

数量上限、付费位置拍卖与搜索引擎的排序行为^{*}

李世杰

摘要:按照行业规制,中文搜索引擎企业对其付费搜索结果列表实施了位置数量上限的约束。这一约束在付费位置不定数的第二价格拍卖规则下,将对搜索引擎的排序行为产生影响。本文基于中国搜索引擎企业实践,以商品关键词检索领域的消费者层面数据为样本,构造契合消费者顺序检索习惯的实证框架,运用多层贝叶斯回归、断尾回归等分析工具,在数量上限的情形下,考察搜索引擎排序行为的匹配效应与利润效应,探究两者变化趋势是否一致及内在原因,进而对经典文献进行证实或证伪。研究发现:付费位置数量有上限时,搜索引擎排序行为的匹配效应随付费位置降低而凸型下降,且仅有最后1个付费位置的匹配效应为负值;搜索引擎排序行为的利润效应随付费位置降低而凹型下降,但仅有第1个付费位置的利润效应为正值,其他付费位置的利润效应皆为负值;两者呈现不一致性的主要原因是搜索引擎所设的拍卖底价并非理论上的最优拍卖底价,降低了搜索引擎汲取利润的能力。本文实证研究结论否定了经典文献所给出的两者一致性观点,有助于纵深推进付费位置拍卖的理论研究。

关键词:搜索引擎排序行为 付费位置拍卖 数量上限

一、问题的提出

2016年6月,国家网络信息化办公室公布了关于搜索引擎付费位置数量上限的规制政策,即搜索引擎在其检索结果页面安排的付费搜索结果(即广告)的数量不得超过搜索结果条目总数的30%。^①这一产业规制政策对以第二价格拍卖(second price auction)规则为付费位置拍卖机制的搜索引擎企业产生影响。原因在于在第二价格拍卖机制的竞价均衡中,付费位置数量是由搜索引擎预设的最低投标金额、广告商投标金额等因素内生决定,具有不定性特征(Yang et al, 2014)。相应地,付费位置数量上限的设置会对搜索引擎造成影响(Katona & Sarvary, 2010)。首先,由于广告商事前已知晓竞价均衡存在付费位置数量上限的约束,竞标激励更强的广告商则更能契合消费者搜索意愿,进而获取到付费位置并收获更高的消费者点击率,从而增加搜索引擎的拍卖收入。其次,在消费者使用活跃关键词进行检索时,付费位置数量上限约束了搜索引擎安排广告商,影响拍卖收入。因而,搜索引擎对付费位置的广告商进行选择及排序时,需内生地考虑付费位置数量上限造成的影响,进而平衡由此带来的两方面的影响。付费位置广告商的选择及排序是由搜索引擎预设的最低投标金额、广告商投标金额及付费位置数量上限等因素内生决定。

针对商品关键词检索排序,搜索引擎企业会通过压缩每项自然搜索结果列表所占空间,增加搜

^{*} 李世杰,海南大学经济学院,邮政编码:570228,电子邮箱:lshijie@foxmail.com。基金项目:国家自然科学基金地区科学项目“互联网平台企业策略性行为及反垄断规制研究”(71963012);海南省自然科学基金高层次人才专项项目“互联网平台企业核心特质、策略性行为与社会福利效应研究”(2019RC040)。在本文的模型推导及实证研究过程中,得到了西南交通大学蔡祖国博士的协助,特此致谢。感谢匿名评审专家的宝贵意见,文责自负。

^① 详情参见 http://www.cac.gov.cn/2016-05/09/c_1118833529.htm。在这一调查结果中,国家网络信息化办公室要求各搜索引擎企业所设定付费搜索结果,不得超过搜索结果总条目的30%。按当时搜索结果页面设计最大搜索结果条目数为15个进行计算,则付费搜索结果总数最大值为4个。

索结果总条目数量,进而增加付费位置的数量,从而变相突破付费位置数量上限的约束。例如,360搜索引擎为消费者检索平均安排 5.41 个付费搜索结果,且超过 5 个付费搜索结果的占比达到 70.6%;百度搜索引擎为消费者关键词检索平均安排 3.87 个付费搜索结果,但超过 5 个付费搜索结果的占比达到 28.5%;搜狗搜索引擎为消费者检索关键词平均安排 3.86 个付费搜索结果,但超过 5 个付费搜索结果的占比达到 43.2%;必应搜索引擎未消费者关键词检索平均安排 3 个付费搜索结果,且超过 5 个付费搜索结果的占比为 11.7%。不难发现,搜索引擎有一定激励安排更多付费搜索结果,以期增加付费位置拍卖收入。尽管已有文献指出搜索引擎降低自然搜索结果质量的行为得不偿失,但搜索引擎企业变相突破付费位置数量上限的举动日益明显。通常而言,搜索引擎平台是集盈利性和公益性的互联网平台企业;在满足消费者的检索需求后,其可进一步追逐利润最大化目标(李俊生、姚东旻,2016)。不仅如此,搜索引擎可通过提升自然搜索结果的信息质量,更好满足消费者信息检索需求,进而以公益性带动消费者的使用,从而间接推高付费搜索结果的点击率、获取更多利润(White,2013)。显然,中文搜索引擎降低自然搜索结果质量的举措同已有研究文献的理论结论是相反的。因而,考察付费位置数量上限的规制政策造成的影响,对搜索引擎本身或消费者用户均有理论和现实意义。

学界通过构造内嵌消费者顺序检索行为的第二价格拍卖模型,考察固定数量上限约束情形下的广告商竞价均衡,进而探索付费位置数量上限对搜索引擎利润及消费者福利的影响,从而解析数量上限约束对搜索引擎排序行为的作用机理。具体地,付费位置数量上限可提高搜索引擎设置的拍卖底价(reserve price),促成参与竞标的广告商普遍提高质量,进而契合消费者搜索意愿、提升信息匹配效率,从而推动搜索引擎利润最大化追求同最优消费者福利相一致(Athey & Ellison,2011)。不仅如此,也有学者构建付费位置固定数量情形的第二价格拍卖模型考察竞争付费位置的广告商同消费者的偏好的匹配均衡,解析搜索引擎排序结果对消费者福利的影响,并发现搜索引擎排序结果同最优消费者福利是一致的(Chen & He,2011)。总体来看,既有研究文献在考察付费位置数量上限的情形下搜索引擎排序行为时,均发现搜索引擎的排序均衡能契合消费者检索行为,即匹配消费者搜索行为。例如,若消费者采取自上而下的检索习惯,则搜索引擎将质量从高到低的广告商群体安排付费搜索结果列表(Athey & Ellison,2011),或者消费者偏好由强到弱的广告商群体安排在付费搜索结果列表(Chen & He,2011)。进一步地,通过调整拍卖底价为参与付费位置拍卖的广告商设置竞标门槛,进而筛选优质广告商并激励其增大投标金额,获取更优付费位置,进而实现消费者福利与平台利润沿着相同方向增长,促成消费者福利与平台利润的一致性。由于消费者顺序检索习惯下付费位置的点击率呈现凹型下降(Agarwal et al,2011),付费位置数量有上限时,搜索引擎排序行为的匹配效应与利润效应均随着付费位置降低而凹型下降且始终为正值。

但是,由于未结合产业实践数据进行实证检验,经典理论研究结论对产业实践特别是规制政策的参考和启示等应用仍有一定的距离。鲜有经验证据对现有理论研究结论进行证实或证伪。例如,Simonov et al(2018)以微软必应搜索为实验平台,探讨了品牌持有者竞标品牌关键词(brand keyword)的行为,及由此引致的竞争效应与挤出效应。该文明确指出微软必应搜索主动为其付费搜索结果列表设置了数量上限,并在其产业实践中得到严格执行。并且,这一上限数量正好为 4 个,同中国互联网行业规制政策所设的数量上限(4 个)相等。此外,Blake et al(2015)在考察消费者异质性同品牌关键词付费位置拍卖的有效性关联性时,同时使用了微软必应搜索和谷歌搜索为实验平台。需要指出的是,这两篇文献的关注焦点是品牌持有者竞争品牌关键词的行为,而非付费位置上限对消费者搜索行为的影响。因而,他们的研究成果不宜被看作是对付费位置数量上限现象的实证分析。

本文的研究重点是拓展及延伸既有研究的实证框架,结合商品关键词检索领域的消费者层面实践数据,实证检验搜索引擎排序行为的匹配效应与利润效应,并考察两者的一致性。本文试图回答付费位置数量情形下,搜索引擎排序行为的三个关键问题:第一,搜索引擎排序行为能否契合消费者顺序搜索习惯,即匹配效应是否凹型下降且始终为正值;第二,搜索引擎排序行为能否始终提升搜索引擎利润,即利润效应是否凹型下降且始终为正值;第三,匹配效应与利润效应是否一致及其内在原

因。由于消费者层面数据可保留排序结果的原始状态并消除加总层面数据带来的加总偏差(Abhishek et al,2015),本文选择消费者层面样本数据具有较高可行性与可信度。

二、文献回顾与实证框架

(一)付费位置数量上限的第二价格拍卖特征

有少量研究文献讨论了付费位置数量有上限时的搜索引擎排序行为。如 Athey & Ellison (2011)、Chen & He(2011)、Jerath et al(2011)、Amaldoss et al(2016)及 Simonov et al(2018)等研究文献均涉及付费位置数量有上限下的搜索引擎排序行为。其中,Jerath et al(2011)、Amaldoss et al(2016)分别将付费搜索结果列表限制为仅有 2 个付费位置、1 个付费位置,且将参与付费位置拍卖的广告商数量设定为 3 个、2 个,旨在表明数量无限的广告商争夺数量有限的付费位置,发现搜索引擎排序行为在增加平台利润的同时,也可能造成广告商投标竞争的困境,即更高付费位置仅有更低的投标利润。不仅如此,他们通过求解搜索引擎设定的拍卖底价(即最低投标金额),以此显示当消费者点击最后一个付费位置时,广告商应当支付的投标费用。Simonov et al(2018)指出,必应搜索对付费搜索结果列表并设定数量上限约束,借助这一特定机制考察非品牌持有商对品牌关键词检索的盗取效应。其研究焦点为广告商特别是非品牌持有商应当积极参与品牌关键词拍卖,特别是知名度较低的品牌关键词拍卖,以此获取品牌关键词检索带来的流量及消费者需求。需要指出的是,非品牌持有商是指获得品牌授权且能够销售品牌商品的广告商,而非销售假冒品牌商品的广告商。Athey & Ellison(2011)、Chen & He(2011)则详细地考察了数量上限情形搜索引擎排序行为,强调即便是在有限付费位置的情形下,搜索引擎平台给出的排序均衡对消费者检索行为进行匹配,能够对搜索引擎平台利润做出正向贡献。具体地,Athey & Ellison(2011)分别讨论了仅有 1 个付费位置的搜索结果列表、仅有 M 个付费位置的搜索结果列表等情形下的排序均衡;此处的 M 个付费位置并非随机设定,而是搜索引擎平台刻意设定。Chen & He(2011)以谷歌搜索引擎为例,将付费位置数量上限设定为 3 个。另外,Chu et al(2020)考察电商平台采用第二价格拍卖规则拍卖搜索结果排序位置的作用机理,发现固定数量广告位置搜索结果列表可为电商平台创造更多的拍卖收入。

不难发现,仅有 Athey & Ellison(2011)及 Chen & He(2011)聚焦付费位置数量上限对搜索引擎排序行为的影响,并且证明了搜索引擎排序行为的匹配效应与利润效应是一致的。不仅如此,有两点共识值得关注。第一,从理论上,两者均以第二价格拍卖为理论基础,即搜索引擎平台排序规则:按广告商投标金额大小依次排定付费位置顺序;占据付费位置的广告商在获取消费者点击后,需要支付的是下一个付费位置的广告商投标金额。第二,广告商特征内生于广告商的投标金额,并外现于搜索引擎平台的排序均衡,即广告商特征不会直接影响搜索引擎利润。

进一步地,比较数量不限与数量有上限的第二价格拍卖研究文献,可知付费位置数量上限与付费位置数量不限的本质差异:搜索引擎是否设立明确的拍卖底价。所谓拍卖底价是指搜索引擎为参与付费位置拍卖的广告商设置的最低投标金额。仅有投标金额不低于拍卖底价的广告商,才会被搜索引擎设定为付费位置候选者。通常而言,数量不限的第二价格拍卖规则没有设立明确的拍卖底价。在理论研究与产业实践中,搜索引擎仅要求广告商投标金额高于 0 即可(Edelman et al,2007; Varian,2007)。不仅如此,学界普遍认为参与付费位置拍卖但未得到付费位置的广告商,其投标金额完全一致。这意味着最后一个付费位置广告商的实际支付金额(即广告商被消费者点击时向搜索引擎支付的投标费用),为参与付费位置拍卖但未得到付费位置的广告商投标金额。故在数量不限情形下,第二价格拍卖规则虽名义上要求广告商投标金额高于 0 即可,实质却将未获得付费位置的广告商投标金额作为隐性拍卖底价,且将之设定最后一个付费位置的支付金额。数量上限的第二价格拍卖规则设立明确的拍卖底价,且选取投标金额名次与付费位置数量相同的广告商,按投标金额大小依次排定顺序(Athey & Ellison,2011)。这意味占据最后一个付费位置的广告商支付金额为明确的拍卖底价。因而,在数量上限的第二价格拍卖规则中,拍卖底价是决定搜索引擎排序均衡的关

键因素。其原因在于拍卖底价同时决定着消费者福利与平台利润。不仅如此,当搜索引擎结合消费者搜索行为与广告商特征内生地设定拍卖底价,则其所定的拍卖底价可实现消费者福利与平台利润一致地最优,即最优拍卖底价。但是,经典研究文献并未给最优拍卖底价明确的表达式,仅给出解析式。即便是搜索引擎仅设置1个付费位置的情形,最优拍卖底价也没有明确的表达式,只有解析式^①(Athey & Ellison, 2011)。相应地,在实践中学者基本不能明确推断出搜索引擎设置的拍卖底价是否为最优拍卖底价,进而默认搜索引擎给出的拍卖底价为最优底价,从而对样本数据实证检验。

(二) 搜索引擎排序行为匹配效应的实证策略

依据前述分析,本文将默认搜索引擎所设拍卖底价为最优拍卖底价,通过设计实证框架,结合消费者层面的样本数据,实证检验产业实践与理论研究结论间的契合性,并在两者出现不一致性时,探索明确拍卖底价的非最优性对实证结果的影响。本文实证框架主要包含两个方面:第一,实证检验搜索引擎排序均衡同消费者检索行为的匹配效果(以下简称“匹配效应”);第二,实证检验搜索排序均衡对平台利润的影响效果(以下简称“利润效应”)。并且,先有良好的匹配效应,后有一定的利润效应。利润效应是搜索引擎排序行为在匹配效应内生出来的结果。经典研究文献证明,在付费位置数量有上限约束时搜索引擎通过调整拍卖底价设置促成消费者福利与平台利润沿着相同方向增长,进而实现匹配效应与利润效应的一致性(Athey & Ellison, 2011)。接下来,本文将详细阐述匹配效应与利润效应的实证框架。

通常而言,搜索引擎排序结果同消费者搜索行为间存在双向因果关系。例如,当消费者采取包含知名品牌名称的关键词进行检索时,搜索引擎倾向于将竞争者(即获得知名品牌销售权的销售商)置于付费搜索结果列表,并将知名品牌的官方销售商置于自然搜索结果的最高位置,期望消费者通过逐个点击的行为,增加付费搜索结果的点击率,进而获取更高付费位置拍卖收入。此排序结果仅会促使不熟悉知名品牌的消费者点击付费搜索结果,而促成熟悉知名品牌的消费者改变搜索行为跨越付费搜索结果列表,直接进入自然搜索结果列表点击官方销售商(Blake et al, 2015)。相反,当消费者采取知名度较弱的品牌关键词进行检索时,其搜索行为将保持顺序搜索特征,进而使搜索引擎将知名度较弱的品牌的官方销售商尽可能置于更高付费位置。此排序结果不仅增大消费者点击官方销售商的概率,并且显著提升了付费位置的消费者点击率(Simonov et al, 2018)。鉴于此,学界通常将排序结果同消费者搜索行为间的关系称为“匹配效果”(曲创、刘重阳, 2019)。进一步地,本文将其概括为“匹配效应”。

反映搜索引擎排序结果同消费者检索意愿匹配效果的指标通常是消费者点击率。相应地,消费者点击率在消费者层面数据格式下的变量:一个反映消费者是否点击付费搜索结果的二元变量。因而,本文以消费者是否点击付费搜索结果的二元变量为因变量、以付费搜索结果的具体特征为自变量,构造一个probit回归模型或logit回归模型。由于仅能对单个付费位置的消费者点击率进行逐个回归分析,probit回归模型或logit回归模型在全面反映消费者搜索行为上有着较大缺陷。其原因在于消费者搜索行为是连续过程且顺序搜索特征十分明显(Kim et al, 2010; Chen & Yao, 2017)。部分学者对消费者在搜索引擎中的搜索行为进行了描绘,进而构造了契合消费者顺序搜索的实证模型,从而较为全面地探索消费者搜索引擎的影响因素和作用机理^②(Park & Agarwal, 2018)。具体地,当消费者向搜索引擎键入搜索项目后,可选择是否进入付费搜索结果列表(即进入行为)。若进

① 尽管已有研究文献对经典研究文献进行多个方向拓展,如考察消费者与广告商的双向信息不对称的情形下搜索引擎最优拍卖机制等(Chen, 2021),但均未提及最优拍卖底价的实现机制与测算规则。因而,最优拍卖底价的测算仍比较模糊。

② 需要指出的是,另一个比较常用的实证框架包含了消费者点击行为和购买行为的分析框架(Ghose & Yang, 2009; Yang & Ghose, 2010; Agarwal et al, 2011)。其内含的理论假设是消费者以关键词检索进入搜索引擎后依据搜索引擎的排序结果,自上而下地顺序点击广告商网址链接,直至找到满足自身偏好的广告商并最终实施购买。显然,若采用此实证框架则需要提供消费者是否在广告商处进行购买的变量数据(即购买行为的数据)。因而,数据格式将是由广告商提供的加总数据,即包含了消费者在较大差异关键词下的行为数据,与本文所用的消费者层面数据结构有着本质差异。故不予考虑。

入付费搜索结果列表,则需依据自上而下的顺序,选择是否点击付费位置的广告商链接(即点击行为)。在获取到相应的信息后,可选择是否离开付费搜索结果列表(即离开行为)。不仅如此,除进入行为外,余下的两个连续行为是彼此部分可观测,需借助多层贝叶斯回归模型考察两者的影响因素,进而探索排序结果对消费者点击率的影响,从而解析排序行为的匹配效应。

有鉴于此,本文尝试对 Park & Agarwal(2018)的实证模型进行拓展,以期构建契合付费位置数量情形下的匹配效应的实证模型。^① 第一,本文在选择表征付费位置排列顺序的广告商时以广告商在付费搜索结果列表中出现的次数作为标准,而非广告商的消费者点击率。尽管搜索引擎在决定广告商排序时会参考其上一期的消费者点击率,但出现次数反映的是搜索引擎依据广告商投标金额以及上一期消费者点击率等影响因素给出排序结果(Ghose & Yang, 2009; Simonov et al, 2018)。因而,采用出现次数作为标准更能体现凸显其在搜索引擎排序中的影响程度,进而真实反映出由其引致的匹配效应,并且可消除出现次数较少但点击率较高的偶然性现象。第二,本文对进入实证模型的广告商数量进行限定,即将参与实证模型的广告商数量设置为付费位置数量上限。在付费位置数量不定情形下,采用第二价格拍卖机制的搜索引擎通常安排付费位置数量均值为 5 个(Yang et al, 2014)。这一均值已经高于我国搜索服务领域 4 个付费位置的上限。相应地,若将参与回归模型的广告商数量同付费位置数量上限联系起来,可带来两点优势:实证模型可反映出付费位置的稀缺性;有助于观察到广告商排序的组合是紧邻型还是空缺型。所谓紧邻型排序是指搜索引擎将最优排序下的广告商,按投标金额大小依次安排付费搜索结果列表的有限位置;空缺型排序是指搜索引擎将最优排序下的广告商不完整地安排在付费搜索结果列表的有限位置,形成类似缺口型的付费位置排序,并由其他广告商填补缺口付费位置。

(三) 搜索引擎排序行为利润效应的实证策略

依据匹配效应的实证模型,匹配效应不仅有消费者点击行为还有消费者进入行为及消费者终止搜索行为。当搜索引擎采用第二价格拍卖规则时,仅在消费者点击付费搜索结果后,搜索引擎才会获得相应的利润。因而,搜索引擎排序行为引致的匹配效应暗含了同一付费搜索结果列表中多次点击行为。有此前置条件的支持,搜索引擎在获得稳定的匹配效应的基础上,试图将虚假宣传且投标金额较高的广告商置于优先位置,进而骗取消费者点击率,从而获得更高的拍卖收入。这一机理反映了社会舆论对搜索引擎给予虚假宣传的广告商更高位置的批评。

进一步地,搜索引擎排序结果能否为消费者所点击直接关系到付费位置拍卖收入。并且搜索引擎排序结果针对的是具体广告商。占据付费位置的具体广告商,能否为消费者点击决定了具体广告商对搜索引擎利润的贡献。在搜索引擎排序行为下,具体广告商对搜索引擎利润的贡献可称为“排序行为的利润效应”。

在消费者层面数据支持下,最优排序结果的表现是不同广告商占据不同付费位置。不同广告商对搜索引擎利润的贡献实质是不同付费位置对搜索引擎利润的贡献。以表征是否为特定广告商的二元变量为自变量、以搜索引擎利润为因变量构建一个线性回归模型是解析广告商对搜索利润的贡献,是常见的实证策略。此实证策略虽简单可行,但容易忽略因变量为其他类型分布的具体特征。原因在于线性回归模型的前提假设之一是因变量正态分布,但搜索引擎利润变量既不是理论上的正态分布,也不是实践意义上接近正态分布。^② 具体地,在消费者层面数据结构下,搜索引擎利润变量是消费者点击行为的二元变量同下一个付费位置广告商投标金额变量(或搜索引擎设置的最低投标金额)相乘的结果。当占据付费位置的广告商未被消费者点击时,即便给出了投标金额也不需支付

^① 本文拓展 Park & Agarwal(2018)实证模型的具体过程因篇幅所限,备索。

^② 需要指出的是,以广告商获取消费者点击率的平均成本作为搜索引擎利润,在经过加总处理后,其数值分布特征将会比较接近正态分布(Agarawal et al, 2011; Yang et al, 2014)。当研究者进行贝叶斯线性回归或最小二乘线性回归时,可得到广告商对搜索引擎利润贡献的较为全面的影响。如前所述,加总数据将不可避免地带来加总误差。而且,本文所用数据为消费者层面数据,与加总数据有着本质差异。因此,不考虑采用加总方式处理搜索利润变量。

任何费用。当其为消费者所点击时,其支付金额为下一个付费位置上的广告商投标金额(或搜索引擎设置的最低投标金额)。

线性回归模型的实证结果反映的是广告商对搜索引擎利润的平均贡献,断尾回归模型的实证结果则反映的是广告商对搜索引擎利润的无偏估计结果。本文尝试分析线性回归模型的实证结果背后的内在逻辑,在此基础上进一步探索断尾回归模型的实证结果背后的内在逻辑。本文假设广告商*i*参与搜索引擎主持的付费位置拍卖。由于实践中广告商投标金额是动态变化的,广告商可被搜索引擎依据投标金额大小安排不同付费位置,并呈现给消费者(Che et al,2017)。广告商对搜索引擎利润的贡献为如下形式:

$$PC_i = \frac{\sum S_{i1} \cdot p_{i1} + \sum S_{i2} \cdot p_{i2} + \sum S_{i3} \cdot p_{i3} + \sum S_{i4} \cdot p_{i4}}{IM_{i1} + IM_{i2} + IM_{i3} + IM_{i4}} - \frac{\sum S_1 \cdot p_1 + \sum S_2 \cdot p_2 + \sum S_3 \cdot p_3 + \sum S_4 \cdot p_4}{IM_1 + IM_2 + IM_3 + IM_4} \quad (1)$$

其中,等式左边的*PC*表示广告商对搜索引擎利润的贡献;*S*表示占据付费位置的广告商是否为消费者点击,若被消费者点击则数值为1,否则数值为0;*p*表示广告商需向搜索引擎支付的投标费用,即下一个付费位置的投标金额或广告商预设的最低投标金额;*IM*表示广告商在付费位置的出现次数。*i*表示广告商,若未加第一个下角标则表示针对所有广告商。下标中的数字1~4表示所在付费位置的排序,如1表示第一个付费位置。

模型(1)表达的是广告商*i*在所有付费位置上对搜索引擎利润的平均贡献与所有广告商在所有付费位置上的平均贡献的差值。在线性回归模型的实证结果中,变量广告商*i*的回归系数虽反映的是广告商*i*的平均贡献程度,却是一个有偏估计。原因是搜索引擎利润变量指标值在0处的断尾特征十分明显。

鉴于此,本文通过对模型(1)进行调整变换,尝试将断尾特征凸显出来,进一步地实证分析广告商对搜索引擎利润的无偏估计的结果。模型(1)改写为:^①

$$PC_{i1} = PC_i - PC_{i2} \quad (2)$$

$$PC_{i2} = 0 \cdot NPC_i \quad (3)$$

$$PC_{i1} = \frac{\sum p_{i1}(S_{i1} = 1) + \sum p_{i2}(S_{i2} = 1) + \sum p_{i3}(S_{i3} = 1) + \sum p_{i4}(S_{i4} = 1)}{IM_{i1} + IM_{i2} + IM_{i3} + IM_{i4}} - \frac{\sum p_1(S_1 = 1) + \sum p_2(S_2 = 1) + \sum p_3(S_3 = 1) + \sum p_4(S_4 = 1)}{IM_1 + IM_2 + IM_3 + IM_4} \quad (4)$$

$$NPC_i = \frac{\sum p_{i1}(S_{i1} = 0) + \sum p_{i2}(S_{i2} = 0) + \sum p_{i3}(S_{i3} = 0) + \sum p_{i4}(S_{i4} = 0)}{IM_{i1} + IM_{i2} + IM_{i3} + IM_{i4}} - \frac{\sum p_1(S_1 = 0) + \sum p_2(S_2 = 0) + \sum p_3(S_3 = 0) + \sum p_4(S_4 = 0)}{IM_1 + IM_2 + IM_3 + IM_4} \quad (5)$$

模型(2)中,尽管广告商*i*对搜索引擎利润的平均贡献*PC_i*仍保持不变,但被分解成两部分:*PC_{i1}*和*PC_{i2}*。并且,在数值上,*PC_{i1}*=*PC_i*,即*PC_{i2}*=0。在模型(5)中,尽管*NPC_i*可能不等于0,但是当*NPC_i*同0相乘后,两者乘积在数值上便等于0。若采用断尾回归模型,可将搜索引擎利润变量在数值0处断尾现象纳入回归模型。实际上,断尾回归模型通过已观察到的*PC_{i1}*的情况,借助两者分布相似性特征,进而反向推出*NPC_i*的分布情况,从而实现无偏的参数估计。这一估计策略可行的原因

①具体的过程及整体实证框架备案。

为:首先,构成搜索引擎利润变量的两个原始变量——消费者点击行为变量和下一个付费位置广告商投标金额(搜索引擎所设的最低投标金额)变量基本上是相互独立。因为消费者在点击付费位置的广告商时不知道广告商的投标金额。其次,不同广告商间的投标策略在经过多轮重复博弈后趋向同质,其投标金额分布无本质差异(Che et al,2017)。因而,断尾回归结果相较于线性回归模型结果更优。

三、数据来源与变量特征

(一)数据来源

本文所用样本数据来自一家国内搜索引擎企业。在数据取样细节方面,本文所用的消费者层面数据产生步骤为:第一,消费者对智能手机等电子产品在搜索引擎进行检索;第二,搜索引擎平台依据广告商竞价及其他排序条件安排广告商在付费搜索结果列表中的位置;第三,搜索引擎记录下消费者是否点击广告商链接、广告商所占据的付费位置及广告商的投标金额等系列数据。样本数据覆盖时段为2017年6月至9月,共四个月,总计50176次检索,平均每天约有418次检索。由于搜索引擎通过用户IP地址记录用户检索行为,本文不能区分同一IP地址下的多次检索是来自同一用户还是不同用户。

为本文提供样本数据的搜索引擎采用第二价格拍卖规则向广告商出售付费位置。其中,投标金额第1高的广告商占据第1个付费位置,投标金额第2高的广告商占据第2个付费位置,依此类推。搜索引擎仍会对参与付费位置拍卖的广告商设置信息质量门槛,筛选高质量的广告商,避免部分广告商抓住付费位置拍卖规则的漏洞,通过恶意投标获取最高付费位置,骗取消费者点击率。所谓恶意投标是指广告商实施逆向选择,提高投标金额进入付费搜索结果列表乃至获取最高付费位置(Abraham et al,2020)。恶意投标的广告商必将向消费者出售虚假服务或劣质商品,损害消费者福利,进而降低付费位置的消费者点击率,从而降低搜索引擎付费位置的拍卖收入。不仅如此,当广告商被置于不同付费位置时,搜索引擎也将同一付费位置的不同广告获取消费者点击率的效率进行测算,并通过信息质量门槛做进一步筛选。经过搜索引擎的不间断迭代测算,占据付费搜索结果列表的广告商不仅是投标金额较高的群体,也是信息质量较高的群体(Park & Agarwal,2018)。在付费位置数量方面,搜索引擎对付费搜索结果列表设置不超过4个付费位置。当消费者点击付费搜索结果时,广告商需向搜索引擎支付下一个付费位置的投标金额。占据最后一个付费位置的广告商支付搜索引擎所设的拍卖底价。拍卖底价是指搜索引擎针对不同类别关键词设置的最低投标金额。本文通过收集4家主流搜索引擎企业安排付费搜索结果列表的数据,发现样本搜索引擎企业较好地执行了数量上限政策,其付费搜索结果列表的平均长度为3.87,处于中等水平。^①

需要说明的是,本文选择电子类商品的原因在于电子产品类关键词的样本数据有助于与既有研究文献进行直接比较分析。例如,Yang et al(2014)、Chan & Park(2015)均讨论的是数量不限的第二价格拍卖规则下广告商投标竞争对搜索引擎拍卖利润的影响。其中,Yang et al(2014)所用样本数据即为电子产品类关键词拍卖数据;Chan & Park(2015)所用样本数据为服装产品类关键词拍卖数据。因此,本文采用电子产品类关键词拍卖数据讨论在数量有上限时搜索引擎平台排序行为作用机理,可与既有研究文献进行直接比较。不仅如此,所用样本数据的特征也可进行比较分析。Yang et al(2014)所用的样本数据中,部分特定付费位置的支付金额并无明确标记。在本文的样本数据中,搜索引擎对最后一个付费位置的支付金额进行明确设定。电子产品的热门程度适中,搜索引擎为之安排的付费位置数量通常不超过4个。在安排付费搜索结果列表方面,本文采用关键词试验法,收集百度搜索引擎平台所设的付费搜索结果的详细数据。其中,所用的关键词基本库是京东商

^①参考曲创和刘洪波(2017)的搜集策略,本研究团队于2021年6月至8月,在商品关键词检索领域,对4家中文主流搜索引擎安排付费搜索结果的情况进行了统计分析。这与注释关键词试验法所提及数据搜集方法是一致的,但覆盖面存在差异。若有兴趣,可向作者索取数据。

城面向所有消费者提供的商品导航目录中的词汇,总计 11 个大类、85 个小类。统计结果显示,包括智能手机在内的电子产品的系列关键词属于热门程度适中的群体,用于反映搜索引擎排序行为较为合适。包括非电子产品在内的其他系列关键词,要么过于热门,如美妆产品系列关键词,要么过于冷门,如艺术品系列关键词。搜索引擎企业在热门关键词上“打擦边球”的现象比较明显,即通常为热门关键词安排的付费搜索结果超过 4 个。但对于冷门关键词,广告商的投标激励不足,搜索引擎企业为之设置的付费搜索结果往往少于 4 个,甚至不安排付费搜索结果。因此,过于热门或冷门的系列关键词均不能较好地体现搜索引擎排序行为的一般性与普遍性。^①

(二) 付费位置的点击率、投标金额及其对搜索引擎利润的贡献

依据前述分析,本文测算了广告商占据付费位置的次数占比及排名,并按付费位置数量上限的约束,即搜索引擎在其搜索结果页面最多可设置 4 个付费位置,选定加入回归模型的广告商分别为:京东商城(JD)、天猫商城(Tmall)、苏宁易购(Suning)及淘宝网(Taobao)。其中,排名第五的广告商为电商平台“1688.com”,其占比为 1.0%,远低于第四名广告商淘宝网的占比。因而,本文将分别单独报告加入回归模型的广告商,不单独报告包括第五名在内的剩余广告商具体情况。为了行文简洁性,本文通常将前四名广告商称为“大型广告商”,将剩余广告商全部统称为“小型广告商”(SAS)。按搜索引擎排序行为先考虑消费者搜索行为影响、后考虑平台自身利润的分析顺序,本文将分别报告各付费位置的消费者点击率、投标金额及其对搜索引擎利润的贡献。其中,各付费位置及大型广告商的消费者点击率情况见表 1。

通过表 1 可以清晰看出所有广告商均对付费搜索结果列表的第一个位置有着强烈追求。具体地,第一个付费位置的消费者点击率几乎是第二个付费位置的 2 倍多,第二个付费位置则是第三个付费位置的 2 倍多,第三个付费位置则是第四个付费位置的 2 倍多,这显示消费者点击率随付费位置降低呈近似等比数列形式下降。换言之,消费者点击率的变化规律几乎以指数函数递减。这基本上印证了 Feng et al(2007)、Yang et al(2014)关于消费者点击次数同付费位置排名指数递减的理论假设。所有付费位置的平均点击率几乎同第二个付费位置的点击率相等,表明广告商第二个及以上付费位置的竞争尤为激烈。在广告商层面,淘宝网、京东商城、苏宁易购等三家大型广告商的消费者点击率均随着付费位置下降而降低,但下降倍数有所不同。天猫商城的消费者点击率随着付费位置先下降而上升的。并且,在第三、四个付费位置上,天猫商城的出现次数及消费者点击率均高于其他大型广告商。从总体看,仅有大型广告商苏宁易购的消费者点击率低于所有广告商的整体水平,其他三家大型广告商则均高于整体水平。小型广告商平均消费者点击率仅在第三个付费位置上超过对应位置的平均水平,在其他付费位置上则均低于对应位置的平均水平,且整体水平也低于所有广告商的整体水平。

表 1 付费位置的消费者点击率分布特征

| 分类 | 总体 | Taobao | Tmall | JD | Suning | SAS |
|---------|--------|------------------|------------------|------------------|------------------|--------------------|
| 第一个付费位置 | 0.266 | 0.393 (4928) | 0.354 (7424) | 0.398 (7552) | 0.274 (3456) | 0.180 (26816) |
| 第二个付费位置 | 0.122 | 0.161 (3584) | 0.0515 (6208) | 0.162 (8704) | 0.177 (5248) | 0.110 (26432) |
| 第三个付费位置 | 0.0517 | 0.0300 (3200) | 0.0748 (6848) | 0.0393 (7296) | 0.0256 (4992) | 0.0571 (24064) |
| 第四个付费位置 | 0.0207 | 0.0128 (2112) | 0.0938 (5120) | 0.0132 (4864) | 0.0180 (5348) | 0.00700 (22428) |
| 总体 | 0.122 | 0.191 (13824) | 0.154 (25600) | 0.168 (28416) | 0.110 (19044) | 0.0929 (99740) |

注:括号中的数字为广告商在付费位置的出现次数。

^① 本文采用关键词试验法,考察对应搜索引擎设置付费搜索结果列表时所搜集。若读者有兴趣,可向作者索取调研数据。

进一步地,本文报告各付费位置及广告商的投标金额情况,具体见表2。

表2 付费位置的平均投标金额分布特征

| 分类 | 总体 | Taobao | Tmall | JD | Suning | SAS |
|---------|-------------------|-------------------|-------------------|-------------------|--------------------|-------------------|
| 第一个付费位置 | 0.193 (0.0236) | 0.202 (0.0141) | 0.201 (0.0217) | 0.202 (0.0234) | 0.188 (0.0241) | 0.188 (0.0237) |
| 第二个付费位置 | 0.171 (0.0245) | 0.177 (0.0232) | 0.174 (0.0238) | 0.170 (0.0233) | 0.178 (0.0252) | 0.160 (0.0188) |
| 第三个付费位置 | 0.151 (0.0271) | 0.154 (0.0168) | 0.153 (0.0162) | 0.150 (0.0176) | 0.150 (0.0169) | 0.151 (0.0172) |
| 第四个付费位置 | 0.124 (0.0276) | 0.118 (0.0143) | 0.127 (0.0127) | 0.121 (0.0127) | 0.122 (0.00787) | 0.124 (0.0134) |
| 总体 | 0.162 (0.0358) | 0.171 (0.0364) | 0.166 (0.0361) | 0.164 (0.0368) | 0.156 (0.0360) | 0.160 (0.0350) |

注:括号中的数字为投标金额指标的标准差。

在表2中,平均投标金额随着付费位置下降而降低。所有广告商的平均投标金额均随付费位置下降而降低。不仅如此,在每一个付费位置上,广告商之间的平均投标金额差异,要低于该付费位置同下一个付费位置的平均投标金额差异。例如,在第一个付费位置上,广告商之间平均投标金额的差异最大值(0.014)要小于第一个付费位置与第二个付费位置的平均投标金额差异(0.022)。将同一个付费位置上四家大型广告商的平均投标金额同该付费位置的平均投标金额进行比较,可知苏宁易购的平均投标金额在第一个付费位置低于对应位置的平均水平;京东商城的平均投标金额在第二个付费位置低于对应位置的平均水平;京东商城、苏宁易购的平均投标金额在第三个付费位置上均低于对应位置的平均水平;淘宝网、京东商城、苏宁易购的平均投标金额在第四个付费位置低于对应位置平均水平;仅有天猫商城的平均投标金额在所有付费位置上均高于对应位置的平均水平。这可能是天猫商城能够获得第二高出现次数的原因之一。小型广告商的平均投标金额在第三、四个付费位置上与对应位置的平均水平相等;在第一、二个付费位置上则低于对应位置的平均水平。在消费者点击率和广告商投标金额数据支持下,借助搜索引擎采用第二价格拍卖作为付费位置拍卖机制的产业实践,本文测算了各付费位置及广告商对搜索引擎利润的贡献,具体见表3。

表3 付费位置对搜索引擎利润的平均贡献的分布特征

| 分类 | 总体 | Taobao | Tmall | JD | Suning | SAS |
|---------|----------------------|-----------------------|---------------------|-----------------------|-----------------------|----------------------|
| 第一个付费位置 | 0.0492 (0.0826) | 0.0739 (0.0921) | 0.0663 (0.0905) | 0.0745 (0.0929) | 0.0474 (0.0783) | 0.0330 (0.0713) |
| 第二个付费位置 | 0.0175 (0.0499) | 0.0239 (0.0572) | 0.00800 (0.0350) | 0.0182 (0.0472) | 0.0265 (0.0605) | 0.0169 (0.0499) |
| 第三个付费位置 | 0.00655 (0.0300) | 0.00472 (0.0271) | 0.00868 (0.0328) | 0.00444 (0.0244) | 0.00344 (0.0213) | 0.00749 (0.0325) |
| 第四个付费位置 | 0.00103 (0.00711) | 0.000639 (0.00562) | 0.00469 (0.0146) | 0.000658 (0.00570) | 0.000898 (0.00664) | 0.00035 (0.00417) |
| 总体 | 0.0200 (0.0556) | 0.0337 (0.0708) | 0.0244 (0.0610) | 0.0267 (0.0633) | 0.0170 (0.0507) | 0.0152 (0.0494) |

注:括号中的数字为搜索引擎利润指标的标准差。

在表3中,付费位置对搜索引擎利润的贡献呈递减趋势。具体地,淘宝网、京东商城、苏宁易购及其他销售商对搜索引擎利润的平均贡献随着付费位置下降而降低;天猫商城对搜索引擎利润的平均贡献先随付费位置下降先降低后上升再降低,且整体上高于平均水平。大型广告商苏宁易购对搜索引擎利润的贡献平均水平(0.0170)要低于其他大型广告商,但与小型广告商的平均水平(0.0152)相比,仍有一定优势。原因在于虽然苏宁易购的平均投标金额低于小型广告商群体,其对付费位置

消费者点击率的平均贡献,要高于小型广告商群体,投标金额能够被转为搜索引擎利润部分要整体高于小型广告。这或许是苏宁易购能够获取较高出现次数的原因之一。

(三)所有变量的描述性统计结果

在考察各付费位置的点击率、投标金额及其对搜索引擎利润的贡献的基础上,本文对纳入回归模型的所有变量进行了描述性统计,结果见表4。

在表4中,变量消费者点击决策 *click* 为二元变量,指标数值1表示消费者点击了付费搜索结果;指标数值0则表示消费者未点击付费搜索结果。从吸引消费者搜索的角度看,消费者点击决策也是广告商对消费者搜索的吸引效果。变量消费者终止搜索 *stop* 为二元变量,指标数值1表示消费者终止点击后续的付费搜索结果;指标数值0则表示消费者未终止点击后续的付费搜索结果。该变量依据 Park & Agarwal(2018)消费者顺序检索行为实证模型中消费者离开付费搜索结果的行为构建而来。^① 若消费者未进入付费搜索结果列表,由于不能判断所有付费位置的终止结果,本文将所有付费位置的哑变量消费者终止搜索 *stop* 设定为缺失值。其实质就是没有进入付费搜索结果列表,就没有退出付费搜索结果行为。若消费者进入付费搜索结果列表,按照有进入便有退出的常识,第一个付费位置至最后一个付费位置的哑变量消费者终止搜索 *stop* 的数值确定方法为:消费者点击的最后一个付费位置的哑变量消费者终止搜索 *stop* 数值为1,在此之前的付费位置的哑变量消费者终止搜索 *stop* 数值为0;付费搜索结果列表的最后一个付费位置的哑变量消费者终止搜索 *stop* 数值设定为缺失值,因为无论消费者是否点击最后一个付费位置,其搜索行为都将在此终止,哑变量消费者终止搜索 *stop* 的数值设定无实质含义。由于54.0%的消费者未进入付费搜索结果列表,以及所有付费搜索结果列表的最后一个付费位置哑变量消费者终止搜索 *stop* 无实际意义,哑变量消费者终止搜索 *stop* 的观测值为73376,占样本观测值的39.32%。

表4 所有变量的描述性统计结果

| 变量 | 均值 | 中位数 | 标准差 | 最小值 | 最大值 | 观测值 |
|---------------|--------|-------|--------|--------|-------|--------|
| <i>click</i> | 0.122 | 0 | 0.327 | 0 | 1 | 186624 |
| <i>stop</i> | 0.0466 | 0 | 0.0726 | 0 | 1 | 73376 |
| <i>bid</i> | 0.162 | 0.160 | 0.0360 | 0.0800 | 0.230 | 186624 |
| <i>profit</i> | 0.0200 | 0 | 0.0560 | 0 | 0.220 | 186624 |
| <i>Id</i> | 0.640 | 1 | 0.480 | 0 | 1 | 186624 |
| <i>Time</i> | 73.35 | 79 | 45.11 | 1 | 209 | 186624 |
| <i>rank</i> | 2.410 | 3 | 1.09 | 1 | 4 | 186624 |
| <i>Taobao</i> | 0.0740 | 0 | 0.262 | 0 | 1 | 186624 |
| <i>Tmall</i> | 0.137 | 0 | 0.344 | 0 | 1 | 186624 |
| <i>JD</i> | 0.152 | 0 | 0.359 | 0 | 1 | 186624 |
| <i>Suning</i> | 0.102 | 0 | 0.303 | 0 | 1 | 186624 |

在表4中,控制变量付费位置投标金额 *bid* 的最小值为0.08元,高于搜索引擎预设的最低投标的0.05元;最大值为0.23元,反映的是广告商为获得付费位置最高需投标0.23元。不仅如此,依据变量搜索引擎利润 *profit* 最大值为0.22元,可知若按付费位置投标金额 *bid* 降低的最小单位0.01元计算,投标金额为0.23元的付费位置曾被消费者点击。占据该付费位置的广告因此向搜索引擎企业支付下一个付费位置的最大投标金额为0.22元,进而形成搜索引擎企业单次的利润0.22元的最大值。反之,若消费者不点击广告商的网址链接,则搜索引擎企业得不到任何利润。依据变量消费者决策 *click* 的均值为12.2%,可知约87.8%的付费位置未被消费者点击。如此,不难理解

①构建变量消费者终止搜索的理论基础来自 Park & Agarwal(2018)的消费者顺序检索的实证模型。

搜索引擎利润 $profit$ 的均值为 0.02 元,最小值为 0,最大值 0.22 元,中位数为 0,即过半数观测值的指标数值为 0。

在其他控制变量方面,认证等级 Id 和合作时长 $Time$ 两个指标均为反映搜索引擎信誉认证制度的指标。为本文提供样本数据的搜索引擎企业实行的是广告商信誉认证制度,主要有两方面。其一,搜索引擎企业通过对广告商进行实地考察,进行等级认证。认证等级有三层次,分别是, V_1 、 V_2 和 V_3 。其中, V_1 和 V_2 等级为淡绿色标志; V_3 等级为金色标志。鉴于此,本文将广告商认证等级设计为一个哑变量。具体而言,若广告商的认证等级为 V_1 和 V_2 ,则 $Id=0$;若广告商认证等级为 V_3 ,则 $Id=1$ 。从指标均值,不难看出超过 60% 的观测值的认证等级为 V_3 等级。其二,搜索引擎企业对已通过信誉认证的广告商,标注合作时间长度(以月为单位)。例如,合作时长的最大值为 209,表示通过搜索引擎的信誉认证制度,并参与付费位置拍卖活动已有 209 个月但未满;如果超过 209 个月有 1 天则显示为 210 个月。合作时长最小值为 1,显示广告商通过搜索引擎认证,但参与付费位置拍卖活动尚未超过 1 个月。不难看出,两个指标是并行的指标,不宜同时置于同一回归方程进行分析。

变量 $Taobao$ 、 $Tmall$ 、 JD 、 $Suning$ 均为哑变量,分别表示广告商是否为淘宝网、是否为天猫商城、是否为京东商城或是否为苏宁易购。四个变量的均值分别显示,淘宝网的出现次数在所有广告商的出现次数的占比为 7.4%;天猫商城的占比为 13.7%;京东商城的占比为 15.2%;苏宁易购的占比为 10.2%。四家广告商占比之和为 46.3%。另外,观测值 186624 表示在消费者 50176 次检索中,搜索引擎共提供了 186624 个付费搜索结果,平均每次检索呈现 3.72 个搜索结果。其中,呈现 2 个付费搜索结果的检索共有 3776 次,呈现 3 个付费搜索结果的检索共有 6528,呈现 4 个付费搜索结果的检索共有 39872 次。

进一步地,考虑到各变量的构建方式并非一致,本文对所有变量进行了相关性系数检验,检验结果参见表 5。

表 5 所有变量的相关性检验

| 变量 | $click$ | bid | $profit$ | Id | $Time$ | $Taobao$ | $Tmall$ | JD | $Suning$ |
|----------|----------|----------|----------|----------|----------|----------|---------|---------|----------|
| $click$ | 1 | | | | | | | | |
| bid | 0.301** | 1 | | | | | | | |
| $profit$ | 0.956*** | 0.350*** | 1 | | | | | | |
| Id | 0.052* | 0.036* | 0.045* | 1 | | | | | |
| $Time$ | 0.047* | 0.034* | 0.044* | 0.707*** | 1 | | | | |
| $Taobao$ | 0.060*** | 0.069*** | 0.071*** | 0.201** | 0.026* | 1 | | | |
| $Tmall$ | 0.039*** | 0.049*** | 0.033*** | 0.092*** | -0.045** | -0.113** | 1 | | |
| JD | 0.060** | 0.028*** | 0.052*** | 0.300** | 0.572*** | -0.120* | -0.169* | 1 | |
| $Suning$ | -0.012* | -0.056** | -0.017* | 0.246*** | 0.042* | -0.095* | -0.134* | -0.143* | 1 |

注: *、**和***分别表示在 10%、5%和 1%水平上显著。下同。

在表 5 中,各变量的相关性均在可控范围。哑变量 $click$ 同搜索引擎利润变量的相关性系数、哑变量 Id 同变量合作时长 $Time$ 的相关性系数,超过 0.7。其余变量间的相关性系数均在合理区间。变量搜索引擎利润 $profit$ 和哑变量消费者点击付费搜索结果决策 $click$,是不同回归方程的因变量,也不会被置于同一回归方程。

四、搜索引擎排序行为的匹配效应

(一)多层贝叶斯回归模型

依据前文分析,本文将采用多层贝叶斯回归模型,对搜索引擎排序行为的匹配效应进行实证检验。在每个付费位置上单独进行估计,考察广告商对消费者点击率的影响,并不能较好地反映消费

者检索行为的连续性。实际上,广告商同时对消费者点击付费位置和离开付费搜索结果列表均存在影响。因此,广告商对消费者点击付费位置和离开付费搜索结果列表的实证模型分别如下:

$$click_i = \beta_{01} + \beta_{11} Taobao_i + \beta_{21} Tmall_i + \beta_{31} Suning_i + \beta_{41} JD_i + \beta_{51} rank_i + \beta_{61} Id_i + \mu_i \quad (6)$$

$$stop_i = \beta_{02} + \beta_{12} Taobao_i + \beta_{22} Tmall_i + \beta_{32} Suning_i + \beta_{42} JD_i + \beta_{52} rank_i + \beta_{62} Id_i + \epsilon_i \quad (7)$$

在模型(6)中,因变量为消费者点击决策 $click$,反映的是占据付费位置的广告商对消费者吸引效果;在模型(7)中,因变量消费者终止搜索 $stop$ 为本文构造变量,反映的是占据付费位置的广告商对消费者点击后续付费搜索结果的阻止效果。其具体规则前文已做详细叙述。因而,因变量消费者点击决策与因变量消费者终止搜索,共同构成搜索引擎排序行为的匹配效应的两个方面:消费者进入付费搜索结果列表;消费者离开付费搜索结果列表。在搜索引擎企业视角下,期望广告商的吸引点击效应越高,越有利于将广告商投标金额转化成搜索引擎利润;期望广告商的阻止点击效应越低,越有利于消费者点击后续付费搜索结果。相应地,搜索引擎排序行为的匹配效应,将消费者点击效应同消费者终止搜索结果效应置于同一分析框架,进行比较分析并测算两者差值。另外,模型(6)(7)的残差项 μ 和 ϵ 的联合分布服从多层正态分布,即 $[\mu, \epsilon] \sim MVN([0, 0], \Omega)$ 。本文运用多层贝叶斯回归(multilevel Bayesian regression),同时估计模型(6)(7),增强回归结果的可比性,进而对广告商的匹配效应进行全面估计,从而解析搜索引擎最优排序。

(二)匹配效应的实证结果

本文的实证分析分两步展开。第一步,本文测算消费者进入付费搜索结果列表行为的后验均值。消费者检索行为是多项连续决策组成的顺序检索决策,包括消费者是否进入付费搜索结果、消费者是否点击付费搜索结果及消费者终止搜索等三项连续决策。具体地,消费者是否进入付费搜索结果是随机概率事件;进入付费搜索结果列表的消费者,其是否点击付费搜索结果则与付费位置的广告商及付费位置的排名有关。故本文首先测算消费者是否进入付费搜索结果列表的后验均值,考察消费者对付费搜索结果列表的态度(表6)。

表6 消费者对付费搜索结果列表进入行为的后验均值

| 测算结果 | 后验均值 | 后验标准差 | 蒙特卡洛标准误 | 95%置信区间 |
|---------|-------|-------|---------|----------------|
| 消费者进入行为 | 0.459 | 0.002 | 0.000 | [0.455, 0.463] |
| 方差 | 0.248 | 0.002 | 0.000 | [0.245, 0.252] |

在表6中,消费者进入付费搜索结果列表的后验均值为0.459。中国是世界上电子商务活动较为活跃的国家之一。电商企业在没有任何经验基础上创造出诸多消费节,基本上覆盖了消费者的日常生活。^①特别是,在移动终端APP未活跃的时候,消费者通过搜索引擎接口进入购物页面的景象更为活跃。面对如此高涨的消费者热情,搜索引擎企业一方面可引入更多广告商参与付费位置拍卖,另一方面可安排更多付费位置的广告商以承接消费者的注意力。但是受到付费位置数量上限的约束,搜索引擎企业需要最优安排各付费位置的广告商,即对广告商进行最优排序。

第二步,结合拓展后的消费者顺序检索的实证模型,本文采用多层贝叶斯回归模型,考察在数量上限约束下对大型广告商匹配效应的影响,进而探索搜索引擎最优排序,从而为进一步分析排序结果对搜索引擎利润的影响奠定基础(表7)。

^①实际上,我国电商市场几乎每个月都有进行大促销的应景主题。例如,1月份的促销标题是元旦新年;紧接着是中国农历新年和春节,此为商品消费旺季;随后是开学季,学生添置学习各类用品和电子消费品;及至五月底端午节,粽子等各类糕点食品大促销;再有“6·18”年中大促销,半价销售的商品充斥市场;7月、8月则为学生暑假,旅游旺季,消暑降温的商品销售旺季,也被称为“家电促销节”;9月开学季,各类学习用品和电子商品销售旺季(注:学生是电子产品消费主力军);10月则是国庆长假,旅游和消费旺季;11月则有全球购物狂欢节,将一年购物促销活动推向顶峰;最后12月,电商企业利用双12消费节,抓住双十一的“尾巴”,乃全年最后促销。

表 7 搜索引擎排序行为的贝叶斯回归结果

| 变量 | 后验均值 | 后验标准差 | 蒙特卡洛标准误 | 95%置信区间 |
|---------------------|--------|-------|---------|------------------|
| 被解释变量: <i>click</i> | | | | |
| β_{01} | 0.331 | 0.002 | 0.000 | [0.326, 0.335] |
| β_{11} | 0.095 | 0.003 | 0.000 | [0.090, 0.102] |
| β_{21} | 0.052 | 0.002 | 0.000 | [0.047, 0.057] |
| β_{31} | 0.085 | 0.003 | 0.001 | [0.080, 0.090] |
| β_{41} | 0.045 | 0.003 | 0.001 | [0.038, 0.052] |
| β_{51} | -0.108 | 0.001 | 0.000 | [-0.110, -0.105] |
| β_{61} | -0.003 | 0.002 | 0.000 | [-0.008, 0.000] |
| 被解释变量: <i>stop</i> | | | | |
| β_{02} | 0.163 | 0.002 | 0.000 | [0.159, 0.166] |
| β_{12} | 0.058 | 0.003 | 0.000 | [0.053, 0.063] |
| β_{22} | 0.036 | 0.002 | 0.000 | [0.033, 0.040] |
| β_{32} | 0.054 | 0.002 | 0.000 | [0.050, 0.058] |
| β_{42} | 0.059 | 0.003 | 0.001 | [0.053, 0.064] |
| β_{52} | -0.049 | 0.001 | 0.000 | [-0.051, -0.047] |
| β_{62} | 0.005 | 0.002 | 0.000 | [0.001, 0.008] |
| $\Sigma_{1,1}$ | 0.118 | 0.000 | 0.000 | [0.117, 0.119] |
| $\Sigma_{2,1}$ | 0.075 | 0.000 | 0.000 | [0.074, 0.076] |
| $\Sigma_{2,2}$ | 0.082 | 0.000 | 0.000 | [0.082, 0.083] |

根据表 7 的上半部分(被解释变量为 *click*), *Taobao*、*Tmall*、*JD* 及 *Suning* 等大型广告商对消费者点击决策影响的后验均值在 95% 的置信区间中均显著为正。这意味着搜索引擎企业赖以吸引消费者点击决策的大型广告商能较好地承接消费者注意力, 满足消费者信息需求。具体地, 淘宝网对消费者点击决策的吸引效应为 0.095; 天猫商城的吸引效应为 0.052; 京东商城的吸引效应为 0.085; 苏宁易购的吸引效应为 0.045。根据表 7 的下半部分(被解释变量为 *stop*), *Taobao*、*Tmall*、*JD* 及 *Suning* 等大型广告商对消费者终止搜索决策影响的后验均值在 95% 的置信区间中均统计显著且为正值; 这表明搜索引擎企业赖以承接消费者注意力的大型广告商, 实质也会抑制消费者继续点击后续搜索结果。原因在于每一个大型广告商实质都是巨型电商企业, 具备完整的消费者购买服务链条, 可抵御其他电商企业对消费者的吸引作用。淘宝网对消费者后续搜索的终止效应为 0.058; 天猫商城的终止效应为 0.036; 京东商城的终止效应为 0.054; 苏宁易购的终止效应为 0.059。在此基础上, 本文进一步测算搜索引擎排序结果同消费者检索行为的匹配效果, 即搜索引擎排序行为的匹配效应。依据前述分析, 搜索引擎排序行为匹配效应的实质为点击行为的后验均值同终止搜索行为的后验均值差异。具体地, 淘宝网匹配效应为 0.037; 京东商城的匹配效应为 0.031; 天猫商城的匹配效应为 0.016; 苏宁易购的匹配效应为 -0.014。

结合上述分析, 可知付费位置数量有上限时, 依据大型广告商匹配效应差异, 搜索引擎企业给出的最优排序是: 将淘宝网安排在第一个付费位置, 京东商城安排在第二个付费位置, 天猫商城安排在第三个付费位置, 苏宁易购安排在第四个付费位置。此最优排序恰好符合搜索引擎企业对付费位置的数量上限约束的最优反映与决策, 即实现付费位置需求与付费位置供给的竞争均衡。

在其他变量方面, 广告商的位置排名 *rank* 对吸引消费者点击决策没有起到抑制作用, 对消费者终止搜索决策也是抑制作用; 并且前者要大于后者, 显示整体上是广告商的位置排名的匹配效应是负向, 基本符合理论预期。搜索引擎的信誉认证对吸引消费者点击决策没有起到促进作用, 反而有

着抑制作用,基本符合理论预期。具体而言,哑变量 Id 的后验均值为负,表明广告商认证越高降低消费者点击付费搜索结果的概率。相反,在对消费者终止搜索决策方面,哑变量 Id 的后验均值为正,显示广告商的更高信誉等级对消费者终止搜索决策有着促进作用。

五、搜索引擎排序行为的利润效应

(一)实证模型

依据上文,可设置如下模型:

$$profit_i = \delta_0 + \delta_1 Taobao_i + \delta_2 Tmall_i + \delta_3 Suning_i + \delta_4 JD_i + \delta_5 Id_i + \delta_6 Con_i + \epsilon_i \quad (8)$$

其中,因变量搜索引擎利润 $profit$ 为搜索引擎利润的原始状态,包含消费者是否点击付费位置的二值结构。控制变量 Con 分别由两个相互替代的指标——广告商投标金额 bid 和广告商位置排序 $rank$ 构成。由于两个指标间有着较强相关性,两者需分开加入实证模型以便更为清晰地了解其中的影响。不仅如此,两个模型的相互配合增强了实证结果的解释力。基于排序行为匹配效应的分析,可知天猫商城等广告商均被搜索引擎认为能契合消费者顺序检索行为,且对自身利润有着较大影响。因而,探讨天猫商城等广告商对搜索引擎总利润的影响,有着较为重要的意义。

(二)OLS 模型的实证结果:参数的有偏估计

本文采用 OLS 模型对搜索引擎利润进行有偏估计验证实证框架的有偏估计的影响因素,并以此为基础同后续的非偏估计结果进行比较分析(表 8)。

表 8 搜索引擎利润的 OLS 模型回归结果

| 变量 | (1) | (2) | (3) | (4) |
|----------|---------------------|----------------------|------------------------|------------------------|
| | $profit$ | $profit$ | $profit$ | $profit$ |
| $Taobao$ | 0.0185*** (3.40) | 0.0187*** (3.45) | 0.0158*** (3.74) | 0.0134*** (3.21) |
| $Tmall$ | 0.00920** (2.51) | 0.00932*** (2.72) | 0.00914*** (3.01) | 0.00618** (2.17) |
| JD | 0.0114*** (3.06) | 0.0116*** (3.06) | 0.0111*** (3.42) | 0.00972*** (3.03) |
| $Suning$ | 0.00178 (0.50) | 0.00199 (0.56) | 0.00633** (2.21) | 0.00451 (1.43) |
| Id | | -0.000351 (-0.19) | -0.00184 (-1.01) | -0.000945 (-0.51) |
| $rank$ | | | -0.0157*** (-21.13) | |
| bid | | | | 0.534*** (13.99) |
| 常数项 | 0.0152*** (6.72) | 0.0154*** (5.86) | 0.0541*** (12.88) | -0.0698*** (-15.19) |
| 观测值 | 186624 | 186624 | 186624 | 186624 |
| R^2 | 0.121 | 0.122 | 0.125 | 0.125 |

在表 8 中,处于不同付费位置的大型广告商均对搜索引擎总利润做出正向贡献。除处于第四个付费位置的苏宁易购外,处于其他付费位置的大型广告商的正向贡献均显著。第(1)–(4)列中,哑变量 $Taobao$ 、 $Tmall$ 及 JD 的回归系数均统计显著且为正值,表明淘宝网所居的第一个付费位置、天猫商城所居的第二个付费位置及京东商城所居的第二个付费位置均能对搜索引擎利润做出正向贡献。不仅如此,处于第一至三个付费位置的大型广告商的回归系数的数值呈现递减趋势。有趣的是,上文表 3 中淘宝网、京东商城及天猫商城的平均利润贡献呈现也递减趋势。尽管表 8 中哑变量

Suning 的回归系数的统计显著性不够稳定,但其回归系数数值始终为正且低于前三个付费位置的大型广告商的回归系数数值,即苏宁易购所居的第四个付费位置对搜索引擎利润的贡献,始终低于前三个付费位置。因而,整体来看,OLS模型的回归结果虽为有偏估计,仍符合第一至四个付费位置对搜索引擎利润贡献的递减趋势。

在控制变量方面,位置排名及投标金额的回归结果均符合理论预期。具体地,表8的第(3)列中,控制变量位置排名 *rank* 的回归系数统计显著为负,表明随着付费位置下降搜索引擎利润随之下降。在表8的第(4)列中,控制变量投标金额 *bid* 的回归系数统计显著为正,反映了投标金额上升有助于提升搜索引擎利润。并且,投标金额上升形成增量仅有 53.4%可转变成搜索引擎利润。由于付费位置浪费造成投标金额未转化成搜索引擎利润,投标金额对搜索引擎利润转化效率存在低估的倾向。表8的第(2)–(4)列中,广告商认证等级 *Id* 系数为负但不显著。这可能是由 OLS 模型有偏估计造成。

(三)断尾回归模型的实证结果:参数无偏估计

1. 断尾回归模型。依据上文实证框架的分析,若考察广告商对搜索引擎利润真实影响,需要剔除付费位置浪费现象的干扰。为此,本文在模型(8)的基础上实施断尾回归分析,对搜索引擎利润贡献的无偏估计。在正式回归分析前,本文对广告商付费位置无效现象进行描述性统计。结果显示,广告商在吸引消费者点击付费搜索结果方面不太能达到 50%以上,表明广告商浪费付费位置是普遍现象。具体地,付费位置无效比排序分别为:苏宁易购(89.0%)、天猫商城(85.6%)、京东商城(83.2%)以及淘宝网(80.9%)。由于四家广告商占据付费位置为搜索引擎最优排序,将他们付费位置无效现象十分剔除后,可得到付费位置对搜索引擎利润贡献的真实影响。不仅如此,本文通过考察搜索引擎付费位置的利润分布特征,确定采用断尾回归模型,从而增强回归结果的可信性。^① 根据结果,可以明确看到搜索引擎在超过 16 万个付费搜索结果上没有获得利润;仅在 22714 个付费搜索结果上获得利润。换言之,搜索引擎利润在 0~0.05 的区间上出现断层,并且非常明显。

2. 无偏估计的回归结果。鉴于上述分析,本文采用断尾回归模型进行回归分析,考察大型广告商对搜索引擎利润的影响,推断搜索引擎的排序决策,从而解析付费位置数量受限后第二价格拍卖的主要的特征(表9)。

表9 搜索引擎利润的断尾回归结果

| 变量 | (1) | (2) | (3) |
|---------------|------------------------|------------------------|-------------------------|
| | <i>profit</i> | <i>profit</i> | <i>profit</i> |
| <i>Tmall</i> | -0.00354*** (-3.88) | -0.00203*** (-3.41) | -0.000436 (-0.64) |
| <i>JD</i> | -0.00144 (-1.52) | -0.00529*** (-8.56) | -0.00621*** (-8.76) |
| <i>Suning</i> | -0.00593*** (-4.87) | 0.000803 (1.01) | -0.00123 (-1.34) |
| <i>Taobao</i> | 0.0166*** (14.69) | 0.00475*** (6.41) | 0.00676*** (7.98) |
| <i>Id</i> | -0.00639*** (-7.77) | -0.00383*** (-7.13) | -0.0120*** (-19.45) |
| <i>bid</i> | | 1.233*** (172.88) | |
| <i>rank</i> | | | -0.0384*** (-132.35) |
| 常数项 | 0.167*** (281.17) | -0.0693*** (-48.82) | 0.233*** (349.01) |

①篇幅所限,具体结果备索。

| 变量 | (1) | (2) | (3) |
|----------|-----------------------|-----------------------|-----------------------|
| | <i>profit</i> | <i>profit</i> | <i>profit</i> |
| σ | 0.0463*** (210.91) | 0.0302*** (212.24) | 0.0346*** (211.39) |
| 观测值 | 22714 | 22714 | 22714 |

在表 9 中,大型广告商对搜索引擎利润的贡献有着明显异质性,实质则是不同付费位置对搜索引擎利润的贡献具有异质性。在四家大型广告商中,仅有一家大型广告商(即淘宝网)对搜索引擎的利润贡献始终稳定为正向。其实质则是仅有第一个付费位置对搜索引擎利润的贡献始终稳定并且为正值。具体地,哑变量 *Taobao* 的断尾回归系数统计显著为,表明淘宝网所居的第一个付费位置对搜索引擎利润的贡献为正值。不仅如此,本文选择表 9 的第(1)列的回归结果为基准来呈现搜索引擎排序行为的利润效应。这是因为回归结果中的四大广告商对搜索引擎利润的影响趋势与表 3 中大型广告商平均利润贡献分布特征相契合。相应地,在表 9 的第(1)列中,淘宝网在剔除付费位置无限的情形后,仍可对搜索引擎的利润做出正向贡献,数值约为 0.0166 元。由于该搜索引擎被检索的次数每日超过 10 亿人次,淘宝网利润贡献累计效应较大。

相反,其他大型广告商对搜索引擎利润贡献系数为负,表明其他付费位置对搜索引擎利润的贡献为负值。在表 9 的第(1)(2)列中,哑变量 *Tmall* 的断尾回归系数显著为负,表明大型广告商天猫商城所居的第三个付费位置对搜索引擎利润的贡献为负值。在表 9 的第(2)(3)列中,哑变量 *JD* 的断尾回归系数显著为负,表明大型广告商京东商城所居的第二个付费位置对搜索引擎利润的贡献为负值。在表 9 第(1)列中,哑变量 *Suning* 的断尾回归系数统计显著且为负值,表明大型广告商苏宁易购所居的第四个付费位置,对搜索引擎利润的贡献为负。尽管由断尾回归分析而来的无偏估计结果并非搜索引擎所愿意看见的结果,但却是必须要接受的利润效应的结果。同表 8 中 OLS 模型回归结果相比,即便是最优排序下的大型广告商,对有限付费位置的浪费效应也较为明显。总结上述分析并以第(1)列为基准回归结果,可明确在付费位置数量有上限时,各个付费位置对搜索引擎的利润贡献,即搜索引擎排序行为的利润效应。具体地,第一个付费位置的利润效应为 0.0166,第二个付费位置的利润效应为-0.00144,第三个付费位置的利润效应为-0.00354,第四个付费位置的利润效应为-0.00593。

在其他控制变量方面,投标金额和位置排名两个变量都对搜索引擎利润贡献有着显著影响,与表 8 中的有偏估计基本相同,但具体定量结果存在一定差异。在表 9 第(2)列中,变量 *bid* 的断尾回归系数显著为正,且回归系数高达 1.233,意味着投标金额向搜索引擎利润的转化效率高达 123.3%。这一数值远高于表 8 中的有偏估计数值,反映出搜索引擎企业对广告商投标金额的转化效率比较高。值得关注的是,在无偏估计下,搜索引擎实名认证制度对搜索引擎利润贡献有着负向影响。这一结果与有偏估计存在显著差异。表 9 第(1)-(3)列中,哑变量 *Id* 的断尾回归系数统计显著且为负值,表明提升广告商的认证等级反而降低搜索引擎利润。不仅如此,这一抑制作用非常稳健且对搜索引擎企业每获取 1 元利润的影响甚至超过 1%。结合表 7 结果,认证等级对消费者点击决策存在负向影响。综合而言,信誉认证制度未能给搜索引擎企业带来收益。其内在机理在于,信誉认证不能推动消费者点击付费搜索结果,不能促使广告商的投标金额转化成搜索引擎利润。换言之,信誉认证制度无效。这意味搜索引擎必须对其信誉认证制度进行改革,避免资源浪费。

六、匹配效应与利润效应不一致性分析

(一)两者不一致性的具体特征

本文以不同付费位置的匹配效应、利润效应的量化结果为数据基础,分别绘制不同付费位置的匹配效应与利润效应的变化趋势图,考察匹配效应与利润效应不一致性的具体特征,进而探究实证

结果与理论结论间的差异,从而提出产业规制政策进一步调整的方向。具体见图1和图2。需要指出的是,本文对匹配效应与利润效应的不一致性分析并非比较匹配效应与利润效应的绝对值差异,而是分析匹配效应与利润效应在付费位置数量有上限时变化趋势的质性差异。

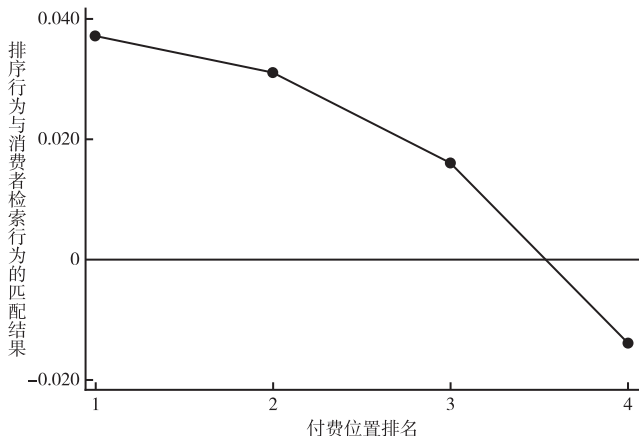


图1 搜索引擎排序行为的匹配效应

注:付费位置排名中占据不同付费位置的广告商排序相一致。下同。

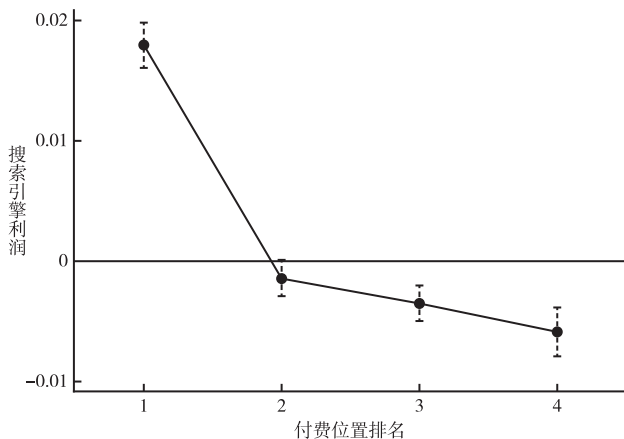


图2 搜索引擎排序行为的利润效应

比较图1和图2的趋势差异,可以发现两者的不一致性大致有两点具体特征:

第一,匹配效应优于利润效应。在图1中,仅有第四个付费位置的匹配效应为负值,其他付费位置的匹配效应均为正值。在图2中,仅有第一个付费位置的利润效应为正值,其他付费位置的利润效应均为负值。付费位置数量有上限时搜索引擎排序行为的匹配效应达到次优状态(顺数第二优状态);搜索引擎排序行为的利润效应达到倒数第二优状态。因而,付费位置数量有上限时,搜索引擎排序行为的匹配效应要优于利润效应。这意味着搜索引擎企业在排序时将会优先考虑排序行为的匹配效应。

第二,匹配效应与利润效应的递减趋势不一致。在图1中,匹配效应随着付费位置降低呈现凸型函数递减。在图2中,利润效应随着付费位置降低呈现凹型函数递减。无论付费位置数量有上限还是不受限制,既有文献均假定消费者点击付费位置概率随着付费位置降低呈指数函数递减趋势,其理论结论则是匹配效应与利润效应均呈现指数函数递减趋势。从理论层面看,匹配效应与利润效应均可放松至呈现凹型函数递减趋势。实践中,付费位置数量有上限时,利润效应呈现凹型函数递减趋势,即随着付费位置降低递减速度呈现下降趋势;匹配效应呈现凸型函数递减趋势,即随着付费位置降低递减速度呈现上升趋势。尽管两者递减趋势差异并非付费位置数量有上限时所独有的特

征,但对付费位置数量设置上限的产业实践会表现出两者递减趋势的不一致性。

总之,当匹配效应与利润效应呈现不一致性时,搜索引擎企业对是否维持数量上限约束进行权衡取舍。一方面,搜索引擎企业可能冒着被处罚的风险在热门关键词拍卖市场,放松付费位置数量上限约束,延长付费搜索结果列表,进而驱动利润效应随匹配效应同步上升,从而增加付费位置拍卖收入。另一方面,搜索引擎企业同样会冒着被处罚风险,提升热门关键词拍卖市场最低投标金额,提升广告商的投标金额并刻意缩短付费搜索结果列表,进而在同等消费者点击率的情形下提升市场竞争强度,使得低质量广告商混入其中,增加消费者购买到劣质产品的可能性。无论延长还是过度缩短付费搜索结果列表的举措,都将违背反垄断规制机构的制定政策意愿,使得规制政策难以达到规制效果,从而损害消费者福利。

(二)对利润效应的倒数第二优状态的讨论

本部分尝试考察搜索引擎所设拍卖底价的最优性,探索数量有上限时搜索引擎排序行为匹配效应与利润效应不一致性的内在原因。根据已有研究,当拍卖底价达到最优时,搜索引擎利润与消费者福利一致地达到最优状态,即两者具有一致性(Athey & Ellison, 2011)。反之,若两者呈现不一致性时,需否定搜索引擎所设的拍卖底价为最优拍卖底价的默认前提。本文在运用此逻辑时要求实证模型与经典文献的理论假设高度一致。沿此研究思路,结合经典文献最优拍卖底价的形成规则(Athey & Ellison, 2011),本文认为:第一,付费位置的平均投标金额反映的是消费者对付费位置的点击概率;第二,将平均投标金额与消费者点击概率进行加权求和(即求解平均投标金额的期望值),得出搜索引擎平台可设定的最低投标金额,可初步测算模拟最优拍卖底价(以下简称“模拟拍卖底价”)。即 $r_{mic} = \sum \overline{bid}_i CR_i$ 。其中, \overline{bid}_i 为付费位置*i*的平均投标金额, CR_i 为付费位置*i*的消费者点击率,可视为样本中“消费者点击概率”。将表1和表2对应的指标数值代入式中,可得到 $r_{mic} = 0.08$ 元。这一模拟拍卖底价较之搜索引擎所设的拍卖底价(0.05元)有一定幅度提升。

本文进一步将模拟拍卖底价代入样本数据,测算不同付费位置的搜索引擎利润,并对不同付费位置的利润效应进行回归分析,从而检验利润效应的模拟效果,最终绘制了不同付费位置利润效应的异质性图谱(图3)。

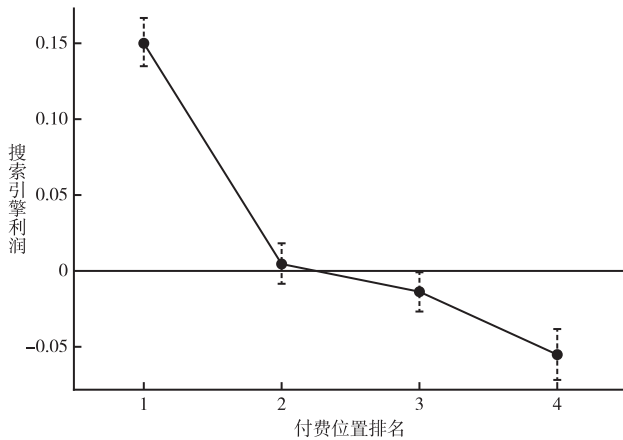


图3 采用模拟拍卖底价的付费位置利润效应异质性

采用模拟拍卖底价后,不同付费位置利润效应异质性的基本趋势保持不变,但其中一些细节有一定变化。基本趋势方面,最优排序下不同付费位置利润效应呈逐层下降趋势。细节方面的变化为:第一,最优排序下付费位置利润效应变化区间由图2的(-0.005, 0.020)缩小为图3的(-0.005, 0.015),这表明不同付费位置的利润效应差异在缩小。第二,后三个位置中有一个位置能够对搜索引擎利润做出微弱的正向贡献。第三,不同付费位置利润效应下降趋势不是始终维持凹型下降,而是在第二至第四个付费位置的下降过程呈现一定轻微的凸型。这一凸型下降趋势与图1的匹配效应凸型下降趋势吻合。综上,可知在适度提高拍卖底价后,搜索引擎利润效应趋势向匹配效

应趋势接近,存在实现两者一致性的可能性。尽管如此,搜索引擎利润在较低付费位置上仍维持较低水平。这反映出两者不一致性现象较为稳健,继续提高拍卖底价的可能性较小。^①

七、结论与建议

已有文献认为,在付费位置数量上限的情形下,搜索引擎排序行为的匹配效应与利润效应随着付费位置降低,均呈现凹型下降且为正值。本文基于经典理论范式建模,结合中国产业实践中商品关键词检索领域的消费者层面数据,考察在付费位置数量有上限时搜索引擎排序行为的匹配效应与利润效应,发现搜索引擎排序行为的匹配效应与利润效应均难以实现经典理论预判。具体地,搜索引擎排序行为的匹配效应呈现凸型下降,并在最低付费位置呈现负值,利润效应虽凹型下降,仅在最高付费位置呈现正值。两者不一致性根源于付费位置竞价的均衡拍卖底价不能等效地改善匹配效应与利润效应。提升均衡拍卖底价可遴选出更高质量广告商,与较高付费位置进行匹配,推动匹配效应快速下降却维持较高水平,但降低了低付费位置对关键词拍卖利润的汲取力度,使利润效应在较低付费位置呈现缓慢下降态势并维持较低水平。

本文的政策建议如下:

第一,反垄断规制部门可以与搜索引擎企业建立协同机制。付费位置数量有上限时,搜索引擎排序行为的较高水平匹配效应激励搜索引擎企业在电子商品关键词拍卖上安排较多付费搜索结果(如3个),以契合消费者检索意愿,从而提升消费者信息检索效率。但较低的利润效应却又推动搜索引擎企业安排较少的付费搜索结果(如1个),以最优地汲取付费位置的拍卖利润。并且无论较长或较短的付费搜索结果列表均在付费位置数量有上限约束范围内具有较好的灵活性。但在其他商品或服务关键词拍卖上,搜索引擎排序行为的匹配效应优于利润效应,意味着搜索引擎企业以契合消费者检索意愿为理由随意增加付费搜索结果数量,进而突破规制政策约束并形成较差的示范效应,侵蚀产业规制政策约束边界。本文建议政府规制部门与搜索引擎企业就关键词搜索热度进行数据信息互通,按关键词特征设置相应的付费位置上限,增强产业规制政策有效性和可行性。从实际操作层面来看,全社会商品关键词类型有85个大类(蔡祖国、李世杰,2020),搜索引擎企业完全可以在政府规制部门指导下构建动态更新的搜索热度目录。

第二,在维持付费位置数量有上限的背景下,搜索引擎企业应对自身设置拍卖底价是否最优进行测算,增强匹配效应与利润效应的一致性。在商品关键词检索领域,通过调整拍卖底价,虽不能完全实现匹配效应与利润效应的一致性,却可以降低匹配效应与利润效应的内在差异性。例如,更高的拍卖底价可使搜索引擎排序行为的利润效应,由“随付费位置降低凹型降低”转变为“随付费位置降低凸型降低”,进而接近匹配效应的变化趋势,增强搜索引擎企业持续接受付费位置数量上限约束的内在激励。对搜索引擎企业而言,调整拍卖底价意味付费位置拍卖收入将呈现较大变化。搜索引擎企业可通过实验方法,在某一时间段或某一关键词领域上调整拍卖底价,测试拍卖底价对拍卖利润及匹配效应的影响,进而使利润效应与匹配效应的变化趋势同向并进。

需要指出的是,本文着重考虑付费位置数量有上限时的商品关键词检索领域搜索引擎排序行为。研究结论适用于市场化特征显著的商品关键词领域,不能简单地推广至知识类、常识类等关键词领域。在知识类、常识类等关键词领域,更应聚焦搜索引擎平台将存在利益关联的内容提供商置于优先排序位置的行为规制等问题,并讨论搜索引擎非中立排序行为引致的市场圈定效应。这方面研究所采用的词条样本数据多为最流行的500个词汇等知识类、常识类关键词(曲创、刘洪波,

^①事实上,在将拍卖底价模拟为0.08元的过程,本文去掉投标金额等于0.08元的付费位置的观测值(2374),造成样本容量损失1.27%,这一微小损失样本容量损失对匹配效应测算的影响可忽略不计。但当继续提升拍卖底价至0.10元时,样本容量损失将达到7.85%;进一步地,拍卖底价升至0.11元时,样本容量损失将达到惊人的11.12%。样本容量过多损失,将直接损害匹配效应先前测算结果的有效性,迫使研究者重新测算匹配效应。匹配效应与利润效应的不一致性将更难分析与估算。频繁地测算匹配效应与利润效应一致性,实质是在挖掘最优拍卖底价。

2017)。学界普遍认为,搜索引擎不应当为知识类或常识类关键词检索等消费者日常信息检索安排付费搜索结果列表,而应当直接将自然搜索结果列表置于显著位置,以提升消费者检索效率,凸显信息检索功能的公益性(李俊生、姚东旻,2016)。如果在知识类、常识类关键词检索领域安排付费搜索结果列表,搜索引擎往往借助信息不对称条件,骗取消费者点击率。政府规制部门应当厘清商品关键词检索与消费者日常信息检索之间的界限,明确将付费位置拍卖机制严格限定于商品关键词检索领域。

参考文献:

- 蔡祖国 李世杰,2020:《产品质量信号能提升竞价排名机制的信息匹配效率吗——基于中国搜索服务市场的分析》,《中国工业经济》第10期。
- 李俊生 姚东旻,2016:《互联网搜索服务的性质与其市场供给方式初探——基于新市场财政学的分析》,《管理世界》第8期。
- 曲创 刘重阳,2019:《平台竞争一定能提高信息匹配效率吗?——基于中国搜索引擎市场的分析》,《经济研究》第8期。
- 曲创 刘洪波,2017:《平台非中立性策略的圈定效应——基于搜索引擎市场的试验研究》,《经济学动态》第1期。
- Abhishek, V. et al(2015), “Aggregation bias in sponsored search data: The curse and the cure”, *Marketing Science* 34(1):59—77.
- Abraham, I. et al(2020), “Peaches, lemons, and cookies: Designing auction markets with dispersed information”, *Games and Economic Behavior* 124:454—477.
- Agarwal, A. et al(2011), “Location, location, location: An analysis of profitability of position in online advertising markets”, *Journal of Marketing Research* 48(6):1057—1073.
- Amaldoss, W. et al(2016), “Keyword management costs and ‘broad match’ in sponsored search advertising”, *Marketing Science* 35(2):259—274.
- Athey, S. & G. Ellison(2011), “Position auction and consumer search”, *Quarterly Journal of Economics* 126(3):1213—1270.
- Blake, T. et al(2015), “Consumer heterogeneity and paid search effectiveness: A large-scale field experiment”, *Econometrica* 83(1):155—174.
- Chan, T. Y. & Y. H. Park(2015), “Consumer search activities and the value of ad positions in sponsored search advertising”, *Marketing Science* 34(4):606—623.
- Che, Y. K. et al(2017), “An experimental study of sponsored-search auctions”, *Games and Economic Behavior* 102:20—43.
- Chen, Y. & C. He(2011), “Paid placement: Advertising and search on the Internet”, *Economic Journal* 121(556):309—328.
- Chen, Y. & S. Yao(2017), “Sequential search with refinement: Model and application with click-stream data”, *Management Science* 63(12):4345—4365.
- Chen, Y.-J. (2021), “Optimal design of revenue-maxing position auctions with consumer search”, *Production and Operations Management* 30(9):3297—3316.
- Chu, Y. L. et al(2020), “Position ranking and auctions for online marketplace”, *Management Science* 66(8):3617—3634.
- Edelman, B. et al(2007), “Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords”, *American Economic Review* 97(1):242—259.
- Feng, J. et al(2007), “Implementing sponsored search in web search engines: Computational evaluation of alternative mechanism”, *INFORMS Journal on Computing* 19(1):137—148.
- Ghose, A. & S. Yang(2009), “An empirical analysis of search engine advertising: Sponsored search in electronic markets”, *Management Science* 55(10):1605—1622.
- Katona, Z. & M. Sarvary(2010), “The race for sponsored links: Bidding patterns for search advertising”, *Marketing Science* 29(2):199—215.
- Kim, J. B. et al(2010), “Online demand under limited consumer search”, *Marketing Science* 29(6):1001—1023.
- Jerath, K. et al(2011), “A ‘position paradox’ in sponsored search auctions”, *Marketing Science* 30(4):612—627.
- Park, C. H. & M. J. Agarwal(2018), “The order effect of advertisers on consumer search behavior in sponsored

- search markets”, *Journal of Business Research* (84):24–33.
- Simonov, A. et al(2018), “Competition and crowd-out for brand keywords in sponsored search”, *Marketing Science* 37(2):200–215.
- Varian, H. R. (2007), “Position auctions”, *International Journal of Industrial Organization* 25(6):1163–1178.
- White, A. (2013), “Search engine: Left side quality versus right side profits”, *International Journal of Industrial Organization* 31(6): 690–701.
- Yang, S. & A. Ghose(2010), “Analysis the relationship between organic and sponsored search advertising: Positive, negative, or zero interdependence”, *Marketing Science* 29(1):1–22.
- Yang, S. et al(2014), “Modeling competition and its impact on paid-search advertising”, *Marketing Science* 33(1): 134–153.

Upper Limit of Quantity, Sponsored Search Auction and Ranking Behavior of Search Engines

LI Shijie

(Hainan University, Haikou, China)

Abstract: The sponsored search auction is the main source of search engine revenue and advertisers' bidding benefits. Due to the requirements of the government's industry regulation departments, search engines impose restrictions on the upper limit of the quantity of the sponsored positions, making it different from the common design of the sponsored positions with variable number. Little attention has been paid to the influence mechanism of auction design with limited quantity of sponsored positions on search engines and advertisers. By combing with the practice of Chinese search engines and using multi-layer Bayesian regression and other regression analysis, this paper constructs an empirical framework suitable for consumer retrieval behavior to investigate the influence mechanism of the sponsored auction with limited quantity of positions and explore the actual effect of industry regulation with that limit. The study found that the auction design with limited positions is equivalent to truncating the list of paid search results, which makes it easier for consumers to slide out of the paid search results area and makes it more difficult for search engines to obtain the click rate from consumers. To reduce the risk, on the basis of meeting the bidding rank rules, search engines need to give priority to the ranking rules considering the preferences of consumers to maximize the click rate from consumers on paid search results and increase the auction revenue of sponsored positions. However, under the best bidding rank rules, the contribution of advertisers to search engine excess profits is heterogeneous, meaning a small number of advertisers have a positive contribution to search engine excess profits, while most advertisers contribute negatively. It means that the limit of sponsored positions contributes to consumer welfare rather than search engine revenue. In essence, the limited quantity of positions provides a better constraint for search engines to improve information matching efficiency.

Keywords: Search Engine's Bidding Behavior; Sponsored Search Auction; Upper Limit of Quantity

(责任编辑:何伟)

(校对:陈建青)

信念与人力资本投资研究进展*

罗卫东 袁月美 许彬

摘要:经典人力资本投资理论认为,决策者根据成本和收益寻找最优人力资本投资水平。但是,现实中人力资本投资收益存在不确定性,个体只能根据对成本和收益的主观判断做出决策。近年来,信念在人力资本投资决策研究领域得到了广泛关注。本文系统梳理了人力资本投资研究中关于信念的研究进展,包括教育决策中信念的分类与衡量;父母和学生的信念对于人力资本投资决策的影响;信念偏差的表现、起因和影响;信念纠偏的信息干预途径及其影响等。本文的梳理有助于深化对人力资本积累理论的理解,同时也为优化家庭人力资本投资提供了重要的经验依据和政策启示。

关键词:信念 人力资本 教育 家庭 不确定性

人力资本在经济增长中发挥着重要作用。人力资本投资决策既关系到个人发展,又关系到国家发展。偏离人力资本投资最优水平,就会造成社会资源的巨大浪费。当前我国教育领域中过度投资和投资不足现象并存,个体焦虑和社会效率损失并存,改进人力资本投资决策不仅是家庭面临的重要课题,也是国家应予重视的重大现实问题。

在人力资本投资模型中,人们根据成本与收益来寻找最优投资水平(Becker,1964)。但是,现实中人力资本投资收益是不确定的,个体做出投资决策时对自身禀赋或能力水平也往往不具备完全信息,实际决策一般依赖于对这些具有不确定性结果或不完全信息的状态的主观判断。这些主观判断即为信念(belief)。近三十年来,关于信念的研究有了长足的发展,但目前鲜有文献对这些新兴研究做出系统梳理。本文考察人力资本研究领域关于信念的新进展,从信念在决策中的角色与重要性、信念偏差形成原因与其社会经济后果、纠正信念偏差的信息干预措施及其效果等方面系统梳理这一研究领域的成果。通过系统归纳与讨论,本文试图呈现教育决策中的主观信念的研究方法、逻辑和结果的主要面貌,以期为今后我国的人力资本研究和实际决策提供参考。

一、信念的含义、分类与衡量

(一)信念的含义与分类

信念是对不确定事件发生的可能性的估计、对现实状态的预期或主观评价,常与主观概率(subjective probability)或主观期望(subjective expectation)通用。不同于其他文化信念或与社会规范有关的道德信念,本文讨论的信念可与真实的结果或状态进行对比,从而可被判断为准确或有偏差。

在涉及教育投资的人力资本研究中,根据内容可将信念划分为两大类型:关于未来结果的信念和关于现实状态的信念。前者包括学生选择专业时关于各专业的未来收入(Arcidiacono et al,2020)、

* 罗卫东,浙江大学经济学院、浙江大学城市学院,邮政编码:310058,电子邮箱:luo.wd@zju.edu.cn;袁月美,浙江大学经济学院,邮政编码:310058,电子邮箱:ymyuan@zju.edu.cn;许彬,浙江工商大学经济学院、浙江工商大学现代商贸研究中心,邮政编码:310018,电子邮箱:xubin211@mail.zjgsu.edu.cn。基金项目:浙江省自然科学基金项目“农村地区稳定脱贫的路径选择:行为机理探索与助推政策设计”(LY18G030012);教育部人文社会科学研究规划基金项目“稳定脱贫长效机制的行为路径研究”(18YJA790092);中央高校基本科研业务费专项资金“偏好的性别差异起源:父母性别偏好与抚养方式”(2020QNA260)。感谢审稿专家的修改意见,文责自负。