

恩格尔系数与居民收入^{*}

——扩展 Working-Leser 模型研究

陈梦根

【摘要】文章以中国 1980~2016 年的城镇和农村数据为样本,从时间维度探讨恩格尔定律在中国的适用性及恩格尔系数变动的影响因素。基于 Working-Leser 模型的研究结果表明,恩格尔系数与收入水平之间呈负相关关系。在此基础上,引入食品价格构建的扩展 Working-Leser 模型也支持了恩格尔定律在中国的适用性,表现为恩格尔系数随居民收入上升而下降,而且食品价格与恩格尔系数显著正相关,若纳入其他控制变量,收入水平和食品价格对恩格尔系数的影响有所弱化。进一步的研究表明,样本期内中国居民家庭恩格尔系数大多数年份下降,但城镇和农村恩格尔系数变动表现出不同的分布特征,二者受不同因素的驱动。GDP 增速、收入与消费水平对恩格尔系数变动具有反向作用,而食品价格、人口城镇化率、第一产业份额对恩格尔系数变动具有同向影响。

【关键词】恩格尔定律 居民收入 消费支出 食品价格

【作者】陈梦根 北京师范大学统计学学院,教授。

改革开放 40 年来,中国城乡居民恩格尔系数总体上呈明显下行趋势。1978 年中国城镇和农村居民恩格尔系数分别为 57.5% 和 67.7%,2016 年分别为 29.3% 和 32.2%,分别下降 28.2 和 35.5 个百分点。按照世界粮食与农业组织(FAO)的标准(恩格尔系数在 30% 以下为最富裕,30%~40% 为富裕,40%~50% 为小康,50%~59% 为温饱,59% 以上为贫困),中国城镇和农村居民已分别处于最富裕和富裕阶段。但也有学者指出,中国正处于经济高速增长期,居民消费结构发生着日新月异的变化,经济发展水平不断提高,各地居民家庭消费习惯、文化习俗差异巨大,直接将恩格尔系数应用于中国的实际情况时可能得出错误的结论,恩格尔定律在中国可能失效(王少飞,2002;尹海洁、唐雨,2009)。本文拟从时间维度考察中国城镇与农村恩格尔系数和收入水平之间的关系,探讨恩格尔定律在衡量中国居民生活质量和经济发展水平中的适用性,以及恩格尔系数变动的主要影响因素。

^{*} 本文为国家社会科学基金重点项目“购买力平价(PPP)汇总方法研究”(编号:19ATJ002)的阶段性成果。

一、文献述评

恩格尔定律被誉为经济学中“迄今最为可靠的实证性规律”,在内涵上是指收入用于食物支出的增长慢于收入的增长,而非随着收入增长食物开支下降(Lewbel, 2008)。恩格尔定律反映了收入增加时居民消费模式的结构变化,旨在测度居民福利,揭示消费结构及其运动规律(Chai 等, 2010)。自恩格尔之后,许多学者对恩格尔定律进行了理论与实证研究。霍利斯·钱纳里、莫伊思·赛尔昆(1988)对 100 多个国家的资料统计发现,随着人均国民生产总值(GNP)的提高,恩格尔系数下降,二者高度相关。此后,恩格尔系数成为衡量一国或地区富裕程度的重要指标,恩格尔定律简单地运用食品支出占生活消费支出的份额来衡量不同居民家庭的生活和福祉水平,受到广泛关注(Lancaster 等, 1999)。长期以来,学者们从多角度、全方位探讨了恩格尔定律的适用性,Gibson(2002)分析了家庭人口数量、年龄及性别构成等对恩格尔系数的影响。Kaus(2013)通过对不同国家和不同时期居民消费模式的研究认为,恩格尔定律代表了人类社会的共性。

从方法上看,早期文献主要采用线性框架描述食物支出与收入或消费支出之间的关系。此后,恩格尔曲线常被表示为商品消费份额关于总消费支出对数的函数关系(Lewbel, 2008)。其中,最经典的框架是 Working-Leser 模型,由 Working(1943)和 Leser(1963)发展,将食品支出与收入的自然对数线性关联,用以检验恩格尔系数与收入水平之间的关系。Muellbauer(1976)将 Working-Leser 恩格尔曲线置于可积消费理论的框架下进行讨论,使该模型具备了更为可靠的理论支撑。Deaton 等(1980)的近似理想需求系统(AIDS)理论可以用来导出 Working-Leser 模型的微观基础,作为 AIDS 在价格给定时的一个特例,该模型可视为真正消费需求函数的一种一阶逼近。Working-Leser 模型在居民消费支出研究中应用广泛,其拟合效果优于其他许多模型。也有学者对 Working-Leser 模型提出质疑,认为该模型在一些新型经济中对数据的拟合没有早期研究那么好,并由此提出了一些新的研究框架,如 Jorgenon 等(1982)的 Translog 模型和 Banks 等(1997)的二次近似理想需求系统(QUAIDS)等。

在中国,恩格尔系数得到广泛应用。甘健胜、黄泽民(2006)认为,恩格尔系数反映了一国经济结构变化及经济增长快慢,可以作为生活水平高低的评价指标,连续测算的恩格尔系数序列可以直接或间接反映主导消费品在消费结构中的更替规律及消费品组合档次变化的情况。尹海洁、唐雨(2009)利用对哈尔滨市和沈阳市贫困人口的调查数据,将恩格尔系数应用于中国城市贫困测量。除了对居民消费结构与生活水平的测度外,一些学者还将其应用于其他相关领域,如许永洪、曾五一(2009)运用恩格尔曲线测度中国居民消费价格指数(CPI)的偏差。

20 世纪 80 年代以来,陆续有学者对恩格尔定律在中国的适用性提出质疑,认为恩

格尔系数在描述中国居民消费结构变动特征时存在一定的局限性,恩格尔定律存在失效的问题(王少飞,2002;张荣山,2007)。由此,不少学者提出应对恩格尔系数进行修正,以提高其在居民生活水平测度时的精度与效度,例如,谢健(1993)指出,若以时间维度为出发点,应采用波动平滑法消除恩格尔系数中不同时期、不同阶段及周期波动的影响。马崇明(1994)认为,应通过生活费价格指数调整的方法,消除价格因素对恩格尔系数的影响。也有学者认为恩格尔定律是适用的,例如,王芳(2003)对中国城乡恩格尔系数变动特征进行研究,探究影响恩格尔系数不规则变动的因素,认为恩格尔定律在中国基本适用。

综上所述,恩格尔系数已成为评价一国或地区生活水平的重要指标之一,恩格尔定律在中国的适用性是一个重要的命题,目前学术界尚未得到一致结论,仍有待进一步深入研究。值得特别关注的是,在国内外现有文献中,针对恩格尔定律适用性的研究没有考虑食品价格的影响。理论上,恩格尔曲线描述消费者在某些食品与服务上的支出占总资源的比例随收入的变化情况,一般假定食品价格不变(Lewbel,2008)。

鉴于此,本文将依据经典的 Working-Leser 模型进行实证研究,并引入食品价格因素构建扩展 Working-Leser 模型,考察食品价格对恩格尔系数与收入关系的影响。

二、研究方案

学术界将居民家庭食物支出占消费支出的比重称为恩格尔系数。计算公式为:恩格尔系数 = $\frac{\text{居民家庭食物支出}}{\text{居民家庭消费支出}} \times 100\%$ 。在经济统计中,恩格尔系数已成为国际上通用的衡量居民生活水平的一项重要指标,学者们据此检验恩格尔定律的适用性,对恩格尔系数与收入之间的关系展开实证研究。从已有文献看,早期相关研究主要以线性函数居多,认为恩格尔曲线表现为一个线性消费系统,将食品支出与收入线性关联。模型基本形式为:

$$E=c+\beta y+\varepsilon \quad (1)$$

式(1)中, E 为恩格尔系数(即食品支出占消费总支出的份额), c 为常数项, y 为居民收入, ε 为随机误差项。该模型与恩格尔最初研究的基本形式一致,在其他研究中也由学者以食物消费量代替食物支出份额作为因变量,模型表示为:

$$q=c+\beta y+\varepsilon \quad (2)$$

Ogburn(1919)与 Allen 等(1935)以式(2)为基准模型对多国样本进行实证检验,但发现该模型有时估计误差较大。Working(1943)将食品支出份额与 $\ln(y)$ 线性关联,提出了以下模型用于检验恩格尔系数与收入水平之间的关系:

$$E=c+\beta \ln(y)+\varepsilon \quad (3)$$

式(3)中, E 为恩格尔系数,收入变量采用 $\ln(y)$ 的形式。Leser(1963)论证了该模型

的拟合效果优于其他许多模型,后来这一方法被称为 Working-Leser 模型。传统意义上恩格尔曲线反映商品价格给定时消费者收入 y 与其对商品的需求量 q 或消费支出份额之间的关系,于此不同,Working-Leser 模型将恩格尔曲线表示成商品消费支出份额关于居民收入对数(或总消费支出对数)的函数关系。该模型具有良好的微观基础,是近似理想需求系统(AIDS)在价格给定情形下的特例,在经验研究中被广泛采纳(Deaton 等,1980)。Hausman 等(1995)、Banks 等(1997)研究发现,多数消费品的恩格尔曲线接近 Working-Leser 模型。

Working-Leser 模型显示,恩格尔系数与居民收入水平的对数呈线性关系。理论上,恩格尔曲线描述消费者在食品上的支出占总资源的比例随收入的变化情况,一般假定食品价格不变。现实中,不同时期食品价格水平可能存在巨大差异,对居民家庭消费支出具有重要影响(Lewbel,2008)。为此,本文探讨恩格尔定律在中国的适用性时纳入食品价格,构建扩展 Working-Leser 模型用于实证研究,基本形式为:

$$E=c+\beta\ln(y)+\gamma p+\varphi Controls+\varepsilon \quad (4)$$

式(4)中, E 代表恩格尔系数, y 代表居民人均收入, p 代表食品价格水平, $Controls$ 代表其他控制变量向量, ε 为随机误差项。其中,根据一般研究惯例,控制变量的选取主要涉及经济增长、货币供应、人口结构、交通条件、产业结构、农业发展水平等方面因素。根据上述模型,本文从时间维度探讨恩格尔定律在中国的适用性,研究样本为中国改革开放后的城镇和农村居民时间序列数据。由于 1978 和 1979 年部分变量数据缺失,故本研究样本期设定为 1980~2016 年,包括全国城镇与农村样本。有关变量具体说明如下。

1. 恩格尔系数

国家统计局曾公布分城镇与农村恩格尔系数,数据截至 2012 年,此后不再发布,因此,本文中 1980~2012 年中国城镇与农村恩格尔系数来自国家统计局数据库,2013~2016 年中国城镇与农村恩格尔系数分别根据城镇和农村居民家庭个人平均消费支出数据计算得到。居民人均消费支出包括历年城镇和农村居民人均消费支出及其分类支出,数据取自历年《中国统计年鉴》。中国城镇与农村恩格尔系数计算公式为:恩格尔系

$$\text{数}(E)=\frac{\text{居民人均食品烟酒支出}}{\text{居民人均消费支出}}\times 100\%。$$

2. 居民人均收入

本文中居民人均收入指标分别为城镇居民人均可支配收入和农村居民人均纯收入,数据来自历年《中国统计年鉴》。

3. 食品价格水平

由于食品种类繁多,现行统计中没有公开发布的专门测度食品价格水平的指标,以往文献很少考察食品价格因素对恩格尔系数的影响。本研究以居民消费价格指数中食

品分类指数即食品 CPI 作为食品价格指数。通常 CPI 反映价格水平的变动状况,但对消费者的感受来说,价格涨跌对实际消费行为的影响可能更为强烈,特别是对于食品,价格变化比价格水平对消费行为的刺激可能更直接,因此以食品 CPI 作为食品价格代理变量是可行的。根据国家统计局公布的数据,中国自 1998 和 1997 年起分别公布了城镇和农村食品类 CPI,其他年份数据均以全国食品 CPI 分类指数代替,但由于 1980~1983 年缺少全国食品 CPI 数据,本研究中以全国零售食品分类物价指数代替。数据分别来自历年《中国统计年鉴》和《新中国六十年统计资料汇编》。

4. 控制变量

为了更准确地评估恩格尔定律在中国的适用性,全面考察恩格尔系数的影响因素,根据以往文献的惯例及有关经济理论,本研究引入一系列统计指标作为控制变量:(1)经济增速(GDPG),以历年国内生产总值(GDP)增长率代表,作为经济发展的最重要指标,能够综合反映某一国家或地区的经济发展水平与速度。(2)货币供应量(M_0),在现代货币经济条件下,货币供应量或社会融资规模决定某一地区的总体流动性水平,进而影响地区物价水平及社会总需求,根据数据的可得性,以 M_0 为代表指标。(3)人口结构指标(People),用城镇和农村常住人口比代表。(4)人均农业总产值(Agri),作为反映农业生产总体状况的指标。(5)第二产业增加值份额(Indust),经济结构与经济发展水平密切相关,同样也会影响居民消费行为与消费结构。(6)基尼系数(Gini),收入差距是一个重要的社会经济问题,可能对居民家庭消费态度与行为产生显著影响,故以反映居民收入分配不平等程度的基尼系数作为控制变量之一。(7)人均公路里程(Road),交通条件决定商品运输成本,也会影响市场发育程度,进而影响食品价格及消费成本。(8)农村居民家庭经营耕地面积(Land),土地是农业之本,农业用地情况可能对农产品供应及价格水平具有显著影响,因此引入农村居民家庭经营耕地面积作为控制变量。以上数据分别来自历年《中国统计年鉴》。

三、实证结果

(一) 初步分析

1980~2016 年中国居民人均收入呈明显的逐年递增趋势(见图 1),城镇居民人均可支配收入由 1980 年的 477.6 元增至 2016 年的 33 616.25 元,农村居民人均纯收入由 1980 年的 191.3 元增至 2016 年的 12 363.41 元。与此同时,城镇和农村恩格尔系数大体呈波动式下降,分别由 1980 年的 56.9%、61.8%降至 2016 年的 29.3%、32.2%,期间先后出现过多次反复,主要是 1982~1983 年、1989~1990 年和 2004~2010 年。直观来看,中国城乡恩格尔系数与居民收入之间的关系与恩格尔定律基本一致。相关分析表明,中国城镇居民恩格尔系数与农村居民恩格尔系数之间的相关系数为 0.95,二者走势基本一

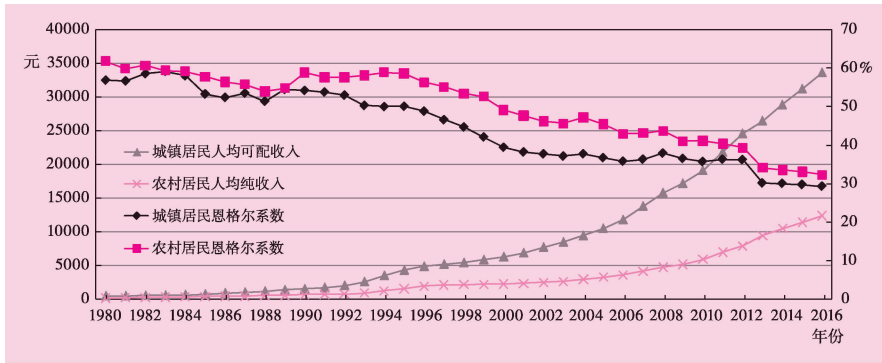


图1 1980~2016年恩格尔系数与居民收入变化趋势

致。城镇恩格尔系数与城镇居民人均可支配收入之间相关系数为-0.88,农村恩格尔系数与农村居民人均纯收入之间相关系数为-0.85,均表现为

显著的负相关关系。

Working-Leser 模型估计,结果如表1所示。对于城镇居民,系数 β 的估计值为-0.0672,在1%的水平上显著;而对于农村居民,系数 β 的估计值为-0.0679,也在1%的水平上显著。可见,恩格尔系数与居民收入之间呈较强的负相关,随着收入水平的上升,居民家庭食物支出占总支出的份额趋于下降。从 R^2 、对数似然比、AIC、SC和F统计量等评判指标来看,Working-Leser模型用于描述中国居民消费支出行为时效果较好,且对城镇恩格尔系数与收入关系的拟合优于农村。

(二) 扩展模型研究

鉴于变量间可能存在多重共线性,本研究采用逐步回归方法依次纳入不同控制变量,扩展Working-Leser模型估计结果如表2所示。

根据扩展Working-Leser模型估计结果,得到以下主要结论:(1)分别以城镇和农村食品CPI作为食品价格的代表指标,先不考虑其他控制变量估计得到扩展Working-Leser模型1和模型5结果显示,城镇和农村恩格尔系数的 β 估计值分别为-0.0662和-0.0675,均在1%水平上显著,与表1Working-Leser模型估计结果一致,恩格尔系数与收入水平呈

负相关关系。(2)在模型1和模型5中,食品价格的系数估计值分别为0.1447和0.0886,均在5%水平上显著,表明食品价格水平越高则恩格尔系数越大。也就是说,食品价格越高,居民家庭食品消费支出在总支出中份额越高,二者呈正相关关系。(3)若考虑其他控制变量的影响,分别估计得到模型2至模型4和模型6至模型8的结果表明,恩格尔系数与收入变量呈负

表1 Working-Leser模型估计结果

变 量	恩格尔系数	
	城镇	农村
C	1.0101***(37.15)	1.0054***(37.15)
ln(y)	-0.0672***(-21.02)	-0.0679***(-14.02)
R ²	0.9266	0.8489
调整 R ²	0.9245	0.8446
对数似然比	83.42	72.21
AIC	-4.4013	-3.7954
SC	-4.3142	-3.7083
F 统计量	442.18	196.76

注:括号内数字为t值;***表示在1%水平上显著。

表 2 恩格尔系数与居民收入:扩展 Working-Leser 模型

变 量	城镇恩格尔系数				农村恩格尔系数			
	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6	模型 7	模型 8
C	0.8474*** (11.45)	0.8227*** (6.88)	-0.1755 (-0.79)	-0.2182 (-0.61)	0.9090*** (8.76)	-0.3484** (-2.28)	-0.4325*** (-2.77)	-0.6063 (-1.60)
ln(y)	-0.0662*** (-21.78)	-0.0574*** (-3.41)	-0.0227 (-1.21)	-0.0380 (-0.97)	-0.0675*** (-13.91)	-0.0727*** (-4.78)	-0.0841*** (-4.93)	-0.0976** (-2.26)
p	0.1447** (2.34)	0.1250** (2.08)	0.0607 (1.13)	0.0780 (1.27)	0.0886** (1.99)	0.0907** (2.30)	0.1232* (1.86)	0.1099** (1.97)
GDPG	—	—	-0.1597 (-1.13)	-0.1249 (-0.76)	—	—	—	0.0309 (0.23)
M ₀	—	-0.1565* (-3.50)	-0.0333** (-2.22)	-0.0591 (-0.87)	—	-0.0780*** (-11.73)	-0.0860*** (-9.89)	-0.0474 (-0.72)
People	—	—	0.1224*** (4.77)	0.1264*** (3.03)	—	0.0991*** (6.52)	0.0895*** (5.70)	0.1089*** (3.07)
Agri	—	0.1474*** (3.13)	—	0.0110 (0.16)	—	—	—	-0.0392 (-0.51)
Indust	—	—	—	-0.0018 (-0.58)	—	—	0.0016 (1.11)	0.0020 (0.89)
Gini	—	-0.0697 (-0.41)	—	-0.0888 (-0.56)	—	—	-0.2159* (-1.83)	-0.2026 (-1.55)
Road	—	—	21.7804 (1.59)	28.2313 (1.69)	—	—	—	0.3276 (0.02)
Land	—	—	0.0414*** (4.13)	0.0364** (2.68)	—	—	—	0.0087 (0.81)
调整 R ²	0.9331	0.9489	0.9616	0.9583	0.8446	0.9699	0.9714	0.9687
F 统计量	252.16	134.96	129.90	83.90	98.84	291.96	204.80	112.65

注:括号内数字为 t 值;*、**、***分别表示在 10%、5%、1%水平上显著。

相关,但 ln(y)系数估计值的显著性水平总体上有所下降,尤其是针对城镇恩格尔系数更为明显。(4)食品价格与恩格尔系数之间仍为正相关关系,但与收入类似,食品价格的系数估计值显著性水平有所下降,尤其是以城镇恩格尔系数明显,当考虑其他因素之后食品价格对恩格尔系数的影响同样有所弱化。(5)从控制变量的情况看,各模型中不同变量系数估计值的显著性存在差异,总体上,货币供应量、人口结构对城镇与农村恩格尔系数都存在一定影响,人均农业总产值和农村居民家庭人均经营耕地面积对城镇居民恩格尔系数存在影响,基尼系数对农村居民恩格尔系数存在影响,但未发现 GDP 增速、产业结构指标和交通条件指标影响恩格尔系数的证据。可见,除居民收入和价格水平外,货币和人口因素是影响中国居民家庭消费支出结构的重要因素,而且城镇和农村地区恩格尔系数可能受不同因素影响。

从拟合效果来看,扩展 Working-Leser 模型 1、模型 5 的 R^2 分别为 0.9368、0.8532,对数似然比分别为 86.19、72.74,结合其他评估指标来看,优于 Working-Leser 模型的估计结果。可见,扩展 Working-Leser 模型纳入食品价格指数对恩格尔系数与收入关系的刻画好于传统模型。当纳入其他控制变量后,虽然 R^2 有所增大,但综合其他评判指标及系数估计值的显著性水平来看,模型拟合效果并没有明显改进,因此,仅纳入食品价格因素的扩展 Working-Leser 模型拟合效果相对更优。

(三) 稳健性检验

基于稳健性的考虑,下面本文分别以替代模型、替换关键变量和地区样本数据进行分析。

1. 替代模型分析

早期文献中有关恩格尔系数与收入关系的研究主要采用式(1)即简单线性系统进行研究,下面将其作为对照模型之一。此外,Leser(1963)还提出另一种实证研究框架,建立以下模型:

$$E = c + \beta \ln(y) + \delta y^{-1} + \varepsilon \quad (5)$$

式(6)中, E 为恩格尔系数, c 为常数项, y 为居民收入, ε 为随机误差项。Leser(1963)认为,这一模型拟合效果优于简单线性框架,甚至也优于 Working-Leser 模型。这一思路通常被称为 Rank-three 模型,因此本研究也将其作为对照模型。从结果看,简单线性模型与 Rank-three 模型估计结果显示,恩格尔系数与居民收入呈负相关关系,与 Working-Leser 模型结论一致。模型中 y 、 $\ln(y)$ 的系数估计值均为负,且在 1%水平上显著。城镇 Rank-three 模型和扩展 Rank-three 模型中, $1/y$ 的系数估计值为负,分别在 5%和 10%水平上显著,而农村 Rank-three 模型和扩展 Rank-three 模型中, $1/y$ 的系数估计值也为负,均在 1%水平上显著。另外,在扩展简单线性模型和扩展 Rank-three 模型中,食品价格变量在统计上不显著,也就是说,替代模型中食品价格因素对恩格尔系数的影响不太明显。从模型拟合效果,综合模型评判指标及系数估计值显著性水平来看,无论是城镇还是农村恩格尔系数,简单线性模型拟合效果均不如 Rank-three 模型,而 Rank-three 模型与 Working-Leser 模型接近。

2. 替换变量分析

以城镇 CPI 和农村 CPI 替换城镇食品 CPI 和农村食品 CPI 作为食品价格的代理变量,重新估计发现,无论是对于城镇还是农村,模型中 y 和 $\ln(y)$ 的系数估计值各自都在 1%的水平上显著为负,而农村扩展 Rank-three 模型中 $1/y$ 的系数估计值在 1%的水平上显著为负,城镇扩展 Rank-three 模型中 $1/y$ 的系数统计上不显著。食品价格变量的系数估计值均为正,但仅城镇和农村扩展 Working-Leser 模型及城镇扩展 Rank-three 模型估计结果在 5%或 10%水平上显著。

以人均 GDP 代替城镇和农村居民收入、CPI 代替城镇和农村食品 CPI,重新估计发

现,与收入指标一样,各模型中人均 GDP、 $\ln(\text{人均 GDP})$ 的系数均在 1%水平上显著为负,扩展 Rank-three 模型中人均 GDP 的系数估计值也为负,但仅在城镇恩格尔系数模型中显著。可见,以人均 GDP 作为收入指标,估计结果仍然支持恩格尔定律的判断。此外,各模型中 CPI 的系数均为正,但城镇扩展简单线性模型和农村扩展 Rank-three 模型结果不显著。可见,替换关键变量的估计结果与前述 Working-Leser 和扩展 Working-Leser 模型的结论基本一致。

3. 地区样本研究

从地区层面出发,本文选取了全国 31 个省份的样本建立面板模型进行稳健性研究。受数据可得性限制,仅以恩格尔系数、收入与价格为变量,样本期为 2002~2016 年。其中,恩格尔系数根据各地区城乡居民家庭人均消费支出计算,收入指标为城镇居民人均可支配收入与农村居民人均纯收入(2013~2016 年为人均可支配收入),价格指标以食品 CPI 代表,数据均来自历年《中国统计年鉴》。从全国来看,无论是城镇还是农村地区,收入指标的系数均显著为负(见表 3)。同时,价格指标与恩格尔系数之间表现为显著正相关。为了考察这种关系是否存在地区差异,本文将全国 31 个省份划分为东部、中部和西部三大地区分别进行研究,结果显示,三大地区城镇与农村恩格尔系数均和收入对数指标呈显著负相关关系。东部城镇地区模型中价格指标系数在 10%水平上显著,中部城镇地区模型中价格指标系数不显著,其余均在 1%或 5%水平上显著。总体上看,城镇地区恩格尔系数受价格水平因素的影响较弱,而农村地区恩格尔系数受价格水平因素的影响较为显著。

四、进一步分析

下面本文进一步探讨不同年份恩格尔系数的变动特征及主要影响因素,首先计算

表 3 地区样本研究结果

变 量	全国		东部		中部		西部	
	城镇	农村	城镇	农村	城镇	农村	城镇	农村
α	0.9106*** (8.43)	0.8488*** (4.39)	0.8782*** (3.77)	1.6158*** (3.14)	0.7566*** (3.74)	0.8401*** (5.15)	0.5318*** (3.14)	0.9740*** (7.06)
$\ln(y)$	-0.0762*** (-12.08)	-0.0927*** (-7.47)	-0.0763*** (-3.42)	-0.1925*** (-4.25)	-0.0538*** (-5.35)	-0.0980*** (-15.80)	-0.0488*** (-4.83)	-0.0957*** (-17.05)
p	0.0017** (2.35)	0.0035** (2.56)	0.0022* (1.85)	0.0048** (2.26)	0.0010 (0.58)	0.0039*** (2.63)	0.0029** (2.09)	0.0024** (1.99)
调整 R^2	0.8674	0.8483	0.9000	0.9124	0.8735	0.8381	0.8430	0.8631
D-W	0.6708	0.7833	0.8187	0.8087	0.7494	0.6600	0.6822	1.0869
F 统计量	95.88	82.11	124.11	66.76	92.31	92.31	74.94	87.85

注:括号内数字为 t 值;*、**、***分别表示在 10%、5%、1%水平上显著;控制地区效应和时间效应。

表4 E-change 的区间频数分布

区 间	均值	标准差	观测值
城镇			
[-0.06, -0.04)	-0.0470	—	1
[-0.04, -0.02)	-0.0246	0.0029	5
[-0.02, 0)	-0.0075	0.0049	21
[0, 0.02)	0.0088	0.0061	8
[0.02, 0.04)	0.0310	—	1
全样本	-0.0063	0.0145	36
农村			
[-0.04, -0.02)	-0.0275	0.0053	4
[-0.02, 0)	-0.0110	0.0051	22
[0, 0.02)	0.0059	0.0050	9
[0.04, 0.06)	0.0400	—	1
全样本	-0.0072	0.0137	36

恩格尔系数变动指标: $E-change = 恩格尔系数_t - 恩格尔系数_{t-1}$ 。其中 $E-change$ 代表恩格尔系数年度变化量, t 为年度。统计分析表明, 城镇恩格尔系数变动指标在样本期内的均值为 -0.0063 , 中位数为 -0.0055 , 最大值为 0.0310 , 最小值为 -0.0470 , 呈明显的尖峰、左偏特征; 农村恩格尔系数变动指标在样本期内的均值为 -0.0072 , 中位数为 -0.0080 , 最大值为 0.0400 , 最小值为 -0.0350 , 呈明显的尖峰、右偏特征(见表4)。

从图2可以看出, 城镇和农村恩格尔系数变动指标的不同分布特征。实际上, 中国经济长期以来存在明显的二元结构特征, 城镇和农村地区经济发展分化严重, 城镇和农村居民的消费行为及支出结构存在差异。受不同因素的影响, 城镇与农村恩格尔系数变动特征有所不同。根据经济学理论及相关研究(Kaus, 2013; 甘健胜、黄泽民, 2006), 本文选取 GDP 增长率(GDPG)、居民人均收入(y)、食品 CPI(p)、居民人均消费支出(Expend)、货币供应量(M_0)、城镇化率(Urban)、人均农牧渔业总产值(Agri)、第一产业增加

价值份额(Share)、基尼系数(Gini)、人均公路里程(Road)、农村居民家庭经营耕地面积(Land)为解释变量, 以 $E-change$ 为因变量, 探讨恩格尔系数变动主要受哪些因素影响。具体模型为:

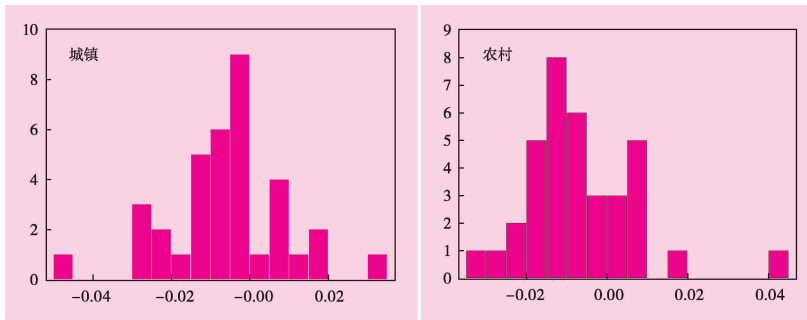


图2 恩格尔系数变动指标(E-change)的分布特征

$$E-change = c + \beta_1 GDPG + \beta_2 y + \beta_3 p + \beta_4 Expend + \beta_5 M_0 + \beta_6 Urban + \beta_7 Agri + \beta_8 Share + \beta_9 Gini + \beta_{10} Road + \beta_{11} Land + \varepsilon \quad (6)$$

如表5所示, 针对城镇恩格尔系数, 经济增速(GDPG)的系数估计值显著为负, 意味着在中国经济增速较快的早期恩格尔系数下降较慢, 而经济增速放缓的近期恩格尔系数下降较快, 主要原因是近十几年来中国经济发展水平明显跃升, 居民家庭富裕程度提

表 5 恩格尔系数变动影响因素分析

变 量	城 镇				农 村			
	模型 9	模型 10	模型 11	模型 12	模型 13	模型 14	模型 15	模型 16
c	-0.3119*** (-3.39)	-0.3593*** (-3.17)	-0.3000*** (-3.39)	-0.3559*** (-2.85)	-0.2357*** (-3.45)	-0.2727** (-2.73)	-0.2956*** (-3.25)	-0.2734** (-2.58)
GDPG	-0.2322** (-2.60)	-0.1968** (-2.12)	-0.2319** (-2.59)	-0.1996* (-2.03)	—	-0.0794 (-0.92)	-0.0697 (-0.81)	-0.0735 (-0.77)
y	-2.76E-06** (-2.27)	-3.07E-06* (-1.88)	—	—	—	-3.75E-06 (-1.58)	—	—
p	0.1112*** (3.05)	0.1088** (2.70)	0.1084*** (3.00)	0.1036** (2.40)	0.1078*** (3.36)	0.1303*** (3.72)	0.1189*** (3.62)	0.1244*** (3.05)
Expend	—	—	-4.17E-06** (-2.24)	-4.98E-06* (-1.92)	—	—	-3.56E-06 (-1.36)	-5.22E-06 (-1.39)
M ₀	—	-1.20E-06 (-0.56)	—	-1.52E-06 (-0.66)	—	—	—	-1.20E-06 (-0.54)
Urban	0.4682*** (2.75)	0.5534** (2.18)	0.4619** (2.72)	0.5699* (1.99)	0.2808** (2.24)	0.4534** (2.35)	0.4169** (2.27)	0.4883** (2.16)
Agri	—	9.93E-06 (0.55)	—	1.25E-05 (0.65)	—	—	—	9.99E-06 (0.53)
Share	0.3667*** (2.91)	0.4494** (2.58)	0.3437*** (2.89)	0.4402** (2.36)	0.2338** (2.33)	0.2578 (1.58)	0.3220** (2.37)	0.2660 (1.50)
Gini	—	0.0526 (0.56)	—	0.0560 (0.60)	—	-0.0755 (-0.92)	—	-0.0896 (-0.97)
Road	—	-0.0092 (-1.50)	—	1.2377 (0.11)	-17.51* (-1.96)	-19.4135* (-2.01)	-18.3133* (-1.91)	-18.24 (-1.66)
Land	—	—	—	-0.0095 (-1.47)	—	—	—	-0.0019 (-0.30)
调整 R ²	0.2580	0.2363	0.2551	0.2142	0.2249	0.2317	0.2303	0.1430
F 统计量	3.43	2.20	3.39	1.95	3.53	2.50	2.74	1.58

注：同表 2。

高,导致在食品消费上的支出份额明显下降。由于居民人均收入与居民人均消费支出相关系数高达 0.99,模型 9 和模型 10 中包含居民人均收入,模型 11 和模型 12 中包含居民人均消费支出,二者系数均为负,且在 5%或 10%水平上显著,表明随着收入或消费水平上升,恩格尔系数降幅逐渐缩小,符合一般预期。食品价格指数的系数为正,分别在 1%或 5%水平上显著,即食品价格变动对恩格尔系数变动有同向影响。城镇化率和第一产业增加值份额的系数估计值也在统计上显著为正。可见,人口结构变化和产业结构变化对城镇恩格尔系数变动具有显著的同向影响。此外,货币供应量的系数为负,其他变量的系数在统计上不显著。

农村恩格尔系数的估计结果存在一定的差异。(1)经济增速、收入和消费水平对农村恩格尔系数变动的的影响在统计上不显著。实际上,农村居民家庭部分食品可能源于自家种植,特别是在20世纪80~90年代这种现象更为明显,故农村居民食品支出所占份额与收入、消费的关系不同于城镇居民,与经济增长不存在明显的直接关联。(2)食品价格指标和城镇化的系数均显著为正,与城镇E-change结果一致。(3)其他变量在统计上不显著,与城镇E-change结果一致。(4)第一产业增加值份额的系数为正,但在模型13和模型15中显著,在模型14和模型16中不显著,表明产业结构变化会对恩格尔系数变动产生正向影响,但考虑其他因素后影响减弱,与城镇E-change结果略有不同。(5)城镇与农村E-change结果中多数自变量系数估计值符号相同,只有人均公路里程二者存在差异。可见,交通条件对城镇和农村恩格尔系数的作用方向有所不同。

五、结 语

本文以1980~2016年中国城镇和农村数据为样本,基于Working-Leser模型发现,恩格尔系数与收入水平之间呈负相关关系,恩格尔定律适用于描述中国居民消费支出结构的变动趋势。纳入食品价格因素构建扩展Working-Leser模型结果表明,恩格尔系数与收入水平之间的负相关性是稳健的,而食品价格与恩格尔系数呈显著的正相关关系,若食品价格上涨则居民家庭食品支出在消费总支出中的份额上升。如果纳入其他控制变量,则收入水平和食品价格对恩格尔系数的影响会有所减弱,也就是说,除收入、价格外,恩格尔系数还受货币供应和人口结构等因素的影响。无论是从全国还是地区层面看,中国城乡居民消费结构的变迁都遵循恩格尔定律,但价格因素对恩格尔系数的影响存在一定的地区差异。1980~2016年中国城镇和农村家庭恩格尔系数绝大多数年份下降,但二者变动特征略有不同。进一步分析发现,城镇和农村居民恩格尔系数变动受不同因素驱动,如经济增长对城镇影响显著,对农村影响不显著。GDP增速、收入与消费水平对恩格尔系数变动具有反向作用,而食品价格、人口城镇化率、第一产业份额对恩格尔系数变动具有正向影响。

本研究具有重要的政策含义:首先,从恩格尔系数角度看,中国城乡居民家庭消费结构变迁遵循恩格尔定律。实际上,恩格尔定律是一种长期趋势,时间越长趋势越明显,个别年份恩格尔系数的波动可能是正常的。其次,正如恩格尔定律所指出的,居民家庭食物支出在总支出中所占份额随着收入增长而下降。因此,加快经济发展,提高广大城乡居民收入水平,防止食品价格大起大落,无疑是降低恩格尔系数的最重要手段。再次,中国城乡居民恩格尔系数变动受不同因素影响,改善人民生活水平,引导居民消费结构升级,应针对城乡居民采取不同举措,政策着力点应有所区别。对于城镇居民,重点在于增加收入水平,降低食品价格;对于农村居民,除提高农村居民收入水平外,重点还应关注对农村地区的金融支持,增加社会流动性,同时大力加强交通等基础设施建设。

参考文献:

1. 甘健胜、黄泽民(2006):《城镇居民恩格尔系数降速放缓趋势及其因素分析——基于运用分形分析方法的研究》,《管理世界》,第12期。
2. 霍利斯·钱纳里、莫伊思·赛尔昆(1988):《发展的型式:1950~1970》,李新华等译,经济科学出版社。
3. 马崇明(1994):《试论恩格尔系数在我国的局限性及其解决办法》,《当代财经》,第3期。
4. 王芳(2003):《有关恩格尔系数的统计分析》,《预测》,第2期。
5. 王少飞(2002):《用恩格尔系数衡量居民生活水平的可行性研究》,《统计研究》,第6期。
6. 谢健(1993):《恩格尔定律的适用性及恩格尔系数的修正》,《统计研究》,第1期。
7. 许永洪、曾五一(2009):《基于中国城市统计数据的CPI偏差估计》,《统计研究》,第4期。
8. 尹海洁、唐雨(2009):《贫困测量中恩格尔系数的失效及分析》,《统计研究》,第5期。
9. 张荣山(2007):《欠发达县城镇居民恩格尔系数的特殊性》,《统计研究》,第12期。
10. Allen R., Bowley A. (1935), *Family Expenditure: A Study of Its Variation*. London: P.S. King and Son.
11. Banks J., Blundell R., Lewbel A. (1997), Quadratic Engel Curves and Consumer Demand. *Review of Economics and Statistics*. 79(4):527-539.
12. Chai A., Moneta A. (2010), Retrospectives: Engel Curves. *Journal of Economic Perspectives*. 24(1):225-240.
13. Deaton A., Muellbauer J. (1980), An Almost Ideal Demand System. *The American Economic Review*. 70(3):312-326.
14. Gibson J. (2002), Why Does the Engel Method Work? Food Demand, Economies of Size and Household Survey Methods. *Oxford Bulletin of Economics and Statistics*. 64(4):341-359.
15. Hausman J., Newey W., Powell J. (1995), Nonlinear Errors in Variables: Estimation of Some Engel Curves. *Journal of Econometrics*. 65(1):205-253.
16. Jorgenson D.W., Lau, L.J., Stoker, T.M. (1982), The Transcendental Logarithmic Model of Aggregate Consumer Behavior. In Basman R., Rhodes G., Eds., *Advances in Econometrics*. Greenwich: JAI Press.
17. Kaus W. (2013), Beyond Engel's Law: A Cross-country Analysis. *The Journal of Socio-Economics*. 47:118-134.
18. Lancaster G., Ray R., Valenzuela M. (1999), A Cross-country Study of Equivalence Scales and Expenditure Inequality on Unit Record Household Budget Data. *The Review of Income and Wealth*. 45(4):455-482.
19. Leser C. (1963), Forms of Engel Functions. *Econometrica*. 31(4):694-703.
20. Lewbel A. (2008), Engel's Curve. In Durlauf S.N., Blume L.E., Eds., *The New Palgrave Dictionary of Economics*, Palgrave Macmillan.
21. Muellbauer J. (1976), Community Preferences and the Representative Consumer. *Econometrica*. 44(5):979-999.
22. Ogburn W.F. (1919), Analysis of the Standard of Living in the District of Columbia in 1916. *Journal of the American Statistical Association*. 16(126):374-389.
23. Working H. (1943), Statistical Laws of Family Expenditures. *Journal of the American Statistical Association*. 38(221):43-56.

(责任编辑:李玉柱)