, 故显著拒绝原假设, 这说明式(1)中的扰动项存在一阶组内自相关. 所以, 在检验中, 将统一采用聚类变量为“肉制品企业法人代表单位

代码”的聚类稳健标准差. 下面分别对模型形式的选择进行检验.

2.2 POOL 与 FE 检验

在这两类模型检验中, 原假设为 $H_0: \text{式}(1)v_{it}=0$ (POOL 模型), 备择假设为 $H_1: \text{式}(1)v_{it} \neq 0$ (FE 模型). 根据运行结果, $p=0.000\ 0 < 0.01$, 故在 1% 水平上拒绝原假设, 即存在个体效应(选择 FE 模型). 需要说明的是, 由于未使用聚类稳健标准误, 故并不能保证这个 F 检验的有效性; 不过由于 p 值很小(接近 0), 即使按聚类标准差来计算 F 值, 大致也能拒绝原假设. 为了使这一结论得到肯定性的验证, 可以进一步采用“最小二乘虚拟变量模型”(Least Square Dummy Variable Model, 简记 LSDVM)进行考察. 从回归结果来看, 大多数个体虚拟变量均都显著($p < 0.1$), 可以确信拒绝“所有个体虚拟变量都为 0”的原假设, 即存在个体效应, 应选择 FE 模型.

2.3 POOL 与 RE 检验

在此检验中, Breusch 和 Pagan 提供了一个检验个体效应的 LM 检验^[19]. 原假设为 $H_0: \text{式}(1)$ 随机扰动项的方差 $\sigma_v^2=0$ (不存在个体效应, 即 POOL), 备择假设为 $H_1: \text{式}(1)\sigma_v^2 \neq 0$ (存在个体效应, 即 RE). 根据运行结果, $p=0.000\ 0 < 0.01$. LM 检验拒绝原假设, 因此, 应该选择 RE 模型.

2.4 FE 与 RE 检验

无论是从 POOL 与 FE 之间的检验还是从 POOL 与 RE 之间的检验来看, 都得出式(1)存在个体效应的结论. 接下来是在 FE 与 RE 之间进行选择检验, 需要注意的是, 如果采用聚类稳健标准差, 则传统的豪斯曼检验会失效. 我们采取 Bootstrap 自助法来解决这一问题.

根据 Hausman 提出的 FE 与 RE 模型检验的基本原理^[20], 在式(1)中, 检验的原假设为 $H_0: \alpha_i$ 与各解释变量不相关(RE), 备择假设为 $H_1: \alpha_i$ 与各解释变量相关(FE). 在备择假设 H_1 情况下, $\hat{\beta}_{FE}$ 为一致估计量, 但 $\hat{\beta}_{RE}$ 是不一致的^[21]. 事实上, 可将原假设和备择假设的检验逻辑思维进行等价转换, 原假设为 $H_0: \alpha_i$ 与各解释变量不相关(RE) $\Leftrightarrow H_0: (\hat{\beta}_{FE} - \hat{\beta}_{RE}) \xrightarrow{p} 0$, 即检验 RE 相当于检验 FE 估计量与 RE 估计量之差依概率收敛于 0; 备择假设为 $H_1: \alpha_i$ 与各解释变量相关(FE) $\Leftrightarrow H_1: (\hat{\beta}_{FE} - \hat{\beta}_{RE})$ 难以依概率收敛于 0. 这种转换将为我们编写 Bootstrap 程序提供了基本依据. Cameron 和 Trivedi 认为, 在进行假设检验或区间估计时, 自助样本的个数取 999^[22]. 因此, 本文将进行 999 次自助抽样. 根据运行结果, 得出包含常数项在内的 21 个解释变量的卡方值为 28.96; 检测统计量 $\chi^2(20, 28.96)$ (20 为 $\hat{\beta}_{FE}$ 的维度, 即随时间而变的解释变量个数, 不包括常数项)的 $p=0.088\ 5 < 0.1$, 即在 10% 显著性水平上拒绝原假设, 从而接受“ α_i 与各解释变量相关”的备择假设, 这表征式(1)应该采用固定效应模型(FE).

2.5 FE 模型的另类表达

由于式(1)采取固定效应模型, 传统估计 FE 模型的方法是消掉无法观测的变量(包括这里的食品安全), 这将影响下一步的分析. 在此, 我们转而对 FE 模型进行另类的表达, 即将式(1)中的方框变量 α_i “打开”, 具体考察 α_i 的形式, 这是本文有别于一般固定效应模型类实证研究的不同之处, 从而拓展了关于固定效应模型实证研究的视野. 当然, 需结合前文的研究结论设定 α_i 的形式: 第一, 参照 α_i 与各解释变量相关; 第二, 遵循超越对数成本函数的基本特征. 最终得到了上文 α_i 的形式.

在式(1)中, α_i 与各解释变量的相关性是通过食品安全变量与这些变量之间的交互来实现的, 这与 Gertler 和 Waldman^[15]、王志刚等^[23]的研究是相似的, 但 Antle^[17] 却对 α_i 的形式作了一些微调, 而 Mocan^[24] 则对 α_i 作了更为灵活的变换. 这表明可以根据研究的需要来灵活设定式 α_i 的形式, 至于如何设定还需要另外进行相关的专题研究.

3 严重多重共线性问题处理

从上文的推导过程来看, 要素成本份额方程式(2)由超越对数成本面板方程式(1)分别对各种要素价格

求偏微分得出, 即前者为后者的一部分, 因此, 可将式(1)看作“母方程”, 而将式(2)看作其“子方程”. 在此计量分析过程中, 经过多次实验发现, 减轻要素成本份额方程的多重共线性将有助于缓解超越对数成本面板方程的多重共线性问题, 进而将缓解由两者组成的系统结构模型的多重共线性. 为此, 归纳出以下结论:

假定 A 方程为“母方程”、B 方程为“子方程”, 那么减轻 B 方程的多重共线性将有助于缓解 A 方程的多重共线性问题, 进而缓解由 A 方程和 B 方程构成的系统结构模型的多重共线性.

该结论对于实证分析中减轻此类多重共线性问题提供了重要思路. 在实证分析中, 相较而言, “母方程”的结构要比“子方程”的结构复杂得多, 这意味着通过“子方程”解释变量的多重共线性分析可以研判“母方程”解释变量的多重共线性问题, 从而简化了分析. 因此, 我们将该问题转化为对式(2)进行多重共线性处理. 将食品安全方程式(3)带入式(2), 得到简化形式:

$$S_l = \gamma_l + \alpha_{yl} \ln Y_{it} + (\beta_{kl} + \gamma_{ls} \tau_k) \ln K_{it} + \gamma_{ll} \ln \frac{W_{it \cdot l}}{W_{it \cdot m}} + (\gamma_{lq} + \gamma_{ls} \tau_q) \ln Q_{it} + \gamma_{ls} \tau_z \ln Z_{it} + \gamma_{ls} \tau_l \ln W_{it \cdot l} + \gamma_{ls} \tau_m \ln W_{it \cdot m} + \tau_l$$

(4)

其中, $\tau_l = \epsilon_l + \gamma_{ls} \zeta_{it}$. 式(4)各解释变量之间的相关系数矩阵见表 1.

表 1 要素成本份额方程的解释变量相关系数矩阵

	lnY	lnK	ln(W _l /W _m)	lnQ	lnZ	lnW _l	lnW _m
lnY	1						
lnK	0.744 2	1					
ln(W _l /W _m)	−0.335 8	−0.012 0	1				
lnQ	0.222 8	−0.121 4	−0.141 0	1			
lnZ	0.014 6	0.059 1	0.051 5	−0.063 4	1		
lnW _l	0.751 7	0.731 7	0.321 8	0.106 9	0.079 9	1	
lnW _m	0.961 7	0.698 2	−0.452 7	0.207 3	0.036 3	0.698 6	1

从表 1 可以看出, lnY 与 lnW_m 之间的相关系数高达 0.961 7, 出现了严重的多重共线性. 经过多次实验, 变换 W_m 的形式. 在对式(4)多重共线性问题的处理过程中发现了另外一个问题, 即变量 lnK_{it} 和 lnQ_{it} 的系数估计值出现了参数线性表达的形式, 而这种形式对于参数求解造成了很大的困扰, 很难求解出各个参数, 导致系统结构模型难以识别(Identification). 进一步观察发现, 要解出系统结构模型的各个参数, 要求式(4)的参数表达式只能以“乘积”的形式出现, 即需要对式(4)中的两个线性参数表达式进行“拆分”. 以此为逻辑起点, 采取“逆推法”, 意味着在将式(3)代入式(1)的过程中, 两个式子相同变量 K 和 Q 不能采取同样的形式(对数形式), 为了与变量 W_m 形式变化取得一致性, 将变换式(3)中 K 和 Q 原本的对数形式.

从上述分析可以看出, 在对多重共线性问题的处理过程中还“掺杂”了方程的可识别问题, 从而使问题变得更为复杂, 因此, 需要将二者同时结合起来考虑.

4 模型可识别问题处理

为了同时考虑两个问题(可识别问题、多重共线性问题), 我们再次聚焦食品安全方程式(3), 并对其中的 W_m、K 和 Q 三个变量进行形式变换. 对数据进行转换时可采用两种技术: BOX-COX 转换和幂阶梯转换, 这两种转换能使数据更接近于正态分布, 其中, 自然对数为幂阶梯转换的一种或 BOX-COX 转换, 其效果是为了减少数据的正偏态分布^[25]. 从这个意义上讲, 式(1)和(3)本质上相当于是对各变量进行幂阶梯转换或 BOX-COX 转换.

4.1 BOX-COX 转换

在转换之前, 需对数据的分布进行考察, 以便采取合适的方法进行转换, 结果见图 1.

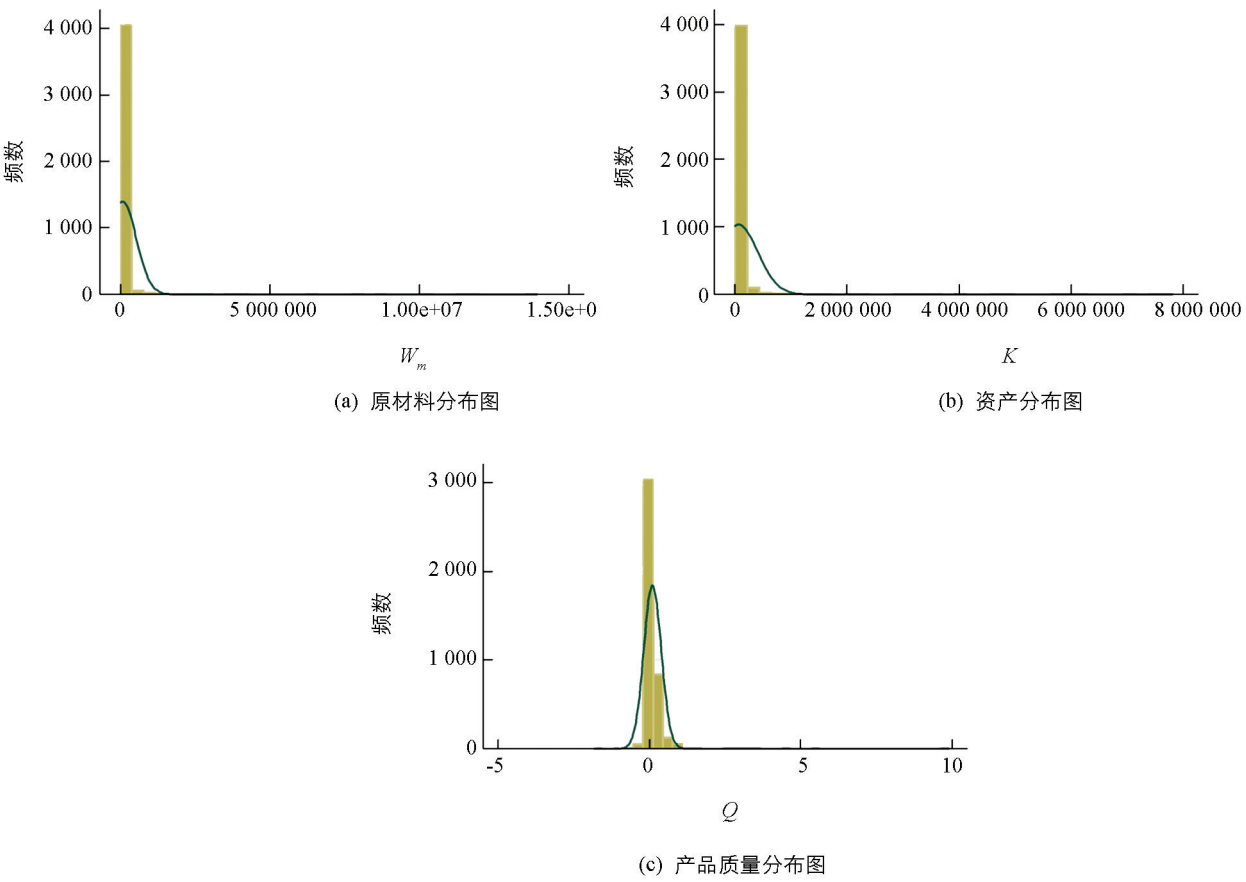


图 1 各变量 W_m 、 K 和 Q 的分布图

从图 1 可以清晰地看出, 变量 W_m 和 K 存在严重正偏态, 但变量 Q 所呈现出的正偏态分布并不是那么严重. 这是因为, 从对产品质量这一变量的处理来看(见上文的分析), Q 为一中性变量(比例性指标), 各观测值之间的数据差异本身就比较小. 为此, 先对这三个变量采取 BOX-COX 转换, 转换后的变量分别表示为 BW_m 、 BK 和 BQ , 其相关统计量见表 2.

表 2 转换以后各变量的相关统计量

转化形式	待确定的常数 L	95% 置信区间		偏度
$BW_m = (W_m^L - 1)/L$	-0.012 8	-0.019 6	-0.005 8	6.66e-05
$BK = (K^L - 1)/L$	-0.085 3	-0.095 5	-0.074 8	0.000 1
$BQ = (Q^L - 1)/L$	0.134 0	0.119 0	0.149 0	-1.68e-06

从表 2 可以看出, BW_m 、 BK 和 BQ 变量的偏度统计量近似于 0, 这说明 W_m 、 K 和 Q 进行 BOX-COX 转换以后均呈现正态分布. 这样, 我们达到了数据转换的目的, 而且系统结构模型的可识别问题也得到了解决. 在此基础上, 进一步考察转换以后各解释变量的多重共线性问题. 根据运行结果, 进行 BOX-COX 转换之后, Y 与 W_m 的相关系数似乎并没有降低多少, 依然高达 0.960 3. 而在对式(4)的两个线性参数表达式进行“拆分”之后, 反而还增加了两类严重多重共线性: $\ln K$ 与 BK 的相关系数为 0.995 7(几乎完全相关); $\ln Q$ 与 BQ 的相关系数为 0.987 2. 这表明 BOX-COX 转换与变量取自然对数的情况比较接近, 因此, 从多重共线性这一视角来看, 不能对 W_m 、 K 和 Q 进行 BOX-COX 转换, 所以, 转而对其进行幂阶梯转换.

4.2 幂阶梯转换

在转换过程中, 首先需要明确的是, 只要对式(3)中 W_m 、 K 和 Q 不采取自然对数形式, 便可以解

决系统结构模型的可识别问题. 在这一前提下, 对这三个变量进行幂阶梯转换以便处理多重共线性的问题, 需要延伸为同时考量减轻多重共线性问题与数据的正态分布, 而二者总是难以兼顾. 因此, 需要进一步厘清一个基本事实: 为了减轻严重多重共线性, 只能“稍微”牺牲一下数据的正态分布; 但也不能过分牺牲这一点, 否则数据的严重偏态分布可能会对计量结果产生较大影响. 经过实验, 最终得到了上文的式(3).

5 回归模型估计及分析

5.1 简化模型估计

经过对多重共线性和系统结构模型的可识别问题处理以后, 将式(3)代入方程(1)和(2), 通过计算整理, 便得到食品安全内生时的系统简化模型, 即为最终估计的模型. 利用 2004—2007 年肉制品面板数据, 采用面板似不相关(Seemingly Unrelated Regression Estimation, 简记 SURE)迭代回归和最小距离估计(Minimum Distance Estimation, 简记 MDE)进行估计.

为了对结果进行对比分析, 将食品安全外生(式(1)和(2)中不存在食品安全变量)时的估计结果作为基准(Bench Mark), 以便考察加入食品安全变量后, 估计结果会发生怎样的变化, 从而更好地了解肉制品企业食品安全生产的行为. 当然, 为何不能采用 OLS 估计系统简化模型, 这需要考察各观测值扰动项之间是否存在相关性, 为此, 列出了经过面板 SURE 迭代回归后的“残差相关系数矩阵”(Correlation Matrix of Residuals), 见表 3.

表 3 面板似不相关迭代回归后的残差相关系数矩阵(食品安全内生)

	$\ln(C/W_m)_{2004}$	$S_{I\ 2004}$	$\ln(C/W_m)_{2005}$	$S_{I\ 2005}$	$\ln(C/W_m)_{2006}$	$S_{I\ 2006}$	$\ln(C/W_m)_{2007}$	$S_{I\ 2007}$
$\ln(C/W_m)_{2004}$	1							
$S_{I\ 2004}$	-0.574 2	1						
$\ln(C/W_m)_{2005}$	0.317 1	-0.146 1	1					
$S_{I\ 2005}$	-0.170 0	0.244 8	-0.671 9	1				
$\ln(C/W_m)_{2006}$	0.268 3	-0.144 9	0.551 7	-0.517 1	1			
$S_{I\ 2006}$	-0.129 9	0.157 2	-0.440 6	0.722 6	-0.649 7	1		
$\ln(C/W_m)_{2007}$	0.110 9	-0.102 7	0.319 8	-0.209 3	0.364 9	-0.212 0	1	
$S_{I\ 2007}$	-0.007 1	0.118 7	-0.216 5	0.419 9	-0.256 7	0.517 3	-0.529 1	1

需要检验的原假设和备择假设分别为, H_0 : 不存在组间截面相关, H_1 : 存在组间截面相关. 一方面, 如果原假设成立, 意味着根据残差计算的个体扰动项之间的相关系数应接近于 0, 即在表 3 的残差相关系数矩阵中, 非主对角线的元素应离 0 不远. 而从上表可以看出, 非主对角线的元素不全为 0, 这意味着原假设并不成立. 另一方面, 从 Breusch-Pagan LM 统计量来看, 在食品安全外生和内生情况下, $p=0.000\ 0<0.01$, 也显著拒绝原假设, 这说明系统模型的扰动项确实存在组间截面相关, 表征使用 SURE 比单一方程的 OLS 更有效率(食品安全外生时的情况类似).

食品安全内生时的系统简化模型一共迭代了 15 次, 并最终收敛于 $9.591\text{e-}07$, 估计结果见表 4. 在表 4 中, 第(1)列至第(4)列分别代表 2004—2007 年式(1)存在食品安全变量的简化形式各年度的回归结果. 可以看出, 食品安全内生时系统简化模型的拟合优度比食品安全外生时系统简化模型的拟合优度(2004—2007 年分别为 0.228、0.204、0.157、0.141)有了较大改善, 表征前者对数据的拟合程度要优于后者, 这说明将食品安全变量纳入模型中是比较理想的选择.

表 4 系统简化模型面板似不相关迭代回归结果(食品安全内生)

变量	(1)	(2)	(3)	(4)
	$\ln(C/W_m)_{2004}$	$\ln(C/W_m)_{2005}$	$\ln(C/W_m)_{2006}$	$\ln(C/W_m)_{2007}$
Q	0.232 0 (0.983 0)	-0.258 0 (0.850 0)	1.110 0 (0.731 0)	0.736 0 (0.775 0)
lnZ	0.099 0 (0.265 0)	-0.059 2 (0.281 0)	-0.054 6 (0.310 0)	0.309 0 (0.360 0)
$\ln W_t$	0.531 0*** (0.112 0)	0.495 0*** (0.096 6)	0.325 0*** (0.091 4)	0.250 0*** (0.091 7)
$\sqrt[3]{W_m}$	-0.330 0*** (0.023 3)	-0.281 0*** (0.018 0)	-0.245 0*** (0.016 4)	-0.231 0*** (0.017 9)
$\sqrt{K}$	-0.008 4 (0.008 5)	-0.001 1 (0.008 1)	-0.010 9 (0.008 1)	0.004 3 (0.008 4)
lnY	-0.225 0 (0.419 0)	-0.641 0* (0.355 0)	-0.155 0 (0.394 0)	-0.699 0* (0.410 0)
$(\ln Y)^2$	0.056 8* (0.031 7)	0.078 1*** (0.030 2)	0.006 4 (0.032 9)	0.046 3 (0.032 9)
$\ln Y \ln K$	0.053 2 (0.033 6)	0.021 2 (0.038 0)	0.099 0** (0.041 9)	0.050 7 (0.041 3)
$\ln Y \ln(W_t/W_m)$	-0.022 8*** (0.004 3)	-0.036 9*** (0.004 5)	-0.032 5*** (0.006 9)	-0.035 7*** (0.005 5)
$Q \ln Y$	-0.089 6 (0.139 0)	0.147 0 (0.115 0)	0.165 0* (0.093 7)	0.054 4 (0.054 2)
$\ln Z \ln Y$	0.034 2 (0.042 7)	0.039 4 (0.040 0)	-0.048 9 (0.041 8)	-0.055 1 (0.047 1)
$\ln W_t \ln Y$	-0.100 0*** (0.016 9)	-0.078 5*** (0.015 1)	-0.058 8*** (0.013 7)	-0.051 1*** (0.013 5)
$\sqrt[3]{W_m} \ln Y$	0.017 4*** (0.002 7)	0.011 9*** (0.002 3)	0.015 7*** (0.002 5)	0.012 9*** (0.002 4)
$\sqrt{K} \ln Y$	-0.001 5*** (0.000 4)	-0.000 8** (0.000 3)	-0.000 8*** (0.000 2)	-0.000 7** (0.000 3)
$\ln Y \ln Q$	0.066 2*** (0.020 9)	0.004 6 (0.016 3)	0.051 7*** (0.015 1)	-0.042 7** (0.019 4)
$\ln(W_t/W_m)$	0.146 0*** (0.020 4)	0.207 0*** (0.028 9)	0.279 0*** (0.048 1)	0.270 0*** (0.032 6)

续表 4

变量	(1)	(2)	(3)	(4)
	$\ln(C/W_m)_{2004}$	$\ln(C/W_m)_{2005}$	$\ln(C/W_m)_{2006}$	$\ln(C/W_m)_{2007}$
$[\ln(W_l/W_m)]^2$	0.009 0*** (0.002 2)	0.006 7*** (0.002 4)	0.017 5*** (0.004 0)	0.010 4*** (0.002 9)
$Q\ln(W_l/W_m)$	0.006 9 (0.007 5)	0.019 6* (0.010 2)	0.041 8*** (0.012 9)	0.008 5 (0.005 6)
$\ln Z\ln(W_l/W_m)$	0.002 9 (0.004 3)	-0.002 7 (0.006 1)	-0.012 6 (0.009 9)	-0.013 4* (0.007 2)
$\ln W_l\ln(W_l/W_m)$	0.029 7*** (0.004 9)	0.039 8*** (0.005 3)	0.022 9*** (0.008 2)	0.029 3*** (0.006 1)
$\sqrt[3]{W_m}\ln(W_l/W_m)$	-0.000 1 (0.000 3)	0.000 3 (0.000 4)	0.001 0* (0.000 6)	0.000 4 (0.000 4)
$\sqrt{K}\ln(W_l/W_m)$	-2.94e-05 (1.96e-05)	-3.68e-05 (2.36e-05)	-6.09e-05* (3.63e-05)	-1.77e-05 (2.35e-05)
$\ln(W_l/W_m)\ln Q$	0.000 9 (0.001 0)	-0.002 8*** (0.001 3)	-0.006 6*** (0.001 9)	0.000 1 (0.001 4)
$\ln K$	-0.498 0 (0.387 0)	0.180 0 (0.396 0)	-0.723 0 (0.446 0)	0.116 0 (0.525 0)
$(\ln K)^2$	-0.021 8 (0.027 7)	-0.044 5* (0.027 1)	-0.010 7 (0.032 2)	-0.052 3 (0.033 3)
$\ln K\ln(W_l/W_m)$	-5.60e-05 (0.002 3)	-0.002 3 (0.003 0)	0.001 9 (0.004 5)	0.000 8 (0.003 1)
$Q\ln K$	0.039 7 (0.122 0)	-0.187 0 (0.120 0)	-0.281 0*** (0.078 7)	-0.168 0*** (0.053 3)
$\ln Z\ln K$	-0.039 2 (0.040 6)	-0.036 0 (0.036 6)	0.060 0 (0.038 1)	0.035 7 (0.043 2)
$\ln W_l\ln K$	0.062 6*** (0.015 4)	0.046 1*** (0.014 0)	0.033 8*** (0.012 0)	0.035 9*** (0.012 4)
$\sqrt[3]{W_m}\ln K$	0.003 8 (0.003 1)	0.005 8** (0.002 6)	-0.001 5 (0.002 3)	0.001 8 (0.002 2)
$\sqrt{K}\ln K$	0.001 8** (0.000 7)	0.000 7 (0.000 7)	0.001 3** (0.000 6)	0.000 3 (0.000 6)
$\ln K\ln Q$	-0.001 7 (0.018 1)	-0.006 7 (0.015 8)	0.014 6 (0.014 8)	0.017 0 (0.016 7)

续表 4

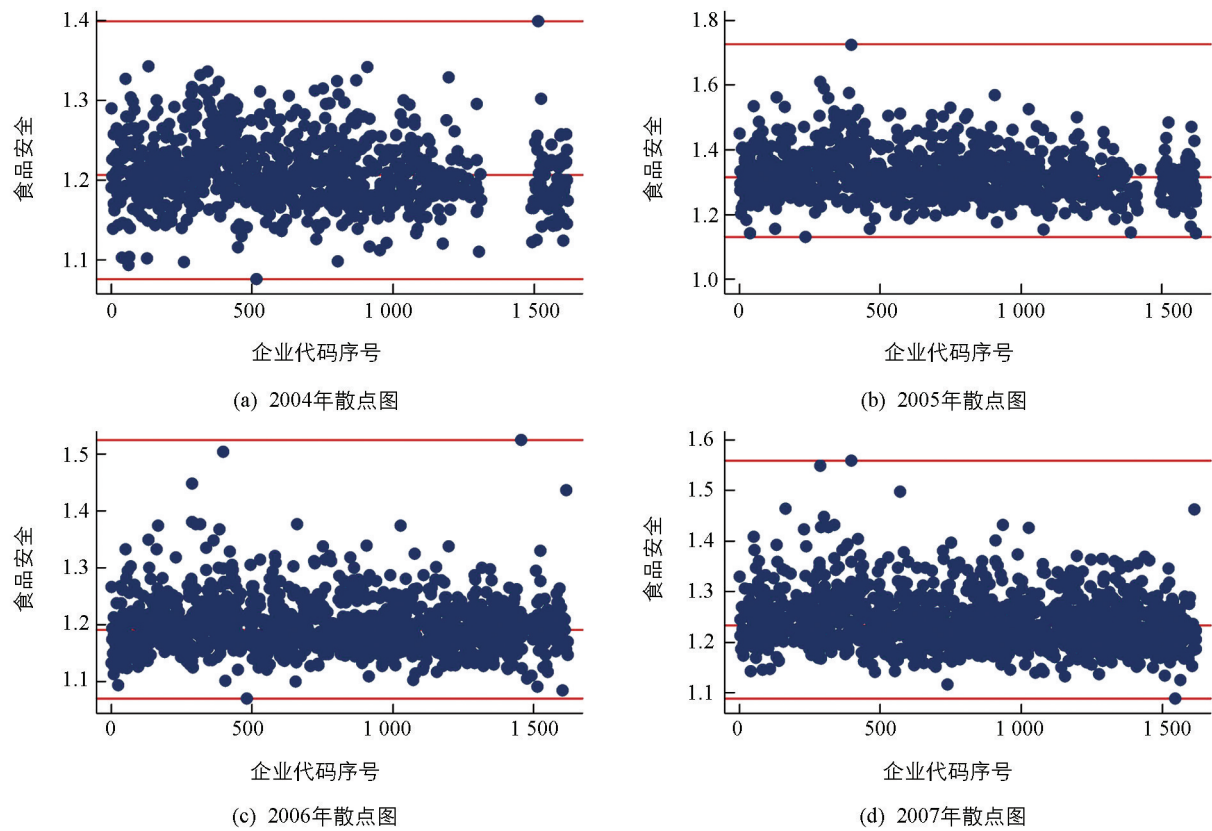
变量	(1)	(2)	(3)	(4)
	$\ln(C/W_m)_{2004}$	$\ln(C/W_m)_{2005}$	$\ln(C/W_m)_{2006}$	$\ln(C/W_m)_{2007}$
$\ln Q$	-0.514 0*** (0.164 0)	-0.022 5 (0.166 0)	-0.662 0*** (0.160 0)	0.217 0 (0.184 0)
$(\ln Q)^2$	0.010 8** (0.005 4)	0.002 4 (0.008 4)	-0.010 7* (0.005 9)	0.004 8 (0.005 2)
$Q \ln Q$	0.247 0 (0.208 0)	0.112 0 (0.302 0)	-0.539 0*** (0.202 0)	-0.001 3 (0.108 0)
$\ln Z \ln Q$	0.009 8 (0.019 4)	-0.011 3 (0.019 4)	0.011 8 (0.019 5)	0.023 9 (0.025 7)
$\ln W_t \ln Q$	0.012 4 (0.008 8)	0.014 5** (0.007 4)	-0.001 3 (0.006 0)	-0.000 6 (0.008 1)
$\sqrt[3]{W_m} \ln Q$	-0.003 4** (0.001 7)	0.000 2 (0.001 4)	-0.004 5*** (0.001 1)	0.004 7*** (0.001 4)
$\sqrt{K} \ln Q$	-0.000 2 (0.000 2)	-0.000 1 (0.000 2)	-6.93e-05 (0.000 1)	-0.000 2** (0.000 1)
常数	1.917 0 (2.436 0)	1.521 0 (2.237 0)	2.486 0 (2.503 0)	3.368 0 (2.746 0)
观测值	365	365	365	365
调整的拟合优度	0.676 0	0.697 0	0.598 0	0.664 0

注：(1) 括号内为参数估计值的标准误；(2) ***、**、* 分别表示参数在 1%、5%、10%水平上显著。

5.2 食品安全数据模拟

本文模拟出了 2004—2007 年的肉制品企业食品安全数据，让我们对食品安全有一个初步的感知，见图 2。从图 2 可以看出，2004—2007 年各企业的食品安全水平比较集中，而且大约集中在 1 与 1.8 之间。为了剔除异端值的影响，我们以中位数来代表各年度食品安全的平均水平，这也就反映了肉制品行业的食品安全水平：2004—2007 年中国肉制品行业的食品安全平均水平分别为 1.206 2、1.314 5、1.191 8 和 1.232 7。图 3 描绘出了 2004—2007 年度肉制品的食品安全水平的分布状况。

从图 3 可以看出，2004—2007 年肉制品企业的食品安全大致呈正态分布，这可以研判本文所模拟的食品安全数据具有合理性。从现有研究来看，我们对这一数据的认知还非常有限。这也不足为奇，正如劳动经济在研究工资时，对“能力”这一变量的控制，也是用“智商”(Intelligence Quotient，简称 IQ)数据来捕捉的；或者如同产品质量一样，也是一个没有单位的中性变量。因此，随着对食品安全的深入研究，将会揭示出更多关于食品安全的相关信息。但有一点是可以肯定的：本文所模拟的食品安全数据正是基于特定行业(肉制品)的数据结构(2004—2007 年面板数据)、并通过相应计量手段(面板 SURE 迭代估计)予以实现，这对当前关于食品安全指数的研究^[26]可提供一定借鉴。



各分图中, 最上面和最下面的水平线分别为最大值线和最小值线, 中间的水平线为中位数线.

图 2 2004—2007 年全国肉制品企业食品安全水平散点图

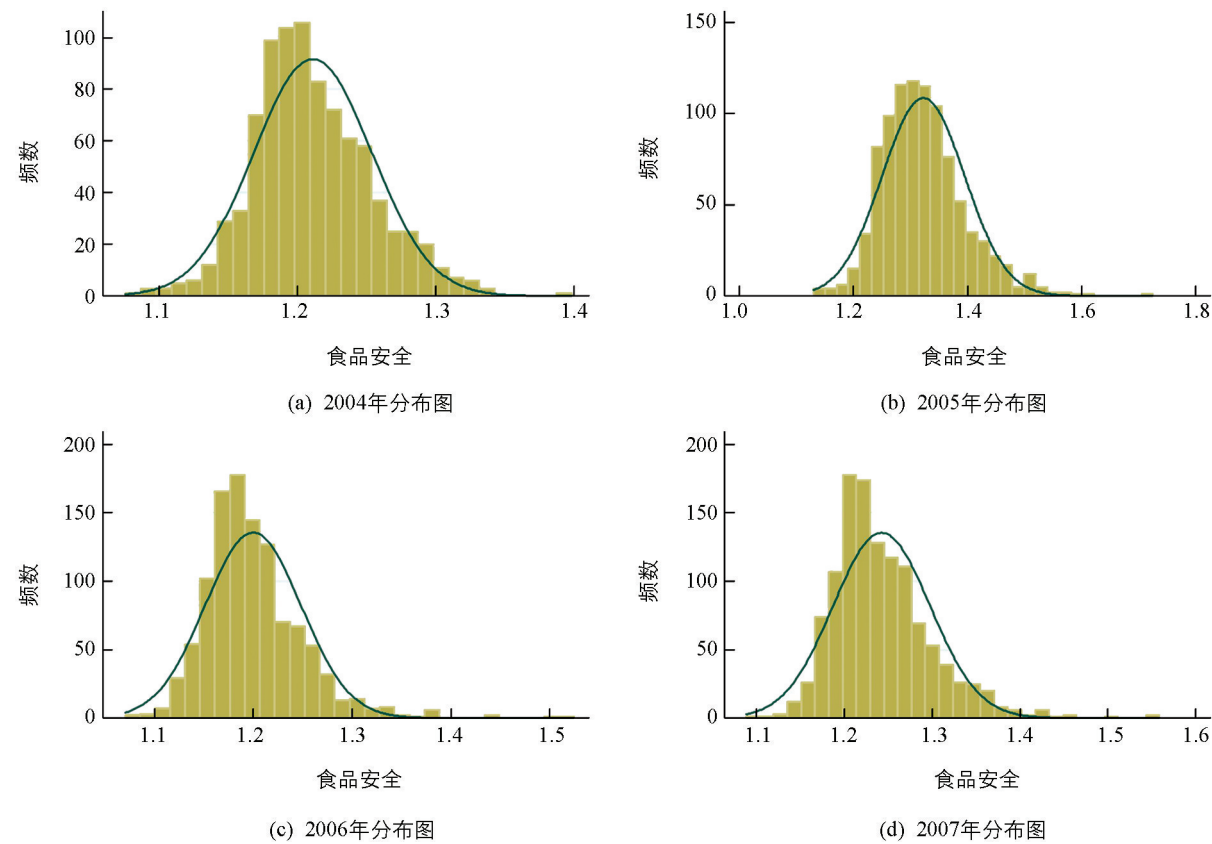


图 3 2004—2007 年全国肉制品企业食品安全水平分布图

6 结语

从上述实证分析本文得出:

第一,模型形式选择.通过模型形式的检验结果来看,拓展后的质量调整成本模型仍然为固定效应模型,属于面板模型实证研究的常见现象,表明对该模型的拓展仍然是科学、合理的.而 FE 模型常见的表达方式是采取离差形式,将无法观测到的扰动项消除,以解决估计的一致性问题;如果按照此方法,那么考虑无法观测的食品安全变量的实证研究便无法进行下去.可能目前的文献并未考虑到这一点,只是将精力集中在思考如何将食品安全变量表达出来.但本文将这一问题有效纳入了面板模型的分析框架之中,既拓展了固定效应模型的相关研究,也让方程的设定变得更有说服力,这应该是本文的一个贡献.

第二,严重多重共线性和模型可识别处理.从幂阶梯转换和 BOX-COX 转换视角,对超越对数成本面板方程进行了再认识.以解释变量数据正态分布的改善为前提进行 BOX-COX 转换,虽然该方法解决了模型识别问题,但却未缓解严重多重共线性问题.而幂阶梯转换的基本逻辑为随着变量形式的变换(变量幂指数的变化),该变量与其他变量之间的相关系数会发生相应的改变(多重共线性);同时,变化后变量的数据分布是否接近于正态分布,这需要在二者之间作出有效的权衡.最终以缓解严重多重共线性为前提(首先要解决模型的识别问题),通过数据接近正态分布时所对应的变量进行重复实验,达到合理解决问题的目的.

第三,模型的优良性和对食品安全水平数据的认识.从估计结果来看,食品安全内生时的模型比外生时的模型对样本数据拟合得更好,因此,在实证研究中如果不考虑食品安全变量,将会导致模型的有偏估计,结论也就失去可信性,最终提出的对食品安全问题监管的措施可能会出现偏差.而且通过模型成功模拟出了食品安全水平的数据,让这一无法观测到的变量具体化,并且具有良好的分布.

第四,现有文献关于超越对数成本函数的研究主要采用截面数据,而本文对其设定为面板模型进行分析,这为后续运用该函数展开面板数据分析奠定了基础.同时,在设定过程中考虑了无法观察到的变量,这将对如何处理不可观测变量模型设定的实证研究起到推动作用.本文对食品安全数据的模拟具有一定的应用价值,这将有助于开发测度企业食品安全水平的相关软件,一旦该软件对接企业的财务报表数据和国家居民收入数据等数据库,便可迅速掌握企业的食品安全水平,从而极大减轻政府相关部门对企业食品安全水平检测的压力.

参考文献:

- [1] KITENGE E. Effects of Food and Agricultural Imports on Domestic Factors in the U. S. Agricultural Sector: A Translog Cost Function Framework [J]. Applied Economics Letters, 2016, 23(2): 132-137.
- [2] SHIN Y, SIM S. Analysis on Economies of Scale in Offshore Fishery Using a Translog Cost Function [J]. Ocean and Polar Research, 2017, 39(1): 61-71.
- [3] BHATTACHARYA A. Technological Innovation in Tobacco Industry in the United States: A Translog Cost Function Framework [J]. American Journal of Economics, 2018, 8(3): 123-137.
- [4] MODREGO F, CANALES A, BAHAMONDE H. Employment Effects of COVID-19 across Chilean Regions: An Application of the Translog Cost Function [J]. Regional Science Policy & Practice, 2020, 12(6): 1151-1167.
- [5] BAGADEEM S. Estimating the Translog Cost Function for the Telecommunications Sector in Saudi Arabia: The Case of Saudi Telecom Company: An Empirical Study [J]. International Journal of Financial Research, 2021, 12(3): 19-30.
- [6] MAZIOTIS A, MOLINOS-SENANTE M. Economies of Integration and Sources of Change in Productivity in the Chilean Water and Sewerage Industry: A Translog Cost Function Approach [J]. Environmental Science and Pollution Research International, 2022, 29(6): 8503-8513.

- [7] 陈林, 胡超凡. 中国上市零售企业的规模经济效应研究——基于超越对数成本函数的实证分析 [J]. 财贸研究, 2015, 26(6): 64-73.
- [8] 章玉, 黄承锋, 许茂增. 特许经营制度下公共交通行业的财政补贴效应——基于超越对数成本函数模型的实证 [J]. 产经评论, 2016, 7(3): 118-129.
- [9] 查冬兰, 司建松, 周德群, 等. 中国工业部门能源与非能源替代弹性研究——基于多弹性测度方法 [J]. 管理评论, 2016, 28(6): 180-191.
- [10] 林善浪, 叶炜, 张丽华. 农村劳动力转移有利于农业机械化发展吗——基于改进的超越对数成本函数的分析 [J]. 农业技术经济, 2017(7): 4-17.
- [11] 陈林. 自然垄断与混合所有制改革——基于自然实验与成本函数的分析 [J]. 经济研究, 2018, 53(1): 81-96.
- [12] 徐成波, 李江一. 食品安全生产对成本影响的实证研究——基于工业企业数据的经验证据 [J]. 贵州财经大学学报, 2018(5): 64-76.
- [13] 徐成波. 一类不可观测变量模型的特征研究 [J]. 统计与决策, 2019, 35(5): 81-84.
- [14] BRAEUTIGAM RR, PAULY M. Cost Function Estimation and Quality Bias: The Regulated Automobile Insurance Industry [J]. The RAND Journal of Economics, 1986, 17(4): 606-617.
- [15] GERTLER P, WALDMAN D. Quality-Adjusted Cost Functions and Policy Evaluation in the Nursing Home Industry [J]. Journal of Political Economy, 1992, 100(6): 1232-1256.
- [16] 威廉·H. 格林. 计量经济分析: 上 [M]. 5 版. 费剑平, 译. 北京: 中国人民大学出版社, 2007: 272.
- [17] ANTLE J M. No such Thing as a Free Safe Lunch: The Cost of Food Safety Regulation in the Meat Industry [J]. American Journal of Agricultural Economics, 2000, 82(2): 310-322.
- [18] WOOLDRIDGE J. Econometric Analysis of Cross Section and Panel Data [M]. Cambridge, MA: MIT Press, 2001: 282-283.
- [19] BREUSCH T, PAGAN A. The Lagrange Multiplier Test and Its Application to Model Specification in Econometrics [J]. Review of Economic Studies, 1980, 47(1): 239-53.
- [20] HAUSMAN J. Specification Tests in Econometrics [J]. Econometrica, 1978, 46(6): 1251-1271.
- [21] 陈强. 高级计量经济学及 Stata 应用 [M]. 北京: 高等教育出版社, 2010: 149-150.
- [22] CAMERON A C, TRIVEDI P K. Microeconometrics Using Stata [M]. College Station, Texas: Stata Press, 2010.
- [23] 王志刚, 李腾飞, 韩剑龙. 食品安全规制对生产成本的影响——基于全国 334 家加工企业的实证分析 [J]. 农业技术经济, 2012(11): 57-68.
- [24] MOCAN H N. Quality-Adjusted Cost Functions for Child-Care Centers [J]. American Economic Review, 1995, 85(2): 409-413.
- [25] 劳伦斯·汉密尔顿. 应用 STATA 做统计分析 [M]. 郭志刚, 等译. 重庆: 重庆大学出版社, 2011: 110-111.
- [26] 李太平. 我国食品安全指数的编制理论与应用研究——以国家食品抽检数据为例 [J]. 农业经济问题, 2017, 38(7): 80-87, 111-112.

责任编辑 任剑乔