

线上客户地理分布对外卖零售商家经营决策的影响

——基于饿了么外卖订单数据的实证研究

焦腾啸 李先国 孟 陆

【摘要】近年来,外卖配送行业与外卖订单用户市场规模发展迅速。对于提供外卖配送服务的零售商家来说,需要根据线上外卖需求客户的地理分布情况来调整选址策略与促销策略。本研究采集了“饿了么”蜂鸟配送平台共 6.5 万余条外卖订单数据,建构模型并验证了外卖配送平均距离与商家的订单量及销售量的之间存在的“倒 U 形”关系,发现当特定期限内的商家外卖订单的平均配送距离,即各下单用户与商家的平均距离达到“最优值”时,订单量达到峰值。研究结论对线上订单外卖配送服务的零售商家在选址策略以及广告、促销地点策略的决策方面具有参考价值。

【关键词】外卖订单;配送距离;配送时长;订单量;面板数据分析

【作者简介】焦腾啸(1988-),男,河南许昌人,中国人民大学商学院博士研究生,研究方向:营销模型、移动营销等;李先国(通讯作者),孟陆,中国人民大学商学院(北京 100872)。

【原文出处】《中国软科学》(京),2020.3.122~133

一、引言

近 10 年来,中国餐饮 O2O(Online to Offline)行业、外卖的即时配送行业经过迅速发展,产业链逐步趋于完善,外卖市场也逐步成熟。在外卖“即时配送”的场景中,针对顾客提出的产品或服务的外卖订单,商家和物流配送方需要在短时间内(通常为 0.5~2 小时)完成发货、取货、送货等一系列服务^[1]。2018 年,中国外卖用户规模已超过 3.5 亿人,外卖市场规模突破 2400 亿元,2019 年估计超过 4 亿人^[2]。随着外卖市场的社会关注度逐步提升,学术界中的营销、物流、信息管理等领域也开始对该行业热点进行关注并取得了各类研究成果,如餐饮企业的优惠政策、价格政策、情境因素及菜品设置对销售量、外卖订单量的影响^[3-5],外卖用户的评论、打分、口碑等因素对销售量的影响^[6-7],餐饮企业外卖配送的成本

及效率的优化研究^[8],还有根据餐饮企业的历史销售订单数据结合各类算法模型进行未来订单量预测的研究等^[9]。

外卖用户数量规模的增长也同时伴随着用户需求的增长,用户网上订外卖的行为频率近年来呈现逐年增加的趋势。据相关机构的行业调查结果显示,超过 8 成用户每周至少订一次外卖,其中超过 6 成用户使用外卖平台或外卖 APP 下单的频率达到每周 3~5 次^[2]。此外,外卖需求用户的多样化消费促使外卖平台不断拓展配送品类,除了主流的餐饮外卖服务外,生鲜水果、生活日用、鲜花绿植等也逐步加入到外卖配送的服务场景中。学者们也认为,商家通过提供外卖服务、加盟线上平台可以有效增加销售收入,提高经营利润。Kimes 等(2011)^[10]探讨了餐饮商家的自营配送服务与加盟第三方平台服务

的区别,发现第三方外卖平台具有更好的发展潜力。范慕晗等(2016)^[11]以“饿了么”平台为例详细分析了外卖配送的商业模式与发展趋势,认为外卖平台为餐饮商家、物流企业、用户都带来了价值增值,提高了社会福利。越来越多的餐饮店、药店等商家开始以线上平台的外卖订单作为主要利润来源。

然而,由于商家自身人力、设施、资源等方面的限制以及外卖配送成本的存在,商家能够服务的线上外卖订单用户数量是有限的,且外卖配送服务所能覆盖的范围也是有限的。赵志伟等(2018)^[12]通过对参加加盟外卖平台前后的销售数据进行分析,发现近一半的商家加盟外卖平台后利润未能显著提高,原因是线上的销售利润难以弥补其线下到店顾客流失的利润。物流配送企业需要根据其客户区域的经济环境、自然环境、客户需求响应能力、企业自身的信息化程度、物流配送能力、配送成本控制能力等多个方面确定其能够辐射的地理范围^[13]。由于线上订单需求、线下顾客到店需求的共同存在,零售商家就需要全面考虑线下、线上顾客的各类行为因素与地理分布因素来进行选址策略制定。而对于营业中的零售商家来说,也需要确定其产品宣传、促销的目标地点与目标人群,在充分调动和利用商家的有限资源前提下为尽可能多的线上、线下顾客提供产品和服务,最终实现销售额增长的目的。

那么,具有外卖订单需求的线上顾客的地理分布有何特征?线上顾客的地理分布与外卖订单量之间存在着怎样的关系?随着外卖配送服务的开展,零售商家需要更加深入地了解其线上客户的地理分布特征,并根据线上顾客的分布特征来确定或调整自身的选址策略或促销策略。然而,目前尚未有基于真实外卖订单数据的实证研究来探讨此类问题,因此,本研究致力于弥补这一空缺。我们与提供即时配送服务的物流公司“饿了么蜂鸟专送”建立合作,在31天内连续收集了6.5万份商家外卖订单的数据并建立面板回归模型进行实证分析,探究订单配送距离与订单量之间的关系,并进一步探讨不同评分等级、不同开店时长的商家群体的差异。研究结论为零售商家尤其是以线上外卖订单为主要收入来源商家的选址、促销策略提供理论指导与实践参

考价值。

二、文献回顾与总结

(一)考虑顾客特征的零售商家选址策略

外卖需求的增长意味着商家的线上用户数量的增长,也伴随着商家的受众用户地理分布范围的扩大,因此,商家的选址问题也需要重新审视。现有的零售商家选址理论主要考虑的因素包括业态^[14]、产品与服务定位^[15]、顾客特征^[16-17]以及地理环境因素^[18-19]等方面,其中顾客特征是零售商家选址的重要考虑因素之一。在前人研究成果中,零售商家选址所考虑的顾客特征主要有两个方面:顾客地理分布特征与顾客行为特征。

在地理特征分布方面,主要是结合传统商圈理论^[20-23]所进行的选址策略研究。商圈通常是指地图上以商店所在地点为中心的圆形面积区域,按范围由近及远分为核心商圈、次级商圈、边缘商圈3个层级,每个层级中的顾客特征有显著差异^[23]。商家需要根据居住人口、流动人口的比例特征确定各层级商圈的地理范围,再确定该范围的中心区域,并以此为基础对店址选择进行决策^[24];在顾客行为特征方面,主要是针对不同的顾客类型、各类顾客行为场景进行的选址决策研究,如Hood等(2015)^[25]以城市居住人口数量及比例、便利店吸引力、高收入人群比例等维度为基准,对全英国超过500家便利店进行聚类分析,总结出每一类便利店的选址策略类型;Marianov等(2018)^[26]考虑消费者单次外出购物的行程中在不同的地点做出多次购买不同产品行为的多任务场景,采用电脑仿真模拟的方法得出商家考虑多任务场景时做出的选址决策优于仅考虑消费者单次购买行为的选址决策。

可见,不考虑零售商家本身因素的情况下,在以商家地理位置为中心的特定区域范围内,不同位置分布的顾客、不同行为模式特征的顾客与商家的黏性不同。然而,大部分理论均只考虑线下到店顾客的情形,考虑线上顾客需求特别是线上顾客的即时性需求的商家选址策略方面的研究还较少。当商家提供外卖配送服务后,商家就获得了线上的顾客群体,因此其行为需求特征、分布以及与商家的黏性就需要通过外卖订单的数据进行分析、验证。

(二) 多样化寻求行为

多样化选择行为是一种消费者或用户表现出的追求新奇、崇尚自由选择、从而达到效用最大化的行为,是一种常见的消费行为特征^[27]。Kahn等(1986)^[28]认为,多样化寻求使得消费者的购买行为与上一次购买行为出现偏离,或与以往购买的产品品牌倾向特征偏离,从而表现出重复购买概率降低的趋势。在市场环境中,企业为满足消费者或用户的多样化需求做出了多方面的努力^[29]。比如,同一类食品要设计尽可能多的口味和形状(售卖多种颜色、口味、形状冰激凌的冷饮店),餐饮店要在明确其定位的基础上设计多种菜品并不断研发推出新菜品,外卖平台也为其用户提供了很多的口味、特色品类的餐饮店列表。

前人研究表明,消费产品特征对多样化寻求行为具有显著的影响^[30]。首先,可替代竞争产品的数量越多,消费者越容易出现多样化寻求行为和购买行为。在可替代产品、竞争产品较多的购买场景中(如饿了么APP中提供外卖服务的餐饮店列表),消费者有很多机会进行探索性质的购买并有较大的选择集,相反在竞争产品较少的场景中,消费者因选择受限而很难进行多样化的选择行为^[31]。其次,对于购买频率较高的产品(如日用品),除非消费者对某个品牌的产品有较高的忠诚度,否则往往会因为多次购买而产生无聊情绪,从而产生多样化寻求行为^[30,32]。

在饿了么、美团等APP客户端或PC客户端上,用户在点外卖时通常可看到数量较多的、一系列的商家列表。多数外卖平台上的用户点外卖的频率较高^[2],因此会表现出多样化寻求行为,较少出现重复购买的情形(连续多次在同一商家下单)。首先,对于与用户的空间距离较近(通常为1公里内)的商家来说,多数会以线下到店的形式进行消费,所以当到店次数逐渐增多之后就会产生厌烦情绪,想要尝试其他的更远一些的店铺产品。另外用户也会认为给近距离商家的外卖订单支付配送费是不划算的。其次,当零售商家距离用户的空间距离较远时(通常为3公里以外),配送成本也会随之增加,配送距离更远、时间更长,这样一来用户的等待耐心也会降低。对于多数外卖用户来说,方便、快捷也是用户选择线

上订外卖的重要因素。因此,对于距离较远的商家,除了少数忠诚度高的用户或少数愿意接受长时间等待的用户之外,多数用户往往会因为担心送达时间过长而放弃在该店铺下单,这部分商家很可能会被排除在用户多样化寻求的选择集之外。因此,本研究认为,多数用户在线上平台点外卖时会表现出多样化寻求、每次尝试不同的商家下单的行为,而对于距离较近与距离较远的商家,用户选择下单的概率较小、频率较低。

(三) 配送服务半径

对于依附于电商平台(京东、淘宝、Amazon等)的多数网络店铺,其顾客群体通常分布在全国各地乃至全世界各地,网络店铺位置与顾客位置可能相距数千公里距离,而对应的物流配送也称作“没有范围限度”的服务^[13,33]。而对于“外卖配送”领域的商家来说,其提供的产品或服务无法承受长途运输,那么“没有范围限度”的配送服务从企业的经济性、产品质量的保证性上都难以实现。“配送服务半径”的概念就在这个背景下应运而生,可理解为物流配送方根据自身资源情况与外部环境条件,以满足客户需求为目的,以低成本高水平为目标而形成的配送空间服务范围的度量,表现为以配送站点为中心的圆的半径^[33]。

配送服务半径是将物流配送企业的服务水平、配送成本、受众需求特征等因素进行权衡后,所确定的最佳服务范围的半径,而非最大服务范围的半径^[13]。关于配送服务半径的研究大多数集中于物流服务领域并结合各类算法模型,如Hassin等(2005)^[34]基于物流服务半径的可变性进行了物流服务中心选址问题研究,提出了服务半径的动态规划算法,但未考虑客户满意度等因素;范月娇等(2009)等^[13]以客户满意度、物流企业硬件设施、人员素质等方面为约束条件,采用综合评判决策模型、层次分析等方法确定物流配送企业的最佳服务半径。肖卡飞等(2017)^[35]研究了物流服务领域中的自提柜的选址与服务半径问题,将客户满意度、客户需求类型、自提柜数量等作为约束条件,通过实验仿真得出混合整数规划模型。总体来看,配送服务半径理论主要应用于物流服务领域,而在营销领域中的应用较少。本研究认

为,由于自身条件的限制,提供外卖服务的零售商家也存在配送服务半径,且该服务半径与各外卖订单用户的距离分布密切相关。

三、数据与变量描述

本研究的数据来自于“饿了么”在线外卖平台,以及与其长期紧密合作的“蜂鸟配送”物流服务公司。“饿了么”是目前中国主流的外卖在线预订服务平台之一,而“蜂鸟配送”是为饿了么提供外卖订单即时运送服务的公司,以骑手配送团队为主要构成。“蜂鸟配送”将每个业务所覆盖的城市划分为多个商圈,每个商圈有多个配送站点,每个站点对应于一个配送团队,在限定的区域范围内提供饿了么等平台所下达的配送任务。“蜂鸟配送”在2018年即时配送行业品质服务指数排名中位居榜首^[2]。用户的订餐、订药等需求在“饿了么”上形成订单并同时发送给“蜂鸟配送”的运送团队和商家,商家按照用户需求即刻准备相应的产品和服务,运送团队调派骑手到商家取到产品后,骑手以最快速度赶路,亲手交付到用户所指定地点,即完成一份订单任务。“饿了么”平台公司与“蜂鸟配送”公司在订单数据层面具有共享机制。

通过与“饿了么”平台的多次合作沟通,本研究从2019年6月11日始,连续收集了北京市3个市中心商圈区域、10个配送站点、持续时长为31天,共计超过6.5万条配送订单数据,每份配送订单对应于某位饿了么平台用户的一次需求,包括点餐、购买药品、购买蔬菜瓜果等多个种类。订单数据的条目包括商家名称、站点位置、商圈位置、订单金额、配送时长、配送距离等多项信息。为保证结果的稳健性,本研究删除了配送信息缺失、因故中途取消、配送失败以及

其他各种异常信息的订单数据,保证分析结果的有效性。最终,本研究共得到61878份有效订单数据。

由于本研究以商家的线上订单量作为因变量,因此需要将订单数据整合成为商家销售订单的面板数据。首先,以商家作为个体单位N,以日期作为时间单位T,对每个商家每天的订单数量进行加总,得到商家每日订单量;其次,对于商家每日的所有有效订单,计算日平均订单金额、日平均订单配送距离与标准差、日平均订单配送时长。此外,为控制天气、空气质量因素的可能影响,本研究从天气后报网(<http://www.tianqihoubao.com>)上收集了31天的天气信息,并从中国空气质量在线监测分析平台(<https://www.aqistudy.cn/historydata>)上收集了31天的当日空气质量信息。本研究采用环境空气质量指数(AQI)表示空气质量(参考2012年中国环境保护部发布的《环境空气质量指数(AQI)技术规定(试行)》标准),AQI越高说明空气中PM2.5雾霾越严重。经过整合后,本研究得到包含530个不同商家、时间跨度为31天的非平衡面板数据,样本量为7974。对整合后数据中主要变量的描述性统计结果见表1。

在表1所示变量的描述性统计结果中,订单金额采用的统计口径为商家提供的货品、配送及各种服务的总价格,反映用户的总购买成本。数据样本中商家日订单量的均值为7.761,最大值为232,最小值1,其中商家日订单量为最小值1的观测样本有2184个,占总样本比例27.4%。订单金额及销售额的最小值为0元,0元的订单为商家针对部分产品进行的促销活动,如“饿了么”上的“首单免费”活动。配送距离是指骑手取货地点(即商家)到用户收货点之间的距离,即代表用户与商家的实际空间距离。

表1 主要变量含义及描述性统计结果

变量	变量含义	均值	标准差	最大值	最小值
订单量	商家每日通过饿了么平台接到的订单数总和	7.761	15.264	232	1
订单金额/元	商家每日通过饿了么平台接到的订单金额平均值	55.529	39.208	743	0
销售额/元	日订单量×日订单金额平均值	394.162	762.746	13527.920	0
配送距离/km	商家每日订单配送距离均值	1.462	0.694	6.100	0.100
配送距离标准差	商家每日订单配送距离标准差	0.782	0.407	3.180	0
配送时长/s	商家每日订单配送时长均值	1895.726	1555.811	51564	361
天气	1=“雨天”,0=“非雨天”	0.362	0.483	1	0
空气质量	AQI指数,该值越高表示空气污染越严重	118.059	50.402	193	33

配送时长是指用户下单到订单送达的持续时间。商家在每个特定时期内(如一天 24 小时内),都可以计算出该时期内所有订单的平均配送距离、平均配送时长。配送时长的均值为 1895.726 秒,约合 30~35 分钟,与“饿了么”平台上大多数商家的“期望送达时间”一致。少数订单的配送时间超过两小时(大于 1 万秒),这部分订单为“预订单”(用户下单并设置规定时间送达,通常为 2~3 个小时后送达)而非“即时单”(用户下单后要求尽快即刻送达)。

四、模型与结果

(一)模型设定

对于面板类型的数据,主要有三种模型供回归分析选择:混合效应模型(Pooled OLS)、固定效应模型(Fixed Effects Model)与随机效应模型(Random Effects Model)^[36]。固定效应模型的假设为个体效应在组内固定不变,个体间的差异反映在一个特定的截距项上,每个个体具有单独的截距项;随机效应模型则假设所有的个体具有相同的截距项,但个体间的差异是随机的且反映在随机干扰项的设定上;混合效应模型等同于普通的 OLS 回归。在本研究中,商家作为独立的个体,其个体效应在时间跨度(一个月)内可假设为固定不变的,因此选用固定效应模型更为适合。从统计结果来看,Wald 检验结果表明采用固定效应模型优于混合效应模型($F(379, 5307)=50.40, p < 0.001$),Hausman 检验结果表明原假设被高度拒绝($\chi^2(8)=48.23, p < 0.001$),表明采用固定效应模型比随机效应模型更为适合。考虑到需要控制时间趋势因素对商家订单量的影响,因此采用包含时间趋势的双向固定效应模型对数据结果进行分析。具体的回归模型设定如下:

$$\ln \text{Bill}_{it} = \beta_0 + \beta_1 \ln \text{Distance}_{it} + \beta_2 \ln \text{Distance}_{it}^2 + \beta_3 \ln \text{Dis_sd}_{it} + \beta_4 \ln \text{Time}_{it} + \beta_5 \text{Weekend}_{it} + \beta_6 \text{Trend}_t + \beta_7 \text{Weather}_t + \beta_8 \ln \text{AQI}_t + \mu_i + \varepsilon_{it}$$

回归模型中的下标 i 代表商家个体, t 代表时间(天), μ_i 为以商家为个体单位的固定效应,用于控制商家的个体差异对回归结果的影响; $\ln \text{Distance}_{it}$ 为商家 i 在 t 时间点的订单配送平均距离,并加入该变量的二次方项来验证其与订单量之间的“倒 U 形”关系; $\ln \text{Dis_sd}_{it}$ 为商家 i 在 t 时间点的订单配送平均距

离的标准差; $\ln \text{Time}_{it}$ 为商家 i 在 t 时间点的日订单配送时长; Weather_t 代表是否为雨天的虚拟变量; $\ln \text{AQI}_t$ 为空气质量 AQI 指数。我们对除时间变量、虚拟变量之外的所有的变量都进行了对数转换,通过压缩变量量纲的方式在一定程度上控制离群值、杠杆值的影响,从而使得回归结果更加稳健。模型中还加入了表示周末与工作日的 Weekend 变量(1=周六或周日,0=工作日)用来控制周末和工作日商家的订单数量差异,又加入了时间趋势变量 Trend 控制时间趋势对结果的影响。

(二)订单平均配送距离与订单量之间的“倒 U 形”关系

研究采用逐步回归法,通过不断加入新变量来验证商家日订单量各影响因素的显著性,数据分析工具采用 Stata 14.0 软件,回归结果如表 2 所示。

在模型 1 中先放入所有的控制变量进行回归,发现 Weekend 变量对因变量商家日订单量有显著负向影响($\beta_5 = -0.036, p < 0.05$),说明周末的订单量相比于工作日更少,原因可能是周末上班的用户少于工作日,因此多选择亲自下厨或外出。时间趋势变量、是否为雨天天气、空气质量指数对因变量的影响均不显著,说明时间、天气等外部因素并不影响用户下单的需求。在模型 2、模型 3、模型 4 中逐步加入了订单配送平均距离、配送距离的平方项与配送距离的标准差, R^2 相对于模型 1 也有逐步明显的增加。

在模型 2 中加入了日订单配送平均距离,来衡量商家的外卖客户与商家的空间距离分布。回归结果显示其对日订单量有显著的正向影响($\beta_1 = 0.075, p < 0.001$)。模型 3 中加入了订单配送平均距离的二次方项,回归结果发现其对订单量有显著影响($\beta_2 = -0.147, p < 0.001$),且系数为负,说明配送平均距离与日订单量之间存在“倒 U 形”关系:当配送平均距离从 0 开始增长时,日订单量随之增长,但增加幅度逐渐减小,增加到最高点后开始呈现随着配送平均距离增长而下降的趋势。图 1(a)为订单平均配送距离与日订单量的散点图与拟合线,图 1(b)为调整坐标轴范围之后更为清楚的拟合线,可以看到拟合线呈现出完美的“抛物线”形状。图 1(b)中抛物线的最高点对应的配送距离约为 1.8 千米,与样本中商家

表 2

商家订单配送距离对订单量的影响

解释变量	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
Weekend	-0.036* (-2.295)	-0.039* (-2.527)	-0.032* (-2.116)	-0.044** (-2.848)	-0.030* (-1.991)	0.023 (1.256)
Trend	-0.000 (-0.235)	-0.000 (-0.259)	-0.000 (-0.217)	0.001 (1.728)	0.001 (1.670)	0.002 (1.935)
Weather	0.023 (1.771)	0.023 (1.818)	0.017 (1.358)	0.018 (1.389)	0.015 (1.165)	0.012 (0.806)
lnAQI	0.004 (0.283)	0.003 (0.202)	0.003 (0.237)	-0.005 (-0.310)	0.001 (0.039)	-0.017 (-0.959)
lnDistance		0.075*** (6.228)	0.018 (1.426)	-0.046* (-2.139)	-0.081*** (-3.747)	0.319*** (8.767)
lnDistance ²			-0.147*** (-14.078)	-0.179*** (-7.819)	-0.190*** (-8.372)	-0.170*** (-6.377)
lnDis_sd				0.137*** (11.445)	0.139*** (11.688)	0.059*** (3.725)
lnTime					0.177*** (7.928)	0.311*** (11.843)
_constant	1.296*** (16.341)	1.284*** (16.227)	1.353*** (17.291)	1.947*** (24.873)	0.620*** (3.355)	3.534*** (16.211)
样本量 N	7974	7974	7974	5697	5695	5694
R ²	0.001	0.006	0.032	0.047	0.058	0.063
F	2.412	9.697	41327	37.104	41.209	44.783

的日订单平均配送距离 1.46 千米较为接近。

在模型 4 中加入了配送距离标准差变量,回归结果发现其对日订单量有显著的正向影响($\beta_3=0.137, p < 0.001$)。外卖平台各用户与商家距离值的标准差,反映了各个下单的线上用户与店家空间距离分布的离散情况。标准差较小说明订单送达目的地分布相对集中,标准差大则说明订单送达的目的地分布更加分散无规律。配送距离标准差变量的显著为正的系数也符合实际情况:由于多样化寻求行为的存在,导致用户的需求具有较大的随机性与多样性,因此难以出现数量较多的、与商家距离均接近某个特定值的“聚堆的”用户群体在特定时期内都具有大量订外卖的需求。所以,距离分布的离散程度越高,对应的外卖订单量越高。同一地点集中出现大量顾客选择同一商家并进行线上下单的可能性低。

在模型 5 中加入了订单平均配送时长变量,发现回归结果为配送时长对日订单量有显著的正向影响($\beta_4=0.177, p < 0.001$),即配送时长越长,日订单量反而越高。出现这种反常结果的原因是,配送时长较长、速度较慢通常对应着订单量大而骑手数量有限,从而需要一位骑手同时送多个订单,甚至出现骑

手跨站点区域配送订单的场景。在这种情况下,骑手每次行程需要取到多份外卖货品,因而无法取货后直接送到用户所在地,需要按照配送站点的安排,按规定顺序配送到各个目的地。模型 6 的各解释变量与模型 5 相同,只将因变量替换为商家日订单量对应的日订单销售额(取对数),结果可知,除配送距离一次方项外,其他各解释变量估计系数的正负方向与显著性均与模型 5 相同。

(三)稳健性、内生性检验

面板数据通常具有三种可能的误差结构:异方差、序列相关、截面相关。当这些因素显著存在时,回归结果中各解释变量估计系数的无偏性、一致性和有效性将可能会受到影响^[37]。因此,本研究拟对回归结果中可能存在的误差结构进行验证,以保证回归结果的稳健性。异方差检验参照 Greene (2003)^[38]提出的针对固定效应模型的检验方法,假定所有截面的方差都相等且等于一个常数,通过建立 Wald 统计量进行检验,结果发现存在显著的异方差($\chi^2(380)=1.6e+33, p < 0.001$)。序列相关采用 Wooldridge(2000)^[39]提出的检验方法,若无序列相关,则假定一阶差分后残差相关系数应为-0.5,本研究

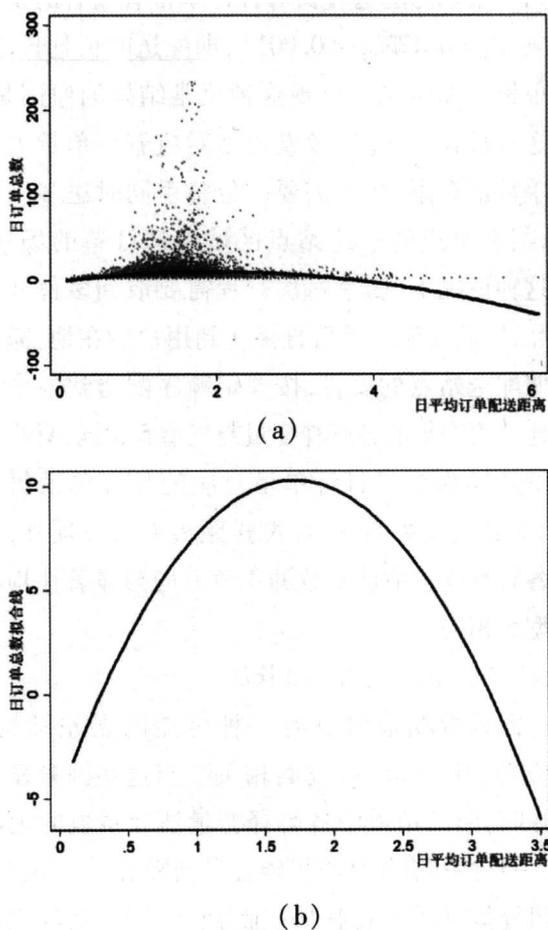


图1 日平均订单配送距离与订单数关系
散点图(a)与拟合线(a)和(b)

中检验结果为显著($F(1, 227)=19.414, p < 0.001$), 说明存在显著的序列相关性。截面相关采用 Pesaran (2004)^[40]提出的平衡与非平衡面板数据均适用的一般性方法, 构造 CD 统计量进行验证, 结果为不拒绝零假设($p > 0.1$)。由于存在异方差、序列相关等问题, 因此需要将各种类型的误差结构对回归结果的影响予以考虑, 并与原固定效应模型进行对比, 结果见表 3。其中, 模型 5 与表 2 相同, 代表原始回归结果, 模型 7 为考虑异方差与序列相关性后采用同样的解释变量进行的回归结果, 模型 8 为同时考虑异方差、序列相关与截面相关进行的回归结果。可以看出, 除控制变量 Weekend 外, 其他各个解释变量系数的正负方向与显著性均是一致的。

本研究采用的回归模型为双向固定效应模型, 但当解释变量存在内生性时, 固定效应模型的估计结果和统计推断结果会存在有偏或不一致的情况。在本研究中, 一方面模型可能存在遗漏变量的问题导致估计结果不准确; 另一方面商家订单量可能反向影响解释变量, 存在双向因果关系。首先, 有需求订单的用户与商家的距离差异以及用户订单量可能会影响商家的日平均配送距离; 其次, 订单量的增加可能导致骑手负荷增加, 延长配送时长, 从而出现商家日订单量反向影响日平均订单配送时长的可能性。

解决面板数据固定效应模型内生性问题的常用

表 3 考虑稳健性、内生性的面板回归结果对比

解释变量	模型 5	模型 7	模型 8	模型 9	模型 10	模型 11
Weekend	-0.030* (-1.991)	-0.030 (-1.178)	-0.030 (-1.779)	-0.031* (-2.010)	-0.042 (-1.791)	-0.042*** (-8.257)
Trend	0.001 (1.670)	0.001 (1.182)	0.001 (1.206)	0.001 (1.708)	-0.002 (-0.548)	-0.001* (-1.995)
Weather	0.015 (1.165)	0.015 (1.268)	0.015 (1.086)	0.013 (1.054)	0.011 (0.523)	0.013*** (3.772)
lnAQI	0.001 (0.039)	0.001 (0.040)	0.001 (0.025)	-0.001 (-0.048)	-0.001 (-0.553)	-0.023*** (-8.969)
lnDistance	-0.081** (-3.747)	-0.081** (-3.196)	-0.081** (-3.575)	-0.074*** (-3.474)	-0.086 (-1.448)	-0.083*** (-3.714)
lnDistance ²	-0.190*** (-8.372)	-0.190*** (-6.136)	-0.190*** (-8.395)	-0.199*** (-8.956)	-0.222*** (-3.551)	-0.335*** (-17.157)
lnDis_sd	0.139*** (11.688)	0.139*** (10.774)	0.139*** (16.460)	0.145*** (12.251)	0.175*** (7.052)	0.194*** (21.574)
lnTime	0.177*** (7.928)	0.177*** (6.487)	0.177*** (5.428)	0.167*** (7.563)	0.237*** (5.020)	0.217*** (9.756)
_constant	0.620*** (3.355)	0.620** (2.869)	0.620* (2.547)	0.304 (1.620)	0.213(0.394)	0.840*** (5.810)

续表 3

解释变量	模型 5	模型 7	模型 8	模型 9	模型 10	模型 11
样本量 N	5695	5695	5695	5695	3855	4475
R ²	0.058	0.058	0.058	0.058	—	—
F	41.209	24.259	112.980	—	—	—
Abond 检验 p 值-AR(1)					< 0.001	< 0.001
Abond 检验 p 值-AR(2)					0.215	0.095
Sargan 检验 p 值					> 0.99	> 0.99

方法为引入工具变量,采用差分广义矩估计(Difference GMM)和系统广义矩估计(System GMM)方法,这两种方法均适用于大 N 小 T 的面板数据^[41]。差分 GMM 采用 Arellano 等(1991)^[42]提出的方法,将所有可能的内生性变量的滞后项作为差分项的工具变量,使用可以更好地处理异方差、序列相关等问题的两阶段估计法进行估计和推断;系统 GMM 相比于差分 GMM 能够更好地缓解差分 GMM 的弱工具变量性等问题,进一步增加了可用的工具变量,且可以同时考虑水平值和差分值。为保证差分 GMM 和系统 GMM 估计和推断结果的有效性,还需要对其进行 Sargan 检验(工具变量的过度识别约束检验,原假设为所有工具变量均有效)和残差项是否存在二阶自相关的 AR(2)统计量 Abond 检验(原假设为残差项不存在二阶自相关),若这两种检验结果不拒绝原假设($p > 0.05$),则说明工具变量选择合理、统计推断结果较为可信。

基于以上考虑,我们将随机效应模型(模型 9)、差分 GMM(模型 10)、系统 GMM(模型 11)的估计结果进行对比分析,如表 3 所示。其中模型 10 中差分 GMM 模型选择一阶差分并加入各内生解释变量的一阶差分项。结果发现,除控制变量外,固定效应、随机效应模型的解释变量系数,以及加入工具变量后的差分 GMM 与系统 GMM 的各解释变量系数在正负方向、显著性水平上仍为一致。Sargan 检验与 Abond 检验的 p 值均大于 0.05,说明工具变量的设定形式可接受。

(四)不同类型特征的商家的“倒 U 形”关系差异

鉴于商家之间的各种特征的差异性,本研究拟进一步对不同特征商家的日订单量与日订单平均配送距离之间的关系差异进行比较,作为延伸性分

析。在本研究中我们考虑商家的开店时长(商家开业日距离样本收集日 2019 年 6 月 10 日的时长,按月计算)与星级评价得分(满分为 5 分)两种情况。我们从“饿了么”App 客户端上收集这两类数据,开店时长在商家页面的营业资质、营业执照上可找到信息,星级评价则是在商家页面的首页上。开店时长反映了商家的生命周期,而生命周期与商家的订单量、销售额高度相关,通常开店后的五年是商家由增长期转为成熟期的时间点^[43],其销售额、接待顾客量的增速开始缓慢甚至下滑。据此,本研究将样本中的商家按照开店时长分为开店一年以内(小于 12 个月)、开店 1 年至 5 年(12 个月—60 个月)、开店 5 年以上(大于 60 个月)3 个类别。星级评价得分则是直接反映商家及其产品形象的综合指标,也是用户在选择商家、下订单时的重要参考标准。通过星级得分数据的收集,发现样本中所有商家的星级得分范围为 3.8 到 5 分,故以评分的中位数为界,将商家分为 3.8-4.4 分、4.5-5 分两类。对不同类别商家的回归结果如表 4 所示。

对于不同开店时长的 3 组商家,其配送距离的二次方项系数均为负且高度显著,说明其“倒 U 形”关系均存在。3 组商家的配送距离标准差系数均为正且高度显著。然而,配送时长变量 lnTime 在开店时长小于 1 年的商家组中不显著,这是因为开店时长小于 1 年的商家的日订单量平均值较小且明显小于其他两组商家(商家组订单量均值 $M_{大于5年}=12.333$, $M_{1年-5年}=10.565$, $M_{小于1年}=2.900$),“骑手资源有限、每趟行程配送多个订单”的场景在该组中不存在;对于两类星级评分的商家,可以发现评分低于 4.5 的商家组回归结果中各个系数的均不显著,可能的原因是评分低于 4.5 的商家数量较少且订单量较少导致结果

表 4 不同开店时间、不同星级评价得分商家回归结果对比

解释变量	开店时长 ≥ 5 年	开店时长1-5年	开店时长 ≤ 1 年	星级评分 ≤ 4.4	星级评分 ≥ 4.5
Weekend	-0.090** (-3.066)	0.003 (0.121)	0.002 (0.085)	0.003 (0.054)	-0.032* (-2.041)
Trend	0.004** (2.963)	0.000 (0.039)	-0.001 (-0.547)	-0.000 (-0.014)	0.001 (1.698)
Weather	0.013 (0.536)	0.019 (0.969)	-0.005 (-0.201)	0.062 (1.145)	0.012 (0.927)
lnAQI	0.020 (0.721)	-0.005 (-0.243)	-0.004 (-0.150)	-0.028 (-0.438)	0.002 (0.155)
lnDistance	-0.070 (-1.281)	-0.107** (-3.145)	-0.025 (-0.776)	-0.016 (-0.181)	-0.085*** (-3.809)
lnDistance ²	-0.257*** (-3.723)	-0.169*** (-5.184)	-0.196*** (-5.385)	-0.106 (-1.008)	-0.192*** (-8.236)
lnDis_sd	0.137*** (4.889)	0.162*** (8.264)	0.099*** (5.571)	0.175** (3.141)	0.137*** (11.278)
lnTime	0.293*** (5.986)	0.219*** (6.074)	0.059 (1.596)	0.188 (1.384)	0.176*** (7.761)
_constant	-0.013 (-0.033)	0.453 (1.531)	1.034** (3.293)	0.498 (0.476)	0.628*** (3.334)
样本量N	1509	2436	1545	276	5419
R ²	0.075	0.063	0.053	0.079	0.059
F	14.577	19.522	9.231	2.733	39.184

不显著。

图 2 中呈现了与图 1 形式相同的,三类开店时长商家的二次项与因变量的拟合线结果。可以发现,开店时长大于 5 年的商家相比于开店时长 1~5 年的商家拟合线的最高点更靠左,即订单量随着订单配送距离的增长速度下降更快。开店 5 年以上商家多数会进入成熟期甚至衰退期,客户数量减少但相对于开店时间较短的商家有更多的忠诚客户且这些客户通常离店家距离也更近。开店时间 1 年以内

的商家虽然也呈现了显著的“倒 U 形”拟合,但由于这部分商家开店时间相对较短,知名度相对较低,因此日订单量相对于开店时间 1 年以上的商家明显较少,所以呈现出较为“平坦”的抛物线形状。

五、研究结论与展望

(一)研究结论

通过线上平台来下订单“点外卖”并享受配送服务的行为,已经逐渐融入到人们的日常生活中。对零售商家来说,外卖市场的兴起意味着其顾客受众从线下为主转变为线上、线下共存,扩大了其利润空间。本研究以零售商家的线上外卖订单为研究对象,基于北京市 3 个城区、10 个商圈、6.5 万条外卖订单数据,验证了配送距离与订单量之间的“倒 U 形”关系。本研究的结果表明,随着商家订单配送距离平均值的增加,商家的订单量会经历先增加后减少的“倒 U 形”过程,订单配送距离标准差、订单配送时长均值正向影响商家的日订单量,从而证明了与商家的距离接近特定值的线上用户点外卖的概率最大、频率最高。另外,对于开店时长大于 5 年的商家来说,其“倒 U 形”的下降趋势早于开店时长小于 5 年的商家,“倒 U 形”关系在饿了么评分大于等于 4.5 的商

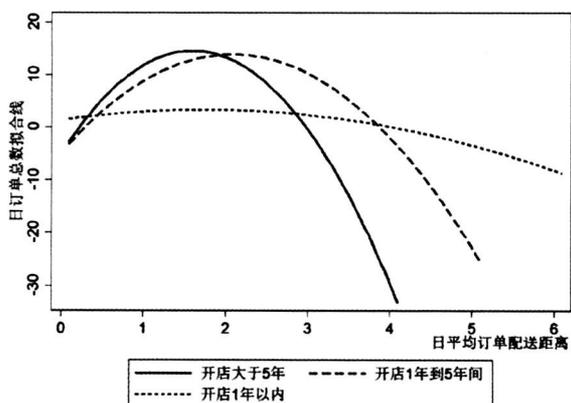


图 2 不同开店时长商家组的“倒 U 形”拟合线

家中显著存在,在低于4.5评分的商家中不显著。

(二)理论贡献

第一,本研究将配送服务半径理论应用到营销领域中,探索了配送距离与外卖零售商销量、订单量之间的内在联系。已有研究发现,在物流配送领域中由于物流企业的成本、资源等条件的约束限制,其配送服务确实存在“最优配送半径”。虽然该理论从逻辑上能够推广到零售商家,即商家的外卖配送服务也存在最优的服务半径,但是缺少理论验证,更缺少大规模、客观的零售商家订单配送数据的实证支撑。因此,本研究基于已有物流领域的研究成果,收集客观数据并从商家的销售量、订单量的角度出发,验证了订单量或销售量与订单配送距离的关系,从而将配送服务半径理论延伸至营销领域中并加以实证检验。

第二,对传统线下零售商家的选址策略理论进行了扩展补充。前人研究表明商家选址一般以顾客集中区域的中心点作为选址参考点,但仅考虑了线下顾客的情形。本研究从线上顾客的视角得出了不同的结论,即“商家选址并不是离中心越近越好”,多数线上顾客与商家之间存在特定的距离。这就为零售商家选址策略的研究领域进行了理论上的补充。

第三,已有的关于零售商家订单量预测的研究更多的是基于商家信息、产品信息、顾客信息、历史订单量信息等视角^[4],而本研究则是从订单配送服务的视角探究了线上客户的地理分布特征对订单量的影响,进而拓展了关于零售商家订单量预测尤其是线上订单量预测的理论体系。

(三)管理启示

第一,选址策略的制定。本研究发现,当零售商家特定时间段内的订单平均配送距离达到某个最优值时,商家的订单量最大。这也意味着当线上外卖订单客户分布区域的中心点与商家的空间距离达到最优时,商家的订单量最大。对于零售商家来说,当目标客户群的分布区域确定后,就需要确定区域的中心点,选择与中心点之间最优距离的地点,作为其选址的目标。特别地,对于门店面积小、类似“小作坊”的,以线上客户作为其主要利润增长来源的小型零售商(如店面小于20平方米的小快餐店)来说,本

研究结论对其在选址策略方面的参考意义更加重要。最优距离可结合过往经验、同类商家的销售订单数据、与商家合作的物流配送公司的数据等多种渠道综合分析后估计。从本研究的数据中得知,对于一线城市人口密集的大型商业区域中的商家来说,最优距离约为1.8千米。

第二,广告促销策略的制定。当商家的店址已经确定时,就需要通过广告促销等宣传手段来吸引更多的顾客,包括线上以外卖需求为主的顾客。根据本研究结论,零售商家需要根据最优配送服务半径来确定线上顾客的目标市场区域,针对区域内的顾客特征确定其广告促销策略。对于分布在该区域的顾客,广告及促销策略需要围绕激励顾客进行线上下单来进行,常用的方法如包含二维码的传单、楼宇广告、APP客户端信息推送等,让顾客通过扫描二维码并链接到商家店铺,点外卖下单可享受各种价格、菜品等多种方面的优惠政策。

(四)研究局限和未来方向

本研究的局限主要存在于数据的地域广度和因果关系的逻辑验证两个方面。首先,本研究主要基于北京地区的外卖订单、饿了么蜂鸟配送公司,研究数据仅限于单一城市、单一配送平台,并未对不同城市、不同外卖平台的情况进行对比分析;其次,本研究的结论仅依赖于客观数据分析,由于数据本身以及研究场景的特征,不能从严格意义上论证因果关系。

未来研究方向可以从跨区域跨平台的收集数据与分析并加入实验以获得一手数据来探索线上顾客选择外卖商家的心理机制两个方面入手。首先,未来研究可以收集跨省份、多级别城市的外卖订单数据,来对比不同区域、不同城市等级的线上顾客地理分布特征对外卖订单销量的影响有何差异,从而进一步扩展本研究结论的外部有效性;其次,可以加入实验室实验或田野实验,从线上顾客选择决策的角度,探索地理距离分布即配送距离对线上顾客选择外卖商家行为的影响机制。

参考文献:

[1]SOLOMON M M. Algorithms for the vehicle routing and scheduling problems with time window constraints[J]. Operations

Research, 1987, 35(2):254-265.

[2]艾媒咨询.2018-2019中国在线外卖行业研究报告[R/OL].<https://www.iimedia.cn/c400/64223.html>.

[3]JEON S, QI J, WANG J. Do local consumers behave differently from visitors? An exploratory study in online group buying[J]. *Electronic Commerce Research & Applications*, 2017, 25:95-104.

[4]赵一星,刘雨萌.基于大学生消费行为的茶餐厅经营对策研究:以武汉商学院“茶语心啡”茶餐厅为例[J]. *中国商论*, 2017(2):9-10.

[5]DOMINIANNI C, LANE K, AHMED M, et al. Hot weather impacts on new york city restaurant food safety violations and operations[J]. *Journal of Occupational & Environmental Medicine*, 2018, 81(7):1048-1054.

[6]张德鹏,陈春峰,张馥丽.在线评价对顾客消费行为的影响:以外卖O2O平台为例[J]. *企业经济*, 2017(3):146-151.

[7]宋悦.基于扎根理论的食客满意度影响因素研究:以大众点评网上海五星级餐厅食客评论为例[J]. *情报探索*, 2018(5):56-61.

[8]YEO V C S, GOH S K, REZAEI S. Consumer experiences, attitude and behavioral intention toward online food delivery(OFD) services[J]. *Journal of Retailing and Consumer Services*, 2017, 35: 150-162.

[9]李凯.基于深度学习的餐饮外卖订单预测模型[D].大连:大连大学,2018.

[10]KIMES S. The future of distribution management in the restaurant industry[J]. *Journal of Revenue and Pricing Management*, 2011, 10(2): 189-194.

[11]范慕哈,周洁如.基于“互联网+”思维饿了么的商业模式创新研究[J]. *上海管理科学*, 2016, 38(6):18-22.

[12]赵志伟,乔晗.餐饮商家加盟外卖平台后能挣得更多么?:基于开封市区267个商家的调查[J]. *经济经纬*, 2018(2): 121-127.

[13]范月娇,潘文军.面向电子商务企业的物流配送服务半径界定[J]. *长安大学学报(社科版)*, 2009, 11(3):21-25.

[14]RAMESH G, PRASAD S, GOYAL S. A pilot study of organised retail formats and their location strategy in Mumbai: A study covering western suburbs from Bandra to Borivali[J]. *Journal of Retail & Leisure Property*, 2011, 9(5): 476-492.

[15]ROSS A, KHAJEHNEZHAD M, OTIENO W, et al. Integrated location-inventory modeling under forward and reverse product flows in the used merchandise retail sector: A multi-echelon formulation[J]. *European Journal of Operational Research*, 2017, 259(2): 664-676.

[16]GREWAL D, LEVY M, KUMAR V. Customer experience management in retailing: An organizing framework[J]. *Journal of Retailing*, 2009, 85(1):1-14.

[17]GLAESER C K, FISHER M, SU X. Optimal retail location: empirical methodology and application to practice: Finalist- 2017 M&SOM practice-based research competition[J]. *Manufacturing &Service Operations Management*, 2019, 21(1): 86-102.

[18]BENOIT D, CLARKE G P. Assessing GIS for retail location planning[J]. *Journal of Retailing and Consumer Services*, 1997, 4(4): 239-258.

[19]MURAD A A. Using GIS for evaluating retail centres location at Jeddah city[J]. *International Journal of Services &Operations Management*, 2011, 10(10): 255-273.

[20]LANGSTON P, CLARKE G P, CLARKE D B. Retail saturation, retail location, and retail competition: An analysis of British grocery retailing[J]. *Location Science*, 1997, 29(1): 77-104.

[21]MCGOLDRICKA P J. Spatial and temporal shifts in the development of international retail images[J]. *Journal of Business Research*, 1998, 42(2): 189-196.

[22]陈金波,邢华,何建敏.中心—外围结构与跨国公司选址研究[J]. *中国软科学*, 2005(4):91-96.

[23]崔振华,杨婷.商业零售店选址分析[J]. *科技与管理*, 2008, 10(2):42-45.

[24]HEMALATHA M, SRIDEVI P, SIVAKUMAR V J. A decision-support system application in retail store location model: A case study of hypermarket in emerging markets[J]. *International Journal of Business &Emerging Markets*, 2017, 3(19): 158-176.

[25]HOOD N, CIARKE G, CLARKE M. Segmenting the growing UK convenience store market for retail location planning [J]. *International Review of Retail Distribution & Consumer Research*, 2016, 26(2): 113-136.

[26]MARIANOV V, EISELT H A, LUER-VILLAGRA A. Effects of multipurpose shopping trips on retail store location in a duopoly[J]. *European Journal of Operational Research*, 2018, 269(2): 782-792.

[27]LEVAV J, ZHU R. Seeking freedom through variety[J]. *Journal of Consumer Research*, 2009, 36(4): 600-610.

[28]KAHN B E, KALWANI M U, MORRISON D G. Measuring variety-seeking and reinforcement behaviors using

panel data[J]. *Journal of Marketing Research*, 1986, 23(2): 89-100.

[29]曹菲,王霞.敬畏情绪对消费者多样化寻求行为的影响及心理机制[J].*中国流通经济*,2018,32(11):104-113.

[30]TRIJP H C M V, HOYER W D, INMAN J J. Why switch? Product category-level explanations for true variety seeking behavior[J]. *Journal of marketing research*, 1996, 33(3): 281-292.

[31]GOURVILLE J T, SOMAN D. Overchoice and assortment type: When and why variety backfires[J]. *Marketing science*, 2005, 24(3): 382-395.

[32]桑辉,贺钰琪,陈小燕.多样化寻求行为文献综述[J].*上海对外经贸大学学报*,2016,23(6):64-72.

[33]范月娇.基于GIS的物流配送中心服务半径的分析与确定[J].*交通运输研究*,2007(9):198-201.

[34]HASSIN R, SEGEV D. The multi-radius cover problem [C]//Workshop on Algorithms and Data Structures. Springer, Berlin, Heidelberg, 2005: 24-35.

[35]肖卡飞,孙咏,王嵩,等.基于渐进服务半径的自提柜选址算法[J].*计算机系统应用*,2017,26(3):187-192.

[36]BALTAGI B. *Econometric analysis of panel data*[M]. John Wiley & Sons, 2008.

[37]BREITUNG J, PESARAN M H. Unit roots and cointegration in panels[M]. Heidelberg: *The Econometrics of Panel Data*. Springer, 2008: 279-322.

[38]GREENE W H. *Econometric analysis*[M]. Pearson Education India, 2003.

[39]WOOLDRIDGE J M. A framework for estimating dynamic, unobserved effects panel data models with possible feedback to future explanatory variables[J]. *Economics Letters*, 2000, 68(3): 245-250.

[40]PESARAN M H. General diagnostic tests for cross section dependence in panels[J]. *Cambridge Working Papers in Economics*, 2004, 69(7):1240.

[41]丁守海,丁洋,沈煜,等.新常态背景下服务业就业的滞后风险[J].*中国软科学*,2016(9):28-36.

[42]ARELLANO M, BOND S. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations[J]. *Review of Economic Studies*, 1991, 58(2): 277-297.

[43]黄明超.餐饮产品生命周期初探[J].*商业研究*,2003(13):173-177.

[44]闫慧丽,彭正银.嵌入视角下社交电商平台信任机制研究:基于扎根理论的探索[J].*科学决策*,2019(3):47-72.

Influence of the Geographic Distribution of Online Customers on the Operation Decisions for Take-out Service Retailers:

Based on the Order Data of Eleme

Jiao Tengxiao Li Xianguo Meng Lu

Abstract: Recently, both take-out delivery service industry and the scale of corresponding user market are developing rapidly. For retailers that provide delivery services, they need to adjust their site location and promotion strategies to cater to the geographic distribution of customers with online delivery needs. This study collected a total of more than 65000 pieces of order data from Eleme Platform, integrated into panel data in unit of retailers, established a regression model to validate the relationship pattern of “Inverted U-curve” between retailer’s take-away service distance and order quantity. We found that during a particular period, while retailer’s average take-away order delivery distance, i. e. the average distance between each online platform customers and the retailer, approximates the “optimal value”, the order quantity approximates the peak. The results of the study provide some reference to the strategic decisions for retailer’s business location, advertising promotion.

Key words: take-out orders; delivery distance; delivery duration; order quantity; panel data analysis