

农村食物市场、食物消费与营养健康:基于 IV-LASSO 方法的检验^{*}

王晓兵 (北京大学现代农学院中国农业政策研究中心,北京,100087)

赵龙强 闵 师 (华中农业大学经济管理学院,武汉,430070)

王 悦 (北京大学现代农学院中国农业政策研究中心,北京,100087)

摘 要:农村食物市场高质量发展是促进城乡公共服务均等化和推进乡村振兴战略的重要举措,对于满足日益增长的食物消费需求和改善农村居民营养健康具有重要的意义。本文基于 2022 年三省农户社会经济调查数据,在应用机器学习 IV-LASSO 算法筛选出工具变量的基础上,利用工具变量法实证估计农村食物市场对农村居民营养健康的影响及其作用机制。研究结果表明,农村食物市场发展显著影响了农村居民营养健康状况,降低了营养不良发生的概率,但同时也增加了超重发生的可能性。此外,农村食物市场对不同群体特征的农村居民营养健康状况存在异质性作用,尤其对家庭人均土地规模较小的群体作用效果较为明显。机制检验结果表明,农村食物市场主要通过提高家庭膳食质量和营养摄入两条路径来实现农村居民营养健康的改善。在实施乡村振兴战略背景下,本文的研究发现为加强农村基础公共服务设施建设,促进城乡公共服务均等化,提高农村居民营养健康提供了经验证据与决策参考。

关键词:农村食物市场;膳食质量;营养健康;IV-LASSO

一、引言

在《“健康中国 2030”规划纲要》《国民营养计划(2017—2030 年)》^①等中长期战略推动下,居民营养健康问题受到政府以及学术界的高度关注。《国民营养计划(2017—2030 年)》指导思想指出,要将营养融入所有健康政策,不断满足人民群众营养健康需求。随着经济发展和居民收入水平的提高,人们对食物的需求从解决温饱问题向追求营养和健康的方向转变,居民膳食结构和营养状况也在

发生巨大变化(青平等,2023;Tian 等,2015)。与此同时,由营养不良、超重或肥胖导致的与饮食密切相关的慢性疾病发生率越来越高,给居民的身体健康造成严重威胁,也给经济发展带来严重负担(Cutler 等,2003;Tafreschi,2015;高杨等,2023)。了解居民的营养健康状况,可为政府制定相关政策措施,例如营养干预、食物补贴、协调食物供给与消费关系等提供参考依据(樊胜根等,2023)。然而,

^{*} 项目来源:国家自然科学基金“乡镇农贸市场对农村居民饮食消费与营养健康的影响机理与效果研究”(编号:72103072),国家社科基金重大项目“推动农业机械化智能化保障粮食安全的路径和机制创新研究”(编号:22&ZD084),国家社科基金重大项目“新形势下我国粮食安全战略问题研究”(编号:22&ZD079),中央高校基本科研业务费专项资金“乡村振兴战略背景下农村经济转型对农村居民食物消费与营养的影响研究”(编号:11041910122),中央高校基本科研业务费专项资金“面向未来的新型食物系统构建及经济社会影响评估研究”(编号:2662021JC003)。闵师为本文通讯作者

① 中共中央 国务院印发《“健康中国 2030”规划纲要》,https://www.gov.cn/zhengce/2016-10/25/content_5124174.htm;国务院办公厅关于印发国民营养计划(2017—2030 年)的通知,https://www.gov.cn/zhengce/content/2017-07/13/content_5210134.htm

我国农村居民的营养健康问题尚未得到足够的重视,缺乏相应的改善措施(高杨等,2023)。健康的人力资本与农村经济发展密切相关,改善农村居民营养健康是实现“全民健康”的重要抓手,也是促进实现乡村振兴的重要动力。

以往大量研究对如何提升我国农村居民营养健康提供了参考(Zhou等,2017;黄泽颖等,2019;Hou等,2021;Ren等,2022),但对于当前我国农村食物市场是否以及多大程度上会影响农村居民营养健康还缺乏足够的认知。随着我国农业从自给自足向市场化转变以及农村居民收入的持续提高,农村居民的食物来源越来越依赖于市场(Huang等,2019)。据相关数据显示,从1978—2019年,农村居民从市场上购买食物的比例由25%增长到80%,农村食物市场发展丰富了农村居民食物消费选择,肉类、水产品、水果等受市场约束较强的食物消费得到增加,在一定程度上可能会提高农村居民膳食质量和营养摄入,并起到改善农村居民营养健康的作用(Zhou等,2017;黄泽颖等,2019;Ren等,2022)。然而,现有研究较多关注农村食物市场可及性对于各类食物消费选择的影响以及由此体现出的饮食多样性、膳食结构和膳食质量等方面的差异(Huang等,2019;Rupa等,2019),少有研究考虑到当地食物市场发展在规模、数量、密度等方面对上述膳食营养指标可能带来的影响。此外,尽管以往的经济学研究对于我国农村食物市场、农村居民饮食消费及其营养健康相关的问题均有涉及(Huang等,2019),但鲜有研究探讨居民饮食消费在农村食物市场影响居民营养健康中的机制作用。

关注农村食物市场发展对农村居民食物消费与营养健康的影响具有重要的理论和现实意义。一方面,拓展了市场约束条件下食物消费行为理论在营养健康方面的应用。基于消费者行为理论,本研究在市场约束下最优食物消费决策基础上(Huang等,1998),进一步探讨农村食物市场发展对农村居民的食物消费数量、膳食质量和饮食结构的影响,进而作用于农村居民营养健康(周津春等,2006)。另一方面,为促进城乡公共服务均等化和提高农村居民营养健康提供了经验证据与决策参考。在“健康中国2030”和国民营养计划(2017—2030年)等政策实施背景下,关注农村食

物市场发展对农村居民营养健康的影响,可以为实现农村居民营养健康和推进健康中国建设提供重要决策依据。同时,在乡村建设行动背景下,推动农村食物市场发展也是实现城乡公共服务均等化以及畅通农村经济内循环的重要措施。

实证检验农村食物市场对居民食物消费与营养健康的影响也面临着如何筛选出最优工具变量,较为精准地处理农村食物市场内生性的问题。首先,农村居民营养健康需求也可能会影响到食物市场上食物店铺的数量和所售卖的食物种类的合理化配置;其次,某些未观测到的因素可能同时影响农村食物市场发展和农村居民营养健康。因此在估计农村食物市场发展对农村居民营养健康的影响时可能会因反向因果关系和遗漏变量而存在内生性问题。以往研究多将食物市场看成外生变量(Kyureghian等,2013;Koppmair等,2017;Headey,2019),仅有少数研究注意到食物市场的内生性问题(Demmler等,2018)。Ren等(2022)运用城市层面的超市环境变量作为社区层面超市环境变量的工具变量,估计了超市环境对农村居民营养健康的影响。然而,衡量超市环境可能存在多种指标,例如超市数量、超市可及性或超市销售食物种类等,在实证分析时并不能直接判断衡量超市环境的哪种指标是最恰当的工具变量。类似地,在解决农村食物市场对农村居民营养健康研究的内生性问题时,也无法直接从多个潜在工具变量集中挑选出最恰当的工具变量。机器学习IV-LASSO方法为解决上述问题提供了潜在的解决方案;该方法突破以往工具变量选择仅基于经济机制或理论存在因果关系的假设,而是基于正则化惩罚提高模型预测精准度,利用调节参数从高维度数据中找出相关性最强的工具变量,更精准地估计内生变量与结果变量之间的因果关系(Bollen,2012)。

本文的研究目标是运用IV-LASSO方法来估计农村食物市场发展对农村居民营养健康的影响,并检验农村居民食物消费的机制作用。为实现该研究目标,本文基于2022年对湖北、吉林和山东三省1080农户的调查数据,并采用IV-LASSO方法选择合适的工具变量与解释变量,更加精准地实证估计了农村食物市场发展对农村居民食物消费与营养健康的影响。研究结果发现当前我国农村食

物市场的发展有助于提升农村居民的营养健康,显著提高了农村居民的 BMI,降低了营养不良发生的可能性,同时也增加了超重发生的概率。对于具有不同社会经济特征的农村居民而言,农村食物市场发展对其营养健康的影响存在异质性。此外,研究结果也验证了农村居民家庭膳食质量和营养摄入是农村食物市场影响农村居民营养健康的重要机制。本研究一方面拓展了 IV-LASSO 方法在我国食物经济研究中的应用,另一方面为发展我国农村食物市场,提升农村居民营养健康的政策设计提供了参考。

本文可能的贡献主要体现在三个方面。首先,本文应用机器学习 IV-LASSO 方法挑选出最佳的工具变量和控制变量,运用工具变量法解决了农村食物市场与农村居民营养健康的内生性问题,通过

控制多个维度的混杂因素,更精准地估计了农村食物市场对农村居民营养健康的影响。其次,本文将农村食物、农村居民食物消费与营养健康三个方面纳入一个分析框架,实证分析农村食物市场对农村居民营养健康的影响,验证农村居民膳食质量与营养摄入是农村食物市场影响农村居民营养健康的重要潜在机制,在一定程度上揭示了我国农村食物市场发展对促进居民营养健康的影响机理。最后,本文在评估我国农村食物市场对农村居民营养健康的影响时,区别已有研究对食物市场可及性的衡量(一定范围内有没有食物市场、距离食物市场的距离),采用所在村委会方圆 5 公里内食物店铺的数量,能够更加全面、准确地反映当地食物市场的发展情况。

二、文献综述与研究假说

(一) 我国农村食物市场发展

新中国成立后我国长期处于计划经济阶段,农村居民食物获取主要是通过集体劳动来实现自给自足,部分加工食品可利用粮票在供销社和食品站购买(陈锡文等,2008;陈丽娜,2018)。由于实行统购统销制度,农村食物市场基本上并不存在。1978 年,我国开启了农村改革的序幕,家庭联产承包责任制的实行推动了农业生产由集体化向家庭化转变。农民可以更加灵活地选择自己种植和养殖的品种,在满足自身食物消费需求的同时,还可将剩余的农产品通过农贸市场、农村集市等场所进行市场交易,所获取的现金收入还可以在市场上购买到其他生活所需的食物(陈锡文等,2008)。为了促进城乡农贸市场的发展,党和政府还提出了一系列发展集市贸易的政策措施^①,鼓励农民在市场上自由交易农产品,为农村居民食物消费提供了更多的选择和机会,农村食物市场得到了初步发展。

20 世纪 80 年代中期,随着农村市场化改革的推进,农村食物市场进入了快速发展阶段。1985 年农产品统购统销体制取消,农村居民在食物获取时不再依赖于粮票供应,通过食物市场可以购买到

种类更加丰富的食物(陈锡文,1988)。同时,为了促进农产品的流通和销售,政府也开始着力推动农村市场体系的建设,大量的农村集市、小卖店、菜店、乡镇农贸市场和集市等相继兴办起来(周洁红等,2004)。需要注意的是,此阶段食物市场规模相对较小,食物供应链以短供应链为主,具体表现为本地农产品的交易。随着城市化的发展和农民进城务工的大趋势,农村人口逐渐流失,农村食物市场发展进入瓶颈期。一方面,农村青壮年劳动力大量减少,影响了农作物生产,使得农产品(特别是高附加值农产品)的数量和种类在一定程度上有所下降,直接影响了本地食物市场的供应能力,进而阻碍了农村食物市场的发展(陈耀庭等,2013;陈丽娜,2018)。另一方面,随着农村人口规模的减少,农村居民从食物市场上获取食物需求相应减少,农村食物市场的规模也将会逐渐缩小。

随着乡村振兴战略的实施,开展乡村建设行动,促进城乡服务均等化,农村食物市场迎来了新的发展契机。在乡村振兴战略的带动下,农村居民收入持续增长,原有的农村食物市场已无法满足农村居民不断增长的多样化和高品质食物消费需求

^① 1979 年《中共中央关于加快农业发展若干问题的决定(草案)》提出,允许农村集市贸易作为社会经济的附属和补充;1983 年的《当前农村经济的若干问题》提出要广辟流通渠道,适当发展个体商业,实行以国营商业为主导、多种商业形式并存的流通局面

(孙凤,2020)。乡村振兴与乡村建设行动的开展为农村食物市场的发展奠定了良好的基础条件,有利于推动农村食物市场的进一步发展。特别是,供销社重返农村,对农村食物市场的繁荣起到推动作用。尽管电商、直播带货等新兴食物销售渠道逐渐进入农村,可能有助于进一步保障农村居民食物供应安全和满足多样化需求。然而,农村居民对这些新兴销售渠道的接受程度远低于城市居民。在未来一定时期内,农村食物市场都将是保障农村居民“菜篮子”安全、满足农村居民对食物的多样化需求的重要公共服务场所。因此,在全面推动乡村振兴的进程中,加强农村本地食物市场建设对于缩小

城乡公共服务差距、提高农村居民膳食质量和营养摄入,进而改善农村居民营养健康状况具有重要的现实意义。

(二)农村食物市场对居民营养健康的影响

参考以往研究(Muthini等,2020;Ren等,2022),本文将农村食物市场发展、农村居民食物消费与农村居民的营养健康状况纳入同一个分析框架。如图1所示,农村食物市场发展可能会通过影响农村居民家庭食物消费行为,包括膳食质量和营养摄入,从而作用于农村居民的营养健康状况(Huang等,2019;Usman等,2022)。

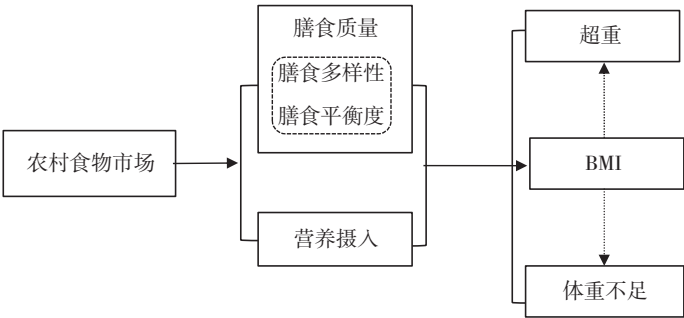


图1 农村食物市场发展对农村居民营养健康的影响机理

首先,农村食物市场发展对农村居民食物消费的影响主要体现在两个方面:一方面,农村食物市场发展可能会影响农村居民的膳食质量(Kihui等,2020;Huang等,2019)。农村食物市场发展有助于农村居民获得更为丰富多样的食物种类选择,例如家庭不易自产的水产品和加工产品(如奶制品)。已有研究证实了农村食物市场可及性的增加显著提高了居民的饮食多样性和膳食质量(Kihui等,2020),尤其是对于那些不从事农业生产的居民来说,这种影响更加显著(Huang等,2019)。此外,农村食物市场发展还增加了居民获取食物的渠道和便利性,促进了各类食物的均衡摄入(Liu等,2014;Huang等,2019)。另一方面,农村食物市场发展可能会影响到农村居民营养摄入的情况(Usman等,2022;Yuan等,2019)。农村食物市场发展能够增加消费者的食物消费选择,促进食物消费种类多样化,进而影响居民对优质能量和蛋白质等营养素的摄入(Usman等,2022;周津春,2006;韩

昕儒等,2012;Yuan等,2019)。

其次,居民的膳食质量、营养素摄入与营养健康有着直接而紧密的联系(Zhou等,2017;黄泽颖等,2019;Ren等,2022)。根据市场约束条件下农村居民的食物消费选择行为(Huang等,1998),农村食物市场发展从不完备到完备,有助于提高农村居民家庭的膳食质量和营养摄入,进而改善了农村居民的营养健康(Zhou等,2017;Ren等,2022)。具体来说:一方面,膳食质量的提高能够改善农村居民的营养健康状况(周津春,2006;Tian等,2015;周莹等,2022)。已有研究表明,膳食多样性不仅可以提高居民营养充足率,降低农村居民疾病发生的概率(Tavakoli等,2016),还能够提高膳食质量,从而改善居民的营养健康水平(Marshall,2001)。此外,Min等(2019)研究证实均衡的膳食摄入会提高营养健康水平,降低居民营养不良发生的概率。另一方面,营养摄入直接影响着农村居民的营养健康。营养摄入量过低可能会导致营养不良(杨月

欣等,2019;李程等,2023),而营养摄入量过多可能会导致营养过剩,引发超重、肥胖和糖尿病等健康问题(Khonje 等,2019;杨月欣等,2019)。近年来,我国农村居民的营养摄入量得到了较大改善,但与中国膳食指南的推荐标准还存在一定差距,农村地区营养摄入不足的问题依旧存在。通过提高农村居民营养摄入量能够有效改善农村居民营养健康状况(李宇等,2023)。营养摄入结构也会对农村居民营养健康状况产生影响。合理的营养摄入结构是改善农村居民营养健康的有效途径(马双等,2011;杨月欣等,2019;李程等,2023)。我国农村居民在营养摄入结构上主要存在碳水化合物摄入过

量、动物蛋白质摄入量比重低等问题(马双等,2011;He 等,2021)。发展农村食物市场能够保障食物供应的多样性和稳定性,在增加优质蛋白摄入的同时降低碳水化合物的摄入量,优化了农村居民的营养摄入结构,进而有助于改善农村居民的营养健康状况(Usman 等,2022;周津春,2006)。根据上述分析,本文提出研究假说:

假说 1:农村食物市场发展能够改善农村居民的营养健康状况。

假说 2:农村食物市场发展能够通过提高农村居民的膳食质量和营养摄入来改善农村居民的营养健康状况。

三、研究设计

(一) 数据来源

本文使用的数据主要来源于两个方面。第一,农户微观调查数据。农户微观调查数据来自课题组于 2022 年 7 月在湖北、山东和吉林三省开展的农户社会经济调查。该数据收集过程中采用分层抽样和简单随机抽样相结合的方法*,共收集了 36 个样本乡镇 108 个样本村中的 1080 个农户。该数据主要包括两个部分:第一部分村级层面调研数据,主要包括村人口规模、经济发展水平和基础设施条件等;第二部分农户层面调研数据,主要包括家庭成员基本信息、家庭生产经营决策、家庭经济状况以及家庭成员的膳食消费等。考虑到本文关注的是农村食物市场对农村居民营养健康的影响,在对调研数据进行预处理后共得到 2248 份成人(18~65 岁)样本数据。在机制分析部分,本文采用农村居民家庭三天膳食消费记录数据来计算家庭膳食质量和营养摄入。第二,农村食物市场数据。农村食物市场数据是通过地理信息数据测量获得,主要包括样本村周围超市、商场、便利店、小卖店、副食店、特产店、熟食店、卤制品店、鲜肉店、水产店、果蔬店、蛋糕店、农贸市场、集市等销售食物的店铺的数量,不涵盖餐馆、饭店等提供在外饮食消费的场所,共获取了 108 个村所在地区的食物市场数据。

(二) 实证策略

1. 应用 IV-LASSO 方法选取工具变量与控制变量。在探究农村食物市场发展农村居民营养健康的关系时可能会存在内生性问题,因此因果效应估计结果存在偏差。具体来看:一方面,农村居民群体的营养健康需求可能会对农村食物市场上食物店铺的数量和市场上售卖的食物种类产生一定影响;另一方面,某些观测到的因素可能会同时影响到农村食物市场发展和农村居民营养健康进而导致内生性问题。工具变量法被广泛用于解决内生性问题,能够得到农村食物市场与农村居民营养健康之间的无偏因果效应估计(Wooldridge, 2010)。传统上关于工具变量的选择主要是基于经济机制和理论基础,但在面临多个符合要求的工具变量时,传统方法难以确定最有效的工具变量。虽然传统的机器学习 LASSO 方法可以进行工具变量的选择,但在操作过程中并不能同时考虑内生变量和控制变量之间的预测关系(Danquah 等,2021)。IV-LASSO 方法综合了工具变量回归和 LASSO 方法的优势,通过引入正则化惩罚来实现变量选择目的,从众多潜在工具变量集中挑选出最优的工具变量和控制变量,并利用工具变量回归来解决内生性问题,实现内生变量对结果变量的高质量预测,更

* 具体抽样过程为:根据湖北、山东和吉林三省内所有县的农业生产总值、农业劳动力数及耕地面积等农业生产信息采用分层随机抽样的方法在上述省份各选取 4 个反映不同层次经济发展水平的代表性县(区),在每个样本县按照相同标准随机选取 3 个代表性的乡镇,在每个乡镇按照相同标准分层选取 3 个样本村,再在每个样本村随机选择 10 户样本农户进行“一对一”入户访谈

精准地估计内生变量与结果变量之间的因果关系 (Tibshirani, 1996; Belloni 等, 2012)。

机器学习 IV-LASSO 方法在选择工具变量时,不需要提前确定相关工具变量的重要性,而是通过数据运行计算推断来挑选出具有重要影响的变量 (Danquah 等, 2021)。通常来讲,工具变量法回归模型可以表示为:

$$y_i = d_i \beta_1 + x'_i \beta_2 + \varepsilon_i \quad (1)$$

$$d_i = Z'_i \pi_1 + x'_i \pi_2 + u_i \quad (2)$$

其中, y_i 代表的是农村居民营养健康结果变量 (BMI 偏离值, 是否超重, 是否低体重), d_i 是内生变量 (农村食物市场), z_i 是工具变量, x_i 是一系列影响农村居民营养健康结果和农村食物市场的控制变量集。其中, 工具变量与控制变量的个数较多, 故可以采用 IV-LASSO 方法来进行工具变量选择 (Belloni 等, 2014; Chernozhukov 等, 2015; Ahrens 等, 2018)。本文遵循 Belloni 等 (2014) 提出的 PDS 算法 (Post Double Selection) 来实现 IV-LASSO 方法选择最佳的工具变量。该算法首先通过 LASSO 等附带正则项的机器学习算法, 经过交叉验证等方法, 识别出一组对结果变量有解释力的变量, 进而重新将结果变量对这些挑选出的特征变量进行普通的线性回归。具体操作步骤为: 第一阶段采用数据驱动的惩罚方法选择内生变量的最佳工具变量和控制变量, 首先对 d_i 进行 LASSO 回归生成向量 x_{1i} 和 z_{1i} , 然后对 d_i 进行再次 LASSO 回归生成向量 x_{2i} , 最后对 y_i 再次运行 LASSO 回归生成 x_{3i} , 即得到最优工具变量 z_i 和具有重要影响的控制变量 x^* (Belloni 等, 2014; Athey 等, 2019; Maruejols 等, 2022)。第二阶段利用选定的工具变量和控制变量来进行工具变量法估计 (Belloni 等, 2014; Chernozhukov 等, 2015)。

具体来说, 第一阶段回归的标准 LASSO 方法是通过最小二乘法的目标函数和惩罚函数之和得到, 该方法的基本思想是在回归系数绝对值总和小于一个常数的约束条件下, 通过最大似然函数使得部分回归系数估计为 0, 从而删除不重要变量并达到提高模型解释力。系数估计如下:

$$\hat{\pi}_1^{Lasso} = \operatorname{argmin}_{\pi_1}$$

* 最终的控制变量包括三个 LASSO 回归方程得到的变量

** $\|x\|_1 = \sum_{n=1}^N |x_n|$ 代表的是参数向量 $x = (x_1, \dots, x_N)^*$ 的 1-范数

$$\left\{ \frac{1}{N} \sum_{i=1}^N (d_i^* - Z'_i \pi_1)^2 + \frac{\lambda}{N} \|\pi_1\|_1 \right\} \quad (3)$$

(3) 式的右侧前半部分为传统的最大似然函数; 后半部分为惩罚项, 可以看作是对回归系数的约束条件, λ 为调节参数, 用于控制惩罚力度通常利用 10 折交叉验证进行选择; 而 $\|\pi_1\|_1$ 为参数向量 π_1 的 1-范数^{**}。当 $\lambda=0$ 时, IV-LASSO 方法就简化为了两阶段最小二乘法估计 (2SLS); 当调节参数过大时, 惩罚项迫使所有的回归系数估计为 0。在适当的惩罚力度参数下, 不重要的变量回归系数将估计为 0, 从而达到剔除不重要变量的目的 (方嫻 等, 2020; Maruejols 等, 2022; Maruejols 等, 2023)。选择最优的力度参数 λ 以挑选出最优的一个或几个工具变量, 将会使得模型预测误差最小。总的来说, 采用 IV-LASSO 方法放宽了工具变量识别的假设条件, 能够从数据中挑选出最佳的工具变量和控制变量, 并进一步利用工具变量回归方法进行实证分析。

2. 应用工具变量法构建实证模型。基于 IV-LASSO 回归选择的工具变量 IV_i^{Lasso} 和控制变量 X_i^{Lasso} , 本文使用工具变量法 (2SLS 和 IV-probit) 进行回归分析, 模型设定如下:

$$M_i = \alpha_0 + \alpha_1 IV_i^{Lasso} + \alpha_2 X_i^{Lasso} + \varepsilon_{ii} \quad (4)$$

$$Health_i = \beta_0 + \beta_1 M_i + \beta_2 IV_i^{Lasso} + \beta_3 X_i^{Lasso} + \mu_{ii} \quad (5)$$

(4) 式为工具变量对农村食物市场发展的第一阶段估计方程, (5) 式是农村食物市场发展对农村居民营养健康的估计方程。其中, M_i 代表的是当地的农村食物市场发展; IV 指应用机器学习在潜在的工具变量集中挑选出的最优工具变量; $Health_i$ 代表农村居民的营养健康状况, 具体包括 (BMI 偏离值, 是否为超重, 是否为低体重), 由于体重不足和超重是二元离散变量, 实证分析中进一步采用 IV-probit 方法; X_i^{Lasso} 指的是一系列影响农村食物市场和农村居民营养健康状况的控制变量, 包括个人特征、家庭特征和村庄特征, 主要是通过 LASSO 挑选出具有重要影响的控制变量; $\alpha_0, \alpha_1, \beta_0, \beta_1, \beta_2$

和 β_3 为待估计参数; ε_{it} 和 μ_{it} 为随机误差项。

(三) 变量说明

1. 被解释变量。农村居民营养健康。当前学界用于衡量营养健康状况的指标较多,主要包括身体质量指数(Body Mass Index, BMI)(Sarkar, 2020)、居民自我评价健康水平以及肥胖、高血压、糖尿病等常见疾病(Choi 等, 2018),本文主要采用 BMI 偏离值来衡量农村居民的营养健康状况。BMI 偏离值是基于受访者实际身高和体重计算出 BMI 并与《中国居民膳食指南(2022 版)》推荐正常值相比的偏离程度,使用 BMI 偏离值衡量农村居民的营养健康更加客观、准确(Ren 等, 2019; 李会等, 2019)。BMI 偏离值的计算方法见(6)式,数字的绝对值趋近于 0,则 BMI 偏离程度越小;反之,数字的绝对值越大, BMI 偏离程度越大。

$$BMI \text{ 偏离值} = \begin{cases} BMI_i - 18.5, & BMI_i < 18.5 \\ 0, & 18.5 \leq BMI_i < 24.0 \\ BMI_i - 24.0, & BMI_i \geq 24.0 \end{cases} \quad (6)$$

此外,根据《中国居民膳食指南(2022 版)》规定,中国成年人的 BMI 在 18.5~24.0(不含 24)为正常,18.5 以下为偏瘦或营养不良,24.0~28.0(不含 28.0)为超重,大于等于 28.0 为肥胖。本文调

查结果表明(见表 1),样本农村居民的平均 BMI 和 BMI 偏离值分别为 23.5334 和 1.1791,有 4% 和 31.23% 的农村居民分别处于营养不良和超重状态。

2. 核心解释变量。农村食物市场。已有文献对食物市场的衡量通常采用居民到最近食物市场的距离或一定区域范围内食物市场的可及性(Koppmair 等, 2017; Ren 等, 2022)。区别于已有研究,本文采用一定区域内食物店铺的数量来衡量农村食物市场发展情况(Kyureghian 等, 2013; Headey, 2019)。食物店铺的数量在一定程度上反映了食物可获得性,一定区域内食物店铺的数量、规模和分布能够更加直接与客观地反映农村食物市场发展程度(Kyureghian 等, 2013; Headey, 2019)。本文根据样本农户所居住的村委会地理信息,采集了以村委会为圆心,半径为 0~5 公里的圆形以内的食物店铺数量的总和*,作为村级食物市场发展情况的衡量指标,调查结果表明村方圆五公里食物店铺数量均值为 27.2362,表明当前多数村庄所在地的食物市场发展水平相对较低。此外,本文还采用村方圆 3 公里以内的食物店铺数量总和作为稳健性检验。

表 1 关键变量说明及描述统计

变量	变量定义	平均值	标准误
被解释变量			
BMI	农村居民的身体质量指数(kg/m ²)	23.5334	3.4551
BMI 偏离值	农村居民实际 BMI 与《中国居民膳食指南(2022)》推荐正常值相比的偏离程度	1.1791	1.9742
营养不良	BMI<18.5	0.0400	0.1961
超重	24.0≤BMI<28.0	0.3123	0.4635
核心解释变量			
农村食物市场	村中心方圆五公里内食物市场数量(个)	27.2362	20.4579
Ln(农村食物市场)	村中心方圆五公里内食物市场数量(个)的对数	2.9872	0.9245
机制变量			
膳食多样性	农村居民食物消费的多样化指数(香农指数)	0.4833	0.1089
膳食平衡度	农村居民食物实际消费数量与《中国居民膳食指南(2016)》推荐摄入量的偏离程度	5.2553	2.1833
蛋白质	农村居民人均每天蛋白质摄入量	60.1450	36.1086
脂肪	农村居民人均每天脂肪摄入量	28.5026	22.3429
碳水化合物	农村居民人均每天碳水化合物摄入量	247.3986	116.9670
样本量	2248		

数据来源:作者根据调研数据整理得到

* 主要包括农贸市场、超市、副食品店、便利店、早餐店、熟食店、生鲜店、肉铺、果蔬店和蛋糕店等

3. 机制变量。

(1) 膳食质量。膳食质量主要从膳食多样性和膳食平衡度两个方面衡量(Bay等,2017;Huang等,2019)。膳食多样性是从消费多元化的角度来评价膳食质量的指标,它鼓励人们消费多种多样的食物,保证人体能够从各类食物中摄入不同种类的营养素,减少单一食物的过量摄入造成的营养不良风险(Bay等,2017)。常用的表示饮食多样性的指标包括饮食多样性得分(Dietary Diversity Score, DDS)、熵值(Entropy Index, EI)、香农期望指数(Shannon Equitability Index, SEI)等,借鉴已有研究(Hou等,2021),本文主要选择香农期望指数作为膳食多样性的衡量指标。膳食平衡度是从均衡饮食的角度来度量膳食质量的指标,它建议人们应该消费适量的各类食物,保证每种食物的摄入量在适宜范围内(Huang等,2019)。膳食平衡度主要是基于中国居民膳食指南2016年中推荐的各类食物消费量,通过赋分形成膳食平衡指数(Dietary Balance Score, DBS)来衡量膳食平衡,膳食平衡指数越高,说明饮食越均衡*。表1调查结果显示样本村居民的膳食多样性指数和膳食平衡得分的整体均值为0.4833和5.2553。

(2) 营养摄入。营养摄入是指从日常食物消费中所获取的各类宏量和微量元素。蛋白质、脂肪和碳水化合物是食物消费与营养摄入相关研究中主要关注的宏量营养元素(李国景等,2018)。根据第六版《中国食物成分表》记载的各类食物主要营养素的含量,本文测算出蛋白质、脂肪和碳水等营养素的摄入情况。计算公式如下:

$$\eta_{ij} = \sum_{k=1}^{12} w_{kj} c_{ik} \quad (7)$$

其中, η_{ij} 表示家庭j每人每日摄入的营养素i的量; w_{kj} 表示家庭j平均每日消费的食物k的重量, c_{ik} 表示每单位食物k中所含营养素i的量。表1调查数据显示,当前农村居民的人均每天蛋白质、脂肪和碳水化合物的摄入量分别为60.1450克/天,

28.5026克/天和247.3986克/天,这与侯明慧(2023)的研究结果较为接近。与《中国居民膳食指南(2022)》推荐的理想膳食模式摄入量相比^①,51.18%和88.42%的样本居民对蛋白质和脂肪的摄入量低于推荐的最低标准,而55.97%的样本居民对碳水化合物的摄入量超过推荐标准。这表明多数农村居民存在碳水化合物摄入过多、蛋白质和脂肪摄入不足的问题,需要进一步改善营养摄入结构(青平等,2023)。

4. 控制变量。为了尽可能排除其他因素对农村居民营养健康的影响,借鉴已有研究(Huang等,2019;Ren等,2022),本文在模型中加入个人、家庭和村庄特征变量。具体看来,个人特征包括性别、年龄、身高、体重、婚姻状况、教育程度等,家庭特征包括家庭规模、同一宗族户数、家庭人均收入、家庭土地规模等,村庄特征包括村人均收入、常住人口数、道路情况等一系列潜在控制变量。考虑到影响农村居民营养健康的潜在因素较多,本文在应用机器学习IV-LASSO方法从众多潜在影响因素中筛选出对农村居民营养健康具有重要影响的控制变量(Belloni等,2014)。

5. 工具变量。本文使用农村居民所在县内其他乡镇的食物市场发展情况作为农村居民所在地食物市场发展的工具变量。乡镇食物市场发展情况通过乡镇中心街道的农贸市场、果蔬生鲜店、综合超市、小超市(商店)、和流动摊贩等食物店铺的数量来衡量。根据居民所在县内其他乡镇的所有类型的食物店铺总数的平均数,以及农贸市场、果蔬生鲜店、综合超市、小超市(商店)、和流动摊贩等5类食物店铺数量的平均数,可以构建出6个潜在的工具变量。表2汇报了所构建6个潜在工具变量的对数。理论上来说,工具变量“县内其他乡镇的食物市场发展情况”与本地农村食物市场发展密切相关,但并不会对生活在本地居民的食物消费与营养健康产生直接影响。具体来说,本研究所选取的工具变量满足相关性要求。同一县内各乡

* 具体测度方法如下:若家庭人均每日消费量刚好在推荐的区间,则记为2分;若家庭人均每日消费某种食物的重量高于推荐消费量上限的2倍或者低于推荐消费量下限的0.5倍,则记为0分;若消费量在推荐量的上限到推荐量上限的2倍之间,或者在推荐量的下限到推荐量下限的0.5倍之间,则记为1分,最后将得分加总即得到家庭膳食平衡指数

① 以轻体力活动作为参考标准,考虑到样本居民的平均年龄为51.23岁,参考《中国居民膳食指南(2022版)》理想膳食模式下人均每天蛋白质摄入量为60~67克,脂肪摄入量为56~64克,碳水化合物摄入量为221~245克

镇食品销售商之间贩卖的食物通常来源于同一个食物批发市场,并且存在着供应链和货物流通的联系,因此同一县内的食物市场之间存在相关性。其次,所选取工具变量也满足外生性要求。农村居民通常更倾向于在本地食物市场进行食物消费,很少或几乎不会去县内其他乡镇食物市场购买食物。因此,县内其他乡镇的食物市场并不会对生活在本地居民的食物消费与营养健康产生直接影响,仅会通过影响本地农村食物市场发展间接影响到农村居民的食物消费和营养健康。此外,尽管随着农村电子商务的发展,农村居民线上外地消费的机会逐

渐增多,但购买食物的比例很小,不会影响工具变量的有效性^①。因此,从理论逻辑上来看,本文所提出的工具变量是合适的。此外,实证检验的结果也表明表 2 中 6 个潜在的工具变量都与核心解释变量显著相关,同时在控制核心解释变量时 6 个潜在工具变量对被解释变量的影响并不显著。因此,所构建的 6 个工具变量无论是从理论上还是实证上来看,都满足了工具变量的相关性与外生性条件。在接下来的实证分析中,本研究将采用机器学习 IV-LASSO 的方法从 6 个潜在工具变量中选择最优工具变量。

表 2 潜在工具变量明及描述统计

变量	变量说明	平均值	标准误
maket1	该县其他乡镇上所有类型食物店铺数的平均数的对数	3.4973	0.5203
maket2	该县其他乡镇上综合超市数量的平均数的对数	0.6090	0.4416
maket3	该县其他乡镇上小超市(商店)数量的平均数的对数	2.1760	0.4524
maket4	该县其他乡镇上农贸市场数量的平均数的对数	0.2634	0.3295
maket5	该县其他乡镇上果蔬店和生鲜超市数量的平均数的对数	0.8390	0.7379
maket6	该县其他乡镇上流动摊贩数量的平均数的对数	3.0212	0.6747

数据来源:作者根据调研数据收集整理得到

四、实证分析结果

(一) 基准回归结果

参考 Belloni 等(2014)的研究,在采用 IV-LASSO 筛选工具变量的回归中采用传统标准误,且在基准回归中限定 IV-LASSO 选择工具变量的数量为 1。表 3 汇报了变量筛选的调节参数估计结果。参考以往研究(Belloni 等,2014;Maruejols 等,2022;Maruejols 等,2023;方炯等,2020),采用十

折交叉验证法选取最终的工具变量。表 3 中十折交叉验证误差最小取值为 3.823706,所对应调节参数 λ 的最优取值为 0.0012,共有 14 个变量纳入回归方程模型,包括 1 个内生解释变量、12 个控制变量、1 个工具变量。变量 market3 为筛选出的最优的工具变量*。

表 3 变量筛选的调节参数估计

调节参数	参数	变量个数	交叉验证误差
第一次调节参数	0.1792	0	3.8984
最优调节参数前一次调节参数	0.0076	14	3.8237
最优调节参数*	0.0069	14	3.8237
最优调节参数后一次调节参数	0.0063	14	3.8237
最后一次调节参数	0.0006	14	3.8241

① 农村居民通过电商平台获取食物的比例较小,并不会影响到工具变量的有效性。当前随着农村电子商务的发展,农村居民通过电商平台进行网络购物,增加了除本地食物市场以外的消费机会。中国消费者协会发布的《2022 年农村消费环境与相关问题调查报告》指出,农村消费者选择在电商平台网络购物的占比仅为 9.7%。且以购买日用商品为主,食物消费占比很少。此外,通过对本文中调研数据进行统计,发现通过网络电商平台购买食物的农村居民样本不足 1%

* 本部分仅以 BMI 偏离值筛选作为样例,其他被解释变量的筛选相类似

图2进一步汇报了交叉验证误差函数 $CV(\lambda)$ 随着 λ 的变化趋势。结果表明随着调节参数 λ 的减小,交叉验证误差函数 $CV(\lambda)$ 值逐渐降低,当 $\lambda=0.0069$ 时 $CV(\lambda)$ 得到最小取值为 3.823706,因

此调节参数最优值 $\lambda_{cv}=3.823706$ 。同时,在最优值 λ_{cv} 附近,交叉验证误差函数 $CV(\lambda)$ 非常平坦,这意味着在最优值附近变化 λ ,对于模型预测能力的影响很小。

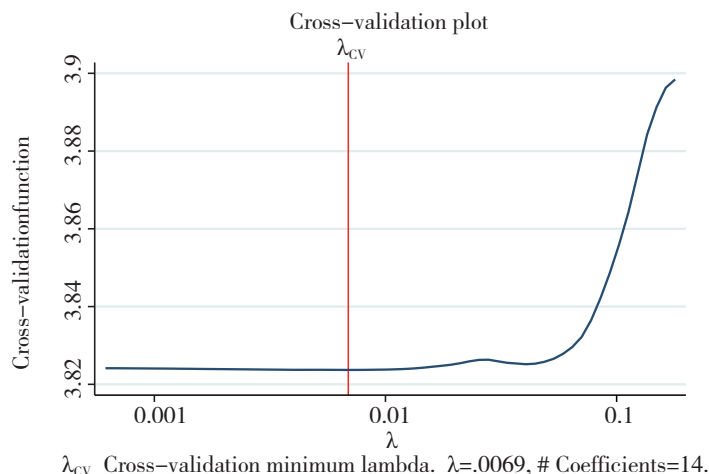


图2 交叉验证

表4(1)(2)(3)列汇报了基于表3筛选出的最优工具变量时农村食物市场对农村居民营养健康影响的估计结果。估计结果表明所筛选出的工具变量与农村食物市场高度相关,且第一阶段F统计值均大于10,不存在弱工具变量问题,说明所选取的工具变量是合理可行的*。估计结果表明,农村食物市场发展显著降低了农村居民BMI偏离程度和体重不足的可能性,同时也增加了农村居民发生超重的概率。具体来看,农村食物市场发展程度每提升1%,农村居民平均BMI偏离值大约减少0.5552,而体重不足概率大约减少3.24%,超重可能性增加约9.73%。农村食物市场的发展虽然有助于提升农村居民的营养健康水平(Ren等,2022),但也可能引发超重肥胖等健康风险(Zhang等,2016;Yuan等,2019)。因此,推动农村食物市场发展的政策设计需要协同考虑提升农村居民的营养健康水平,避免可能带来的超重肥胖风险。

表4(4)(5)(6)列进一步汇报了IV-LASSO应用于多个工具变量的场景。借鉴方娴等

(2020)的研究,本文选择改变IV-LASSO筛选工具变量的个数,拟筛选出2个工具变量,并重新估计农村食物市场对农村居民营养健康的影响。调节参数估计结果表明最优的工具变量为market 2和market 5。估计结果表明,第一阶段F值大于经验值10,不存在弱工具变量问题;过度性识别检验结果显示,不存在过度性识别问题。应用IV-LASSO筛选2个工具变量的模型估计结果与采用1个工具变量时的估计结果基本一致,只是参数大小存在略微差异。该结果再次证实了农村食物市场发展可以显著降低农村居民BMI偏离程度和营养不良发生的可能性,但也增加了超重的可能性。

(二) 机制分析

1. 膳食质量。表5汇报了农村食物市场对农村居民膳食质量影响的工具变量回归结果。基于机器学习IV-LASSO筛选出最佳的工具变量为market 2变量,运用工具变量法回归分析表明农村市场发展程度对农村居民膳食质量的提升具有显

* 本文还采用两阶段最小二乘(2SLS)法和“杜宾-吴-豪斯曼检验”(Durbin-Wu-Hausman Test, DWH)进行内生性检验。检验结果显示, Wu-Hausman F统计量显著(P-value=0.012),拒绝“所有解释变量均为外生、不存在内生变量”的原假设。因此,农村食物市场在解释农村居民营养健康状况时存在内生性问题

著的正向影响。在其他条件不变的情况下,农村食物市场发展程度每提高 1%,膳食多样性增加 0.1897(约 39.34%),膳食平衡得分增加 1.2872 (约 24.49%)。农村食物市场发展会促进农村居民膳食多样性和膳食平衡度的提高,进而改善农村居民营养健康(Huang 等,2019)。

表 4 农村食物市场对农村居民营养健康影响的回归结果

变量	1 个工具变量			2 个工具变量		
	2SLS	IV-probit		2SLS	IV-probit	
	(1) BMI 偏离值	(2) 体重不足	(3) 超重	(4) BMI 偏离值	(5) 体重不足	(6) 超重
Ln(农村食物市场)	-0.5552 ** (0.2395)	-0.4063 ** (0.1956) [0.0324]	0.2861 ** (0.1123) [0.0973]	-0.5409 ** (0.2388)	-0.4136 ** (0.1799) [0.0331]	0.2026 * (0.1091) [0.0686]
控制变量	控制	控制	控制	控制	控制	控制
常数项	0.5836 (1.4735)	2.2944 (1.6017)	-3.0811 *** (0.8920)	2.0547 *** (1.2052)	2.3459 (1.598)	-2.8546 *** (0.8942)
第一阶段回归						
第一阶段 F 值	234.2386	97.7854	119.9888	102.2421	99.1110	114.6742
过度性识别检验	—	—	—	0.2425	0.1932	0.2941
样本量	2248	2248	2248	2248	2248	2248

注:1 个工具变量情形下 BMI 偏离值所对应的最优工具变量为 maket3,超重和体重不足所对应的最优工具变量为 maket5;2 个工具变量情形下 BMI 偏离值所对应的最优工具变量为 maket2 和 maket3,超重和体重不足所对应的最优工具变量为 maket2 和 maket5。表中括号内为标准误;*、**、*** 分别表示在 10%、5%、1%的水平上显著。下同

表 5 农村食物市场对农村居民膳食质量影响的回归结果

变量	2SLS	
	(1)膳食多样性	(2)膳食平衡度
Ln(农村食物市场)	0.1897 ** (0.0510)	1.2872 * (0.7656)
控制变量	控制	控制
常数项	-0.5532 *** (0.1744)	-3.8977 (3.1458)
第一阶段回归结果		
工具变量	0.1653 *** (0.0352)	0.1575 *** (0.0352)
第一阶段 F 值	22.0773	20.0744
样本量	2248	2248

注:膳食多样性和膳食平衡度所对应的最优工具变量为 maket2

2. 营养摄入。表 6 汇报了农村食物市场对农村居民营养摄入影响的工具变量回归结果。IV-LASSO 方法挑选出的脂肪、蛋白质、碳水化合物所对应的工具变量为 maket3。工具变量估计结果表明农村食物市场发展程度每提升 1%,农村居民对蛋白质和脂肪的摄入大约分别增加 0.079% 和

0.086%,而对碳水化合物的摄入并没有显著影响。因此,农村食物市场发展有助于促进农村居民蛋白质和脂肪的摄入,缓解现阶段农村居民蛋白质和脂肪摄入不足的问题,从而影响农村居民的营养健康。

表 6 农村食物市场对农村居民营养摄入影响的回归结果

变量	2SLS		
	(1) Ln(蛋白质)	(2) Ln(脂肪)	(3) Ln(碳水化合物)
Ln(农村食物市场)	0.0792 ** (0.0416)	0.0860 ** (0.0468)	-0.0296 (0.0365)
控制变量	控制	控制	控制
常数项	1.6409 *** (0.3399)	2.5020 *** (0.3764)	6.1225 *** (0.2617)
第一阶段回归结果			
工具变量	0.6682 *** (0.0356)	0.7178 *** (0.0355)	0.6697 *** (0.0356)
第一阶段 F 值	351.7930	409.4144	353.1884
样本量	2248	2248	2248

注:脂肪、蛋白质、碳水化合物所对应的工具变量为 maket3

(三) 异质性分析

考虑到不同土地规模的农户,自产食物及可用于购买食物的收入存在差异,其营养健康与食物消费可能受到市场的影响也有所不同(黄泽颖等,2019)。因此,本文将土地规模按照均值分为两组,进一步探讨农村食物市场对居民个体的 BMI 偏离值、体重不足概率、超重概率的异质性影响。首先,利用机器学习挑选出的 BMI 偏离值、体重不足和超重所对应的最佳工具变量分别为 market 3、market 5 和 market 5;其次,实证结果表明对于土地规模较小的农户家庭而言,农村食物市场对居民个体的 BMI 偏离值和体重不足概率具有显著的正向影响,对超重发生概率具有显著的负向影响(见表

7);而对于土地规模较大的农户家庭而言,农村食物市场对居民个体的 BMI 偏离值、体重不足和超重概率均不存在显著影响。这可能是因为,土地规模较小的农户家庭,其农业生产活动条件相对不利,可能更需要依赖市场获取食物(黄泽颖等,2019)。与自产食物相比,市场所提供的食物更加多样化,且具备更高的营养价值。对于土地规模较大的农户家庭而言,通常能够充分利用现有的农业生产条件,并通过作物种植、畜禽养殖等传统农业生产方式获取日常所需食物,对食物市场的依赖性相对较弱(黄泽颖等,2019)。因此,农村食物市场对家庭土地规模较小的群体的营养健康状况提升更多。

表 7 农村食物市场对不同土地规模农村居民营养健康影响的回归结果

变量	2SLS		IV-probit			
	BMI 偏离值		体重不足		超重	
	小	大	小	大	小	大
Ln(农村食物市场)	-0.8049 * (0.4811)	-0.2967 (0.2346)	-0.4554 * (0.2645) [0.0406]	0.0325 (0.3133) [0.0021]	0.2927 * (0.1639) [0.0976]	0.3091 (0.1948) [0.1063]
控制变量	控制	控制	控制	控制	控制	控制
常数项	0.9503 (1.9455)	0.8652 (2.5746)	0.9401 (1.9002)	7.2573 * (3.7707)	-2.7226 ** (1.074)	-2.3463 *** (1.7704)
第一阶段回归						
工具变量	0.4279 *** (0.0615)	0.8131 *** (0.0454)	0.3737 *** (0.0269)	0.8124 *** (0.0569)	0.3705 *** (0.0270)	0.3981 *** (0.0390)
第一阶段 F 值	48.4833	320.7183	61.2915	37.5744	74.3886	63.0844
样本量	1349	899	1349	899	1349	899

注:农村食物市场变量取对数处理;BMI 偏离值所对应的最优工具变量为 maket3;超重和体重不足所对应的最优工具变量为 maket5

(四) 稳健性检验

1. 更换核心解释变量。考虑到测量过程中可能存在的偏差,本文缩小区域检索范围,采用村方圆三公里内食物市场的数量重新估计农村食物市场发展对农村居民营养健康的影响。首先,基于机器学习 IV-LASSO 算法筛选出用于解决 BMI 偏离值、体重不足和超重内生性问题的最合适工具变量分别为 market 3、market 5 和 market 5 变量。其次,

表 8 回归结果表明农村食物市场发展每提升 1%,农村居民的 BMI 偏离值大约减少 0.9824,体重不足概率大约减少 6.55%,而农村居民超重概率大约增加 17.83%。因此,在更新核心解释变量之后,农村食物市场发展仍然可以显著改善了农村居民的营养健康状况,降低体重不足的概率,同时也增加超重发生的可能性,验证基准回归结果的稳健性。

表 8 农村食物市场对农村居民营养健康影响的回归结果

变量	2SLS	IV-probit	
	(1) BMI 偏离值	(2) 体重不足	(3) 超重
Ln(农村食物市场)	-0.9824* (0.5212)	-0.7778** (0.3594) [0.0655]	0.5121** (0.2245) [0.1783]
控制变量	控制	控制	控制
常数项	3.1202* (1.6628)	1.8227 (1.3245)	-2.8100*** (0.761)
第一阶段回归			
工具变量	0.2433*** (0.0315)	0.1901*** (0.0189)	0.1907*** (0.0189)
第一阶段 F 值	59.7125	115.5339	123.8430
样本量	2248	2248	2248

注:农村食物市场变量取对数处理;BMI 偏离值所对应的最优工具变量为 maket3;超重和体重不足所对应的最优工具变量为 maket5

2. 剔除异常值。为了避免核心解释变量极端值对基准回归结果产生的影响,本文对核心解释变量农村食物市场以 5%的门槛进行双侧截尾,使其更加接近正态分布后重新回归,以尽可能减小异常值对估计结果的影响(王锋等,2022)。首先,基于机器学习 IV-LASSO 算法挑选出最佳的工具变量分别为 market 3、market 5 和 market 5 变量,且第一阶段 F 值大于 10,不存在弱工具变量问题。其次,表 9 结果显示农村食物市场发展程度每提升 1%,农村居民的 BMI 大约增加 0.4870,体重不足概率大约减少 3.26%,而超重概率大约增加 8.47%。除了影响效果大小的差异,剔除异常值后的估计结果与基准回归结果没有明显不同,进一步表明了本文基准回归结果的稳健性和可靠性。

3. 加权后的农村食物市场。考虑到综合超市

和农贸市场内商品数量和种类丰富,和单一的食物店铺相比,更有助于农村居民获得多样化的食物。为更准确量化综合超市和农贸市场对本地食物市场发展的贡献,本文采取对综合超市和农贸市场的数量进行加权处理,并重新估计农村食物市场发展对农村居民营养健康的影响。具体来看说,首先将方圆 5 公里内的综合超市、农贸市场的数量分别赋予 3 种不同的权重,分别为 10、15 或 20;然后,重新计算本地食物市场发展水平;最后,采用加权后的农村食物市场来估计对居民营养健康的影响。表 10 汇报了以 10 作为权重进行加权的农村食物市场对农村居民营养健康影响的估计结果^①。基于机器学习 IV-LASSO 算法筛选出 BMI 偏离值、体重不足和超重分别对应的最合适工具变量分别为 market 1、market 5 和 market 5 变量。模型估计结果

① 篇幅所限,以 15 和 20 为权重进行加权后的模型估计结果在此省略,感兴趣的读者可以向作者索取

表明农村食物市场发展水平每提升 1%, 农村居民的 BMI 偏离值大约减少 0.3676, 体重不足概率大约减少 2.56%, 而农村居民超重概率大约增加 6.37%。无论是选择那种加权方式, 其回归结果均与基准回归结果保持一致, 验证基准回归结果的稳健性。

表 9 农村食物市场对农村居民营养健康影响的回归结果

变量	2SLS	IV-probit	
	(1) BMI 偏离值	(2) 体重不足	(3) 超重
Ln(农村食物市场)	-0.4870 * (0.2673)	-0.4309 * (0.2229) [0.0326]	0.2520 ** (0.118) [0.0847]
控制变量	控制	控制	控制
常数项	1.9495 ** (1.2248)	2.5835 ** (1.8684)	-3.7218 *** (0.9989)
第一阶段回归			
工具变量	0.5057 *** (0.0337)	0.3991 *** (0.0190)	0.3995 *** (0.0191)
第一阶段 F 值	225.6093	85.5318	116.4042
样本量	2103	2103	2103

注:农村食物市场变量取对数处理;BMI 偏离值所对应的最优工具变量为 maket3;超重和体重不足所对应的最优工具变量为 maket5

表 10 农村食物市场(加权)对农村居民营养健康影响的回归结果

变量	2SLS	IV-probit	
	(1) BMI 偏离值	(2) 体重不足	(3) 超重
Ln(农村食物市场)	-0.3676 ** (0.2188)	-0.3226 ** (0.1871) [0.0256]	0.1873 ** (0.1027) [0.0637]
控制变量	控制	控制	控制
常数项	1.0212 (0.9306)	0.6732 (0.9800)	-1.8075 *** (0.5525)
第一阶段回归			
工具变量	0.6200 *** (0.0322)	0.4546 *** (0.0202)	0.4524 *** (0.0202)
第一阶段 F 值	370.5754	89.9885	100.9548
样本量	2248	2248	2248

注:农村食物市场变量取对数处理;BMI 偏离值所对应的最优工具变量为 maket1;超重和体重不足所对应的最优工具变量为 maket5

五、研究结论与政策启示

本文基于 2022 年吉林、湖北和山东三省农村居民家庭三天膳食消费记录数据,利用机器学习 IV-LASSO 算法挑选出最佳工具变量并结合工具变量法实证检验农村食物市场对农村居民营养健康的影响,并按照家庭人均土地规模进行异质性分

析,同时探索了农村食物市场对农村居民营养健康的潜在作用机制。研究发现:第一,农村食物市场显著降低了农民 BMI 偏离值和营养不良发生的概率,但同时也增加了超重发生的可能性。第二,农村食物市场对不同家庭土地规模的农村居民

营养健康状况影响效果不同,对家庭土地规模较小群体的营养健康改善作用效果较为明显。第三,农村食物市场主要是通过提高家庭膳食质量和营养摄入两条路径来改善农村居民的营养健康状况。

研究结果有着重要的实践意义,具体表现在三个方面。第一,积极推广 IV-LASSO 方法在食物经济或其他农林经济管理研究中的应用。突破以往研究在对工具变量筛选中,仅依靠经济机制或理论进行判断,机器学习 IV-LASSO 方法可以帮助在多个潜在工具变量中选出最优的工具变量,同时满足工具变量有效性的检验。第二,在国家发展县域经

济的大背景下,扶持食物市场发展落后地区的市场建设,缩小区域内食物市场差距。通过资金补贴、税收优惠等财政政策,支持发展农村食物市场,对现有的食物市场进行改造或数字化升级等,提高食物供给保障能力,以满足农村居民日益增长的食物消费需求。第三,充分利用农村食物市场,发挥其在提升农村居民膳食质量和营养水平,改善农村居民营养健康状况中的作用。特别是对于家庭人均土地规模较小、自产食物较少或缺乏多样性的地区,增强农村居民通过市场获取食物的便利性,助力提升居民膳食多样性、促进膳食平衡。

参考文献

1. Ahrens, A., Hansen, C. B., Schaffer, M. E. Lasso pack: Model Selection and Prediction with Regularized Regression in Stata. *The Stata Journal*, 2020(1):176~235
2. Athey, S., Imbens, G. W. Machine Learning Methods that Economists Should Know About. *Annual Review of Economics*, 2019:685~725
3. Belloni, A., Chen, D., Chernozhukov, V., Hansen, C. Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain. *Econometrica*, 2012(6):2369~2429
4. Belloni, A., Chernozhukov, V., Hansen, C. Inference on Treatment Effects after Selection amongst High Dimensional Controls. *Review of Economic Studies*, 2014(2):608~650
5. Choi, A., Cawley, J. Health Disparities across Education: The Role of Differential Reporting Error. *Health Economics*, 2018(3):e1~e29
6. Cutler, D., Glaeser, E., Shapiro, J. Why Have Americans Become More Obese. *Journal of Economic Perspectives*, 2003(3):93~118
7. Danquah, M., Iddrisu, A. M., Boakye, E. O., Owusu, S. Do Gender Wage Differences within Households Influence Women's Empowerment and Welfare? Evidence from Ghana. *Journal of Economic Behavior & Organization*, 2021,188:916~932
8. Demmler, K. M., Ecker, O., Qaim, M. Supermarket Shopping and Nutritional Outcomes: A Panel Data Analysis for Urban Kenya. *World Development*, 2018,102:292~303
9. He, G. H., Geng, C. F., Zhai, J. Q., Zhao, Y., Wang, Q. M., Jiang, S., Zhu, Y. N., Wang, L. Z. Impact of Food Consumption Patterns Change on Agricultural Water Requirements: An Urban-Rural Comparison in China. *Agricultural Water Management*, 2021,243:106~504
10. Headey, D., Hirvonen, K., Hoddinott, J., Stifel, D. Rural Food Markets and Child Nutrition. *American Journal of Agricultural Economics*, 2019(5):1311~1327
11. Hou, M. H., Qing, P., Min, S. Multiple Indicators of Household Dietary Diversity in Rural China: Effects of Income and Dietary Knowledge. *Nutrition*, 2021,91:111406
12. Huang, Y., Tian, X. Food Accessibility, Diversity of Agricultural Production and Dietary Pattern in Rural China. *Food Policy*, 2019,84:92~102
13. Khonje, M. G., Qaim, M. Modernization of African Food Retailing and (Un) healthy Food Consumption. *Sustainability*, 2019(16):4306
14. Kihui, E., Amuakwa-Mensah, F. Agricultural Market Access and Dietary Diversity in Kenya, Gender Considerations Towards Improved Household Nutritional Outcomes. *Food Policy*, 2020,100:102004
15. Koppmair, S., Kassie, M., Qaim, M. Farm Production, Market Access and Dietary Diversity in Malawi. *Public Health Nutrition*, 2017(2):325~335
16. Kyureghian, G., Nayga, R. M., Bhattacharya, S. The Effect of Food Store Access and Income on Household Purchases of Fruits and Vegetables: A Mixed Effects Analysis. *Applied Economic Perspectives and Policy*, 2013(1):69~88
17. Liu, J., Shively, G. E., Binkley, J. K. Access to Variety Contributes to Dietary Diversity in China. *Food Policy*, 2014,49:323~331
18. Marshall, T., Stumbo, P. J., Warren, J. J., 2001. Inadequate Nutrient Intakes Are Common and Are Associated with Low Diet Variety in Rural Community-dwelling Elderly. *Journal of Nutrition*, 131,2192~2196
19. Maruejols, L., Höschle, L., Yu, X. Vietnam Between Economic Growth and Ethnic Divergence: A LASSO Examination of Income-Mediated Energy Consumption. *Energy Economics*, 2022,114:106222

20. Maruejols, L., Wang, H., Zhao, Q., Bai, Y., Zhang, L. Comparison of Machine Learning Predictions of Subjective Poverty in Rural China. *China Agricultural Economic Review*, 2023(2): 379~399
21. Muthini, D., Nzuma, J., Qaim, M. Subsistence Production, Markets, and Dietary Diversity in the Kenyan Small Farm Sector. *Food Policy*, 2020, 97: 101956
22. Ren, Y., Zhao, J., Glauben, T., Campos, B. C. Supermarket Environment and Nutrition Outcomes: Evidence from Rural China. *Journal of Rural Studies*, 2022, 92: 79~92.
23. Rooke, J. A., Flockhart, J. F., Sparks, N. H. The Potential for Increasing the Concentrations of Micro-Nutrients Relevant to Human Nutrition in Meat, Milk and Eggs. *The Journal of Agricultural Science*, 2010(5): 603~614
24. Rupa, J., W., Umberger, W. J., Zeng, D. Does Food Market Modernisation Lead to Improved Dietary Diversity and Diet Quality for Urban Vietnamese Households. *Australian Journal of Agricultural and Resource Economics*, 2019(3): 499~520
25. Sarkar, J. Occupational Variation in the Relationship between Child Health and Family Size. *Health Economics*, 2020(1): 98~103
26. Tafreschi, D. The Income Body Weight Gradients in the Developing Economy of China. *Economics & Human Biology*, 2015, 16: 115~134
27. Tavakoli S, Dorosty-Motlagh A R, Hoshiar-Rad A, et al. Is Dietary Diversity a Proxy Measurement of Nutrient Adequacy in Iranian Elderly Women. *Appetite*, 2016, 105: 468~476
28. Tian, X., Yu, X. Using Semiparametric Models to Study Nutrition Improvement and Dietary Change with Different Indices, The Case of China. *Food Policy*, 2015, 53: 67~81
29. Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 1996(1): 267~288
30. Usman, M. A., Haile, M. G. Market Access, Household Dietary Diversity and Food Security: Evidence from Eastern Africa. *Food Policy*, 2022, 113: 102374
31. Woodridge, J. M. Introduction to Econometrics. China Renmin University Press, 2010
32. Yuan, M., Seale Jr J. L., Wahl, T., Bai, J. The Changing Dietary Patterns and Health Issues in China. *China Agricultural Economic Review*, 2019(1): 143~159
33. Zhou, L., Zeng, Q., Jin, S., Cheng, G. The Impact of Changes in Dietary Knowledge on Adult Overweight and Obesity in China. *Plos One*, 2017(6): e0179551
34. 陈丽娜. 取消统购统销 农产品市场化改革的重大步骤. *农村工作通讯*, 2018(Z1): 29~30
35. 陈锡文, 赵 阳, 罗 丹. 中国农村改革 30 年回顾与展望. 人民出版社, 2008
36. 陈锡文. 改善城乡利益关系 深化农村经济改革. *农业经济问题*, 1988(4): 17~21+1
37. 陈耀庭, 蔡贤恩, 戴俊玉. 生鲜农产品流通模式的演进——从农贸市场到生鲜超市. *中国流通经济*, 2013(3): 19~23
38. 樊胜根, 张玉梅. 践行大食物观促进全民营养健康和可持续发展的战略选择. *农业经济问题*, 2023(5): 11~21
39. 方 炯, 金 刚. 社会学习与消费升级——来自中国电影市场的经验证据. *中国工业经济*, 2020(1): 43~61
40. 高 杨, 郑志浩, 张哲晰. 不同收入阶层农村居民的膳食质量演变: 1997—2018 年. *农业技术经济*, 2023(11): 19~37
41. 韩昕儒, 陈永福. 基于收入分层 QUAIDS 模型对农民工家庭内食物消费影响因素的实证分析. *农村经济*, 2016(3): 61~66
42. 侯明慧. 农村居民的收入水平和膳食知识对食物消费的影响研究. 华中农业大学博士学位论文, 2023
43. 黄泽颖, 孙君茂, 郭燕枝, 王秀丽, 马云倩. 农民的农业生产多样性对其饮食多样化和营养健康的影响. *中国农业科学*, 2016(18): 3108~3121
44. 李 程, 朱 婧, 向雪松, 杨月欣. 三大宏量营养素摄入比例与慢性疾病的研究进展. *营养学报*, 2023(2): 1~7
45. 李 会, 王晓兵, 任彦军. 中介效应机制分析的比较研究——来自农民收入与健康的证据. *农业技术经济*, 2019(9): 58~72
46. 青 平, 王玉泽, 李 剑, 闵 师. 大食物观与国民营养健康. *农业经济问题*, 2023(5): 61~73
47. 孙 凤. 新时代我们需要怎样的菜市场. *人民论坛*, 2020(28): 54~56
48. 王 锋, 葛 星. 低碳转型冲击就业吗——来自低碳城市试点的经验证据. *中国工业经济*, 2022(5): 81~99
49. 杨月欣, 葛可佑. 中国营养科学全书. 人民卫生出版社, 2019
50. 周洁红, 金少胜. 农贸市场超市化改造对农产品流通的影响. *浙江大学学报(人文社会科学版)*, 2004(3): 46~53
51. 周津春. 农村居民食物消费的 AIDS 模型研究. *中国农村观察*, 2006(6): 17~22
52. 周 莹, 谢清心, 张林秀, 田 旭. 新冠肺炎疫情对农村居民食物消费的影响——基于江苏省调查数据的实证分析. *农业技术经济*, 2022(7): 34~47

Rural Food Market, Food Consumption, and Nutritional Health: Examination Based on the IV-LASSO Method

WANG Xiaobing, ZHAO Longqiang, MIN Shi, WANG Yue

Abstract: The high-quality development of rural food markets is an important measure to promote the equalization of urban and rural public services and advance rural revitalization. It is of great significance to meet the growing demand for food consumption and improve the nutritional health of rural residents. Based on the 2022 social-economic survey data of three provinces, this study applies the machine learning IV-LASSO algorithm to select instrumental variables and utilizes the instrumental variable method to empirically estimate the impact of rural food markets on the nutritional health of rural residents and its mechanism. The research results indicate that the development of rural food markets significantly influences the nutritional health status of rural residents, reducing the probability of malnutrition while increasing the likelihood of overweight. Furthermore, the impact of rural food markets on the nutritional health status of rural residents with different characteristics exhibits heterogeneity, particularly for groups consisting of households with smaller per capita land size. Mechanism tests reveal that rural food markets improve the nutritional health of rural residents mainly by enhancing household dietary quality and nutrient intake. Under the background of implementing the rural revitalization strategy, this research provides empirical evidence and decision-making references for strengthening the construction of basic public service facilities in rural areas, promoting equalization of urban and rural public services, and improving the nutritional health of rural residents.

Keywords: Rural food market; Dietary quality; Nutritional health; IV-LASSO

责任编辑:吕新业