

2022年12月14日

基于搜索指数还原疫情水平——宏观量化策论之一

宏观研究团队

——宏观经济点评

何宁（分析师）

hening@kysec.cn

证书编号：S0790522110002

● 防疫持续优化

11月以来我国防疫优化稳步迈进新阶段。11月11日“二十条”出台后，各地“因城施策”边际优化防疫政策，截至12月9日我们统计已有19个省市落实“新十条”。在“缩小核酸检测范围、减少频次”、“非必要不核酸”等导向下，以及多个省市取消常态化核酸，如何更好跟踪各地疫情水平以及判断第一轮冲击达峰情况？

● 搜索指数映射疫情水平

表观疫情数据的传导链条为新冠→居民→卫健委。居民在感染或认为感染概率较高时，其可能发生的行为包括搜索新冠本体、新冠体征以及购买药品等。若能有效观测到前半段传导，理论上或可领先于卫健委的表观疫情数据。

陈沁博士在《各城市首轮感染高峰期预测》提出以发烧指数衡量疫情感染情况，我们将其扩展为三类：搜索新冠本体、新冠体征以及购买药品，具体为“新冠”、“发烧”、“莲花清瘟”、“布洛芬”等关键词。

从与新增确诊的相关性来看，药品搜索指数>疫情搜索指数>体征搜索指数。有趣的是，“莲花清瘟”在三类关键词中相关性最好，附带的“莲花清瘟”同样具备映射疫情水平的作用。“二十条”发布后，莲花清瘟指数迅速上升而后下行，与疫情数据有小幅背离，在于防疫优化推进时药品包含了对未来疫情扩散的预期、因而提前上升；

发烧指数在中国的有效性似乎并没有那么高，这可能与我国奥密克戎变异路径有关。2022年上半年主要是ba.2和ba.2.2变异株、下半年为BA.5和BF.7，前两者感染后临床表现相对温和、无症状感染者占比高，但BF.7感染症状包括高烧、持续性咳嗽、身体疼痛不适、头痛、喉咙痛等，与新冠奥密克戎早期亚型变异株感染后的症状类似。因此，高烧指数无法描绘前三个季度的疫情水平，但11月以来有效性显著提升。

搜索指数具备高频、领先性、抗干扰等特点。我们将时间窗口选择为2022年1月1日-2022年12月7日，并计算各搜索指数与确诊病例的相关性。从全国来看，莲花清瘟指数、新冠指数、发烧指数与确诊病例的相关性分别为0.70、0.63、0.03；莲花清瘟指数的领先期主要分布在5-10日、平均为7.1日，新冠指数的领先期主要分布在2-8日、平均为6.1日。

● 利用搜索指数拟合疫情水平

“小步走、不停步”的防疫优化已初步迎来第一波冲击，我们尝试利用搜索指数拟合疫情水平。回归统计数据显示，20个地区调整R方超过0.5，表明该方法具有一定解释力度。

从时间序列来看，本轮疫情始于11月初，河南、河北、内蒙古一带较快蔓延，11月初以广州多区宣布解除疫情防控临时管控区为标志、广东主动迎接第一波冲击，12月初北京被动进入第一波冲击。目前上海、吉林、黑龙江等地疫情可能以较快的速度扩散。

● **风险提示：**国内疫情反复超预期，搜索指数有效性下降。

相关研究报告

《通胀连续下行，美联储内部可能分化——美国11月CPI点评—宏观经济点评》-2022.12.14

《杠杆仍加在企业端——11月社融数据点评—宏观经济点评》-2022.12.12

《后疫情时代来临，突出做好稳增长—宏观周报》-2022.12.11

目 录

1、 防疫持续优化	3
2、 搜索指数映射疫情水平	3
2.1、 药品搜索指数最为有效	3
2.2、 区域搜索指数具有高频、领先、抗干扰特征	4
3、 利用搜索指数拟合疫情水平	6
4、 风险提示	7

图表目录

图 1： 各地积极落实“新十条”	3
图 2： 药品搜索指数较为有效	4
图 3： 新冠搜索指数有一定相关性	4
图 4： 发烧搜索指数从 11 月开始生效	4
图 5： “连花清瘟”搜索指数领先确诊病例 5-10 日	5
图 6： “新冠”搜索指数领先确诊病例 2-8 日	6
图 7： 连花清瘟指数与疫情节点较为吻合	6
图 8： 各地搜索指数与确诊病例拟合效果较好	7
图 9： 基于百度指数拟合日均确诊病例	7
表 1： 搜索指数与确诊病例相关性	5

1、防疫持续优化

11月以来我国防疫优化稳步迈进新阶段。11月11日“二十条”出台后，各地“因城施策”边际优化防疫政策。12月7日卫健委发布“新十条”，截至12月9日我们统计已有19个省市落实“新十条”。在“缩小核酸检测范围、减少频次”、“非必要不核酸”等导向下，以及多个省市取消常态化核酸，如何更好跟踪各地疫情水平以及判断第一轮冲击的达峰情况？

图1：各地积极落实“新十条”



数据来源：央视新闻、各地人民政府网、开源证券研究所

2、搜索指数映射疫情水平

表现疫情数据的传导链条为新冠→居民→卫健委，我们不妨针对前半段传导做文章。居民在感染或认为感染概率较高时，其可能发生的行为包括搜索新冠本体、新冠体征以及购买药品等。若能有效观测到前半段传导，理论上或可领先于卫健委的表现疫情数据。

2020年以来多篇学术论文基于谷歌趋势分析工具测算新冠疫情病例，例如2022年9月15日发表于《政策研究与战略管理》的一篇论文¹指出，“嗅觉丧失”、“味觉丧失”、“发烧”（按数量级计算）与官方病例显著相关。《中国-世卫组织新型冠状病毒肺炎（COVID-19）联合考察报告》对中国55924例实验室确诊病例的统计²显示，发烧、干咳、疲劳、咳痰、呼吸急促为前五大症状，出现比例分别为87.9%、67.7%、38.1%、33.4%、18.6%。因而“发烧”可能是居民搜索新冠体征的常用项。

2.1、药品搜索指数最为有效

陈沁博士在《各城市首轮感染高峰期预测》提出以发烧指数衡量疫情感染情况³，

¹ Lateef, B.A., Hossana, T., and Chinedu, W.O. (2022). Modeling COVID-19 incidence with Google Trends. Research Policy and Strategic Management.

² <http://www.nhc.gov.cn/xcs/yqfkdt/202002/87fd92510d094e4b9bad597608f5cc2c/files/fa3ab9461d0540c294b9982ac22af64d.pdf>

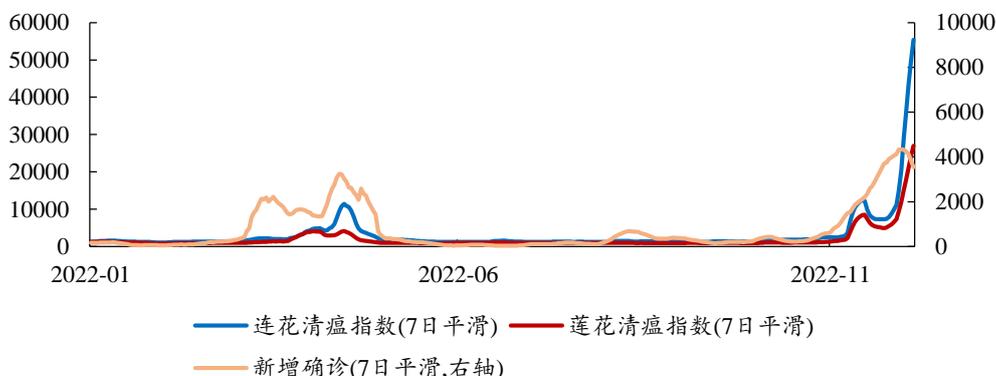
³ https://zhuanlan.zhihu.com/p/590989182?utm_medium=social&utm_oi=556272282197110784&utm_psn=1585332724964745217&utm_source=

请务必参阅正文后面的信息披露和法律声明

我们将其扩展为三类：搜索新冠本体、新冠体征以及购买药品，具体为“新冠”、“发烧”、“莲花清瘟”、“布洛芬”等关键词。

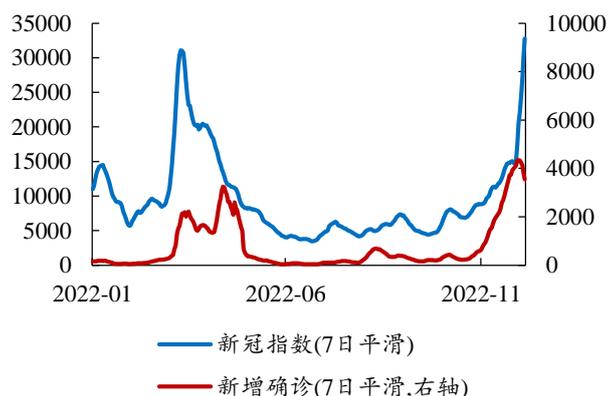
从与新增确诊的相关性来看，药品搜索指数>疫情搜索指数>体征搜索指数。有趣的是，“莲花清瘟”在三类关键词中相关性最好，附带的“莲花清瘟”同样具备映射疫情水平的作用。“二十条”发布后，莲花清瘟指数迅速上升而后下行，与疫情数据有小幅背离，在于防疫优化推进时药品包含了对未来疫情扩散的预期、因而提前上升；发烧指数在中国的有效性似乎并没有那么高，这可能与我国奥密克戎变异路径有关。2022年上半年主要是ba.2和ba.2.2变异株、下半年为BA.5和BF.7，前两者感染后临床表现相对温和、无症状感染者占比高，但BF.7感染症状包括高烧、持续性咳嗽、身体疼痛不适、头痛、喉咙痛等，与新冠奥密克戎早期亚型变异株感染后的症状类似。因此，高烧指数无法描绘前三个季度的疫情水平，但11月以来有效性显著提升。

图2：药品搜索指数较为有效



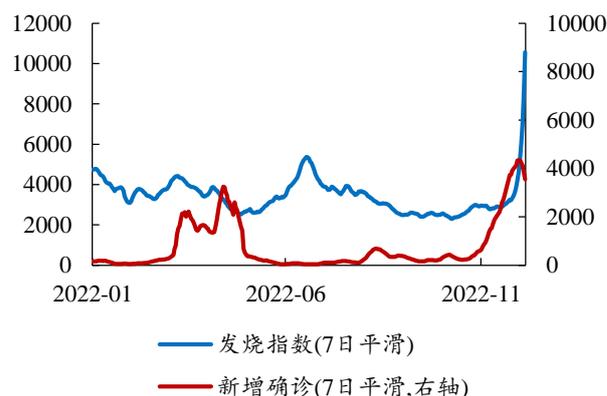
数据来源：百度指数、Wind、开源证券研究所

图3：新冠搜索指数有一定相关性



数据来源：百度指数、Wind、开源证券研究所

图4：发烧搜索指数从11月开始生效



数据来源：百度指数、Wind、开源证券研究所

2.2、区域搜索指数具有高频、领先、抗干扰特征

基于上述分析，“莲花清瘟”指数和“新冠”指数宜着重考察，我们发现具备

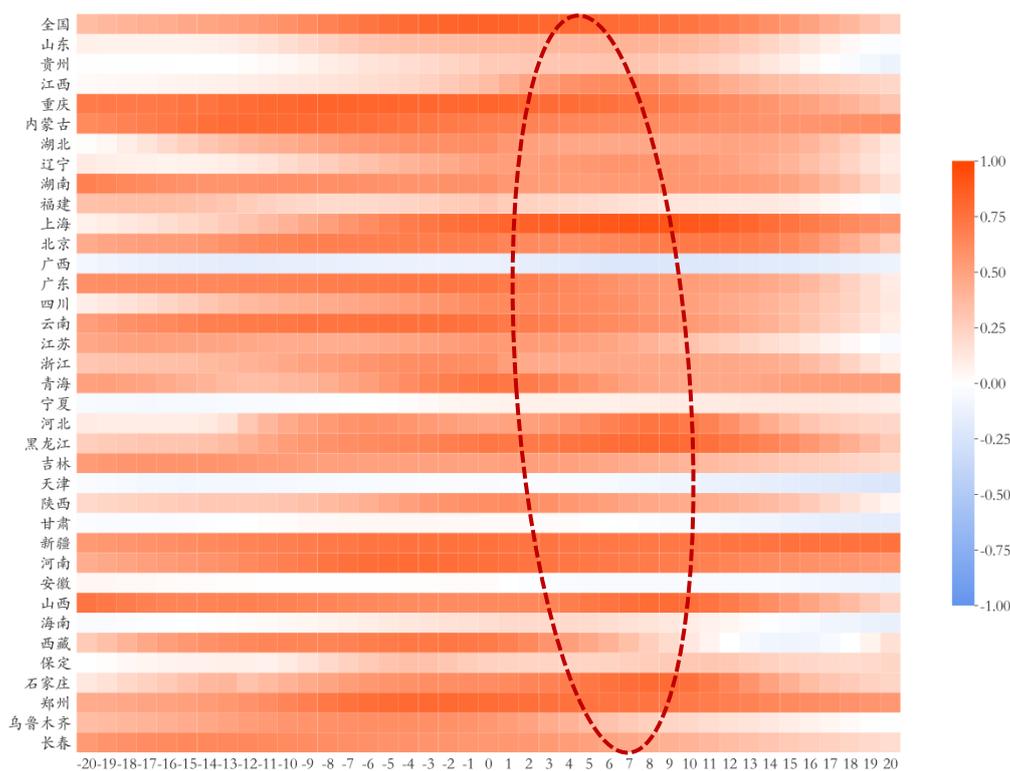
高频、领先性、抗干扰等特点。考虑到“新十条”对确诊病例的干扰，我们将时间窗口选择为2022年1月1日-2022年12月7日，并计算各搜索指数与确诊病例的相关性。从全国来看，连花清瘟指数、新冠指数、发烧指数与确诊病例的相关性分别为0.70、0.63、0.03；连花清瘟指数的领先期主要分布在5-10日、平均为7.1日，新冠指数的领先期主要分布在2-8日、平均为6.1日。

表1：搜索指数与确诊病例相关性

搜索指数	区域	相关性	相关性 (7日平滑)	相关性 (最优领先期)	相关性 (7日平滑, 最优领先期)
连花清瘟	全国	0.70	0.72	0.84	0.84
	省市平均	0.42	0.50	0.51	0.57
新冠	全国	0.63	0.63	0.64	0.65
	省市平均	0.43	0.47	0.49	0.53
发烧	全国	0.03	0.03	-0.06	-0.02
	省市平均	0.08	0.18	0.10	0.21

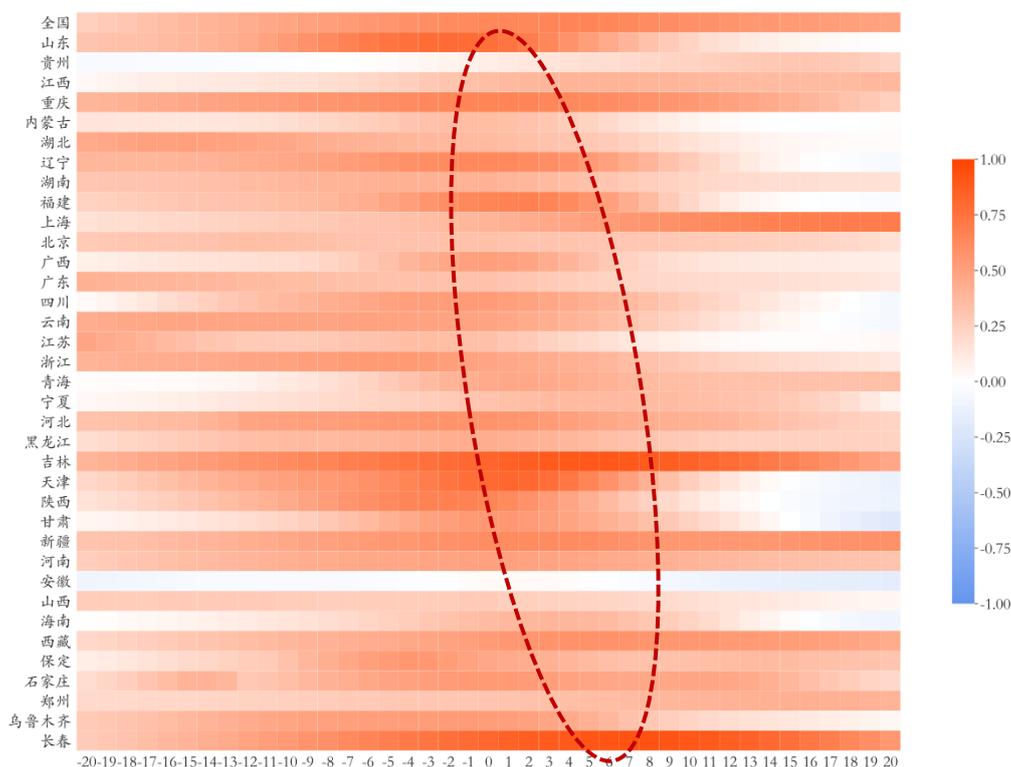
数据来源：百度指数、Wind、开源证券研究所

图5：“连花清瘟”搜索指数领先确诊病例5-10日



数据来源：百度指数、Wind、开源证券研究所；注：横坐标表示对搜索指数作平移处理，例如“1”列表示搜索指数前置1日时与确诊病例的相关性，“-1”列表示滞后1日，以此类推。

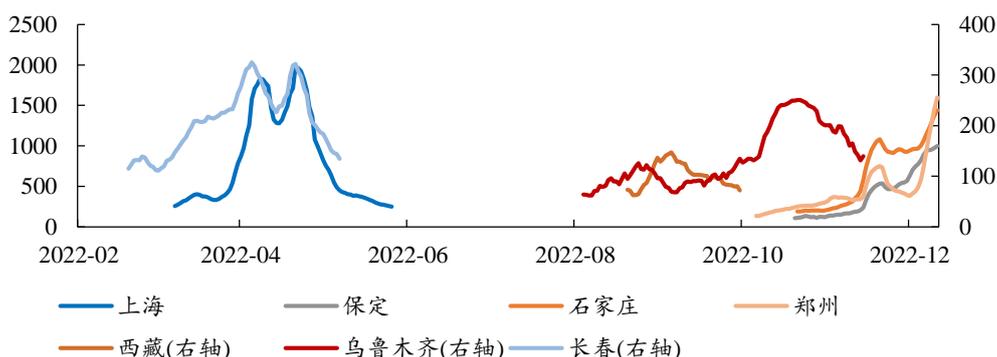
图6：“新冠”搜索指数领先确诊病例 2-8 日



数据来源：百度指数、Wind、开源证券研究所；注：横坐标表示对搜索指数作平移处理，例如“1”列表示搜索指数前置 1 日时与确诊病例的相关性，“-1”列表示滞后 1 日，以此类推。

表观疫情数据的收集、处理与披露受外部因素影响。当人力、医疗、设备资源匮乏时，表观疫情数据可能与真实疫情存在一定出入。从这个角度来看，搜索指数具有一定的抗干扰能力。

图7：连花清瘟指数与疫情节点较为吻合



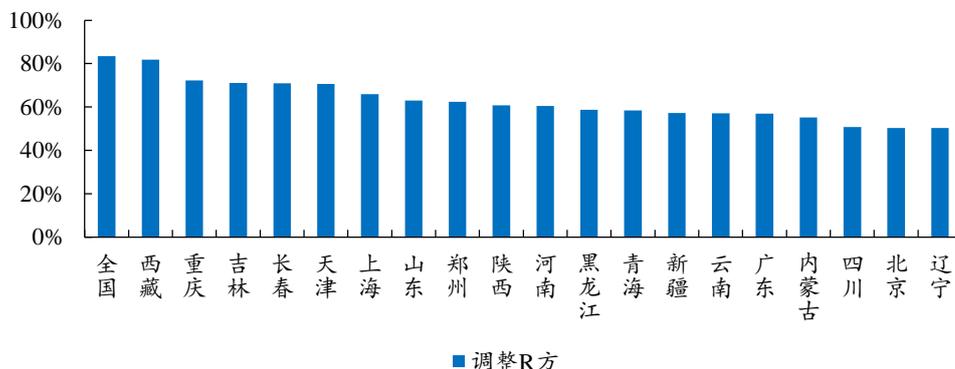
数据来源：百度指数、Wind、开源证券研究所

3、利用搜索指数拟合疫情水平

目前，“小步走、不停步”的防疫优化已初步迎来第一波冲击，我们尝试将搜索指数拟合疫情水平。回归统计数据显示，20 个地区调整 R 方超过 0.5，表明该方

法具有一定解释力度。

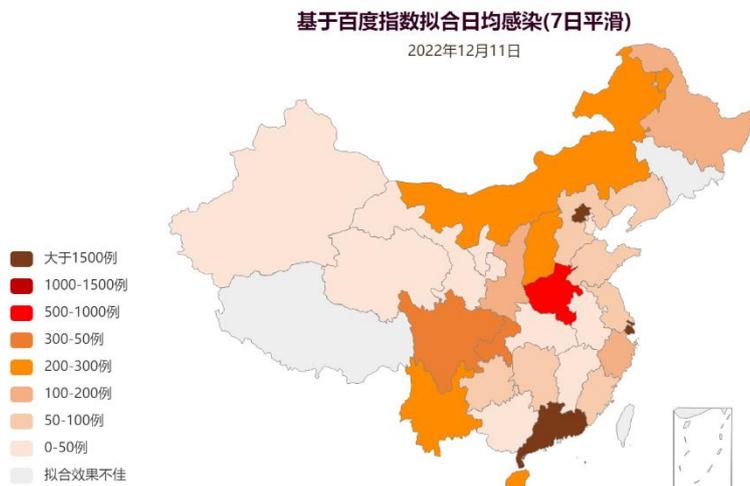
图8：各地搜索指数与确诊病例拟合效果较好



数据来源：百度指数、Wind、开源证券研究所

从时间序列来看，本轮疫情始于11月初，河南、河北、内蒙古一带较快蔓延，11月初以广州多区宣布解除疫情防控临时管控区为标志、广东主动迎接第一波冲击，12月初北京被动进入第一波冲击。目前上海、吉林、黑龙江等地疫情可能以较快的速度扩散。

图9：基于百度指数拟合日均感染病例



数据来源：百度指数、Wind、开源证券研究所

4、风险提示

1. 国内疫情反复超预期：如果疫情防控优化政策反复或第一波冲击应对不及时，那么疫情可能扩散超预期。
2. 搜索指数有效性下降：若搜索指数无法真实反映搜索情况，则利用搜索指数

拟合新增确诊的模型有效性下降。

特别声明

《证券期货投资者适当性管理办法》、《证券经营机构投资者适当性管理实施指引（试行）》已于2017年7月1日起正式实施。根据上述规定，开源证券评定此研报的风险等级为R3（中风险），因此通过公共平台推送的研报其适用的投资者类别仅限定为专业投资者及风险承受能力为C3、C4、C5的普通投资者。若您并非专业投资者及风险承受能力为C3、C4、C5的普通投资者，请取消阅读，请勿收藏、接收或使用本研报中的任何信息。

因此受限于访问权限的设置，若给您造成不便，烦请见谅！感谢您给予的理解与配合。

分析师承诺

负责准备本报告以及撰写本报告的所有研究分析师或工作人员在此保证，本研究报告中关于任何发行商或证券所发表的观点均如实反映分析人员的个人观点。负责准备本报告的分析师获取报酬的评判因素包括研究的质量和准确性、客户的反馈、竞争性因素以及开源证券股份有限公司的整体收益。所有研究分析师或工作人员保证他们报酬的任何一部分不曾与，不与，也将不会与本报告中具体的推荐意见或观点有直接或间接的联系。

股票投资评级说明

	评级	说明
证券评级	买入（Buy）	预计相对强于市场表现 20%以上；
	增持（outperform）	预计相对强于市场表现 5%~20%；
	中性（Neutral）	预计相对市场表现在-5%~+5%之间波动；
	减持（underperform）	预计相对弱于市场表现 5%以下。
行业评级	看好（overweight）	预计行业超越整体市场表现；
	中性（Neutral）	预计行业与整体市场表现基本持平；
	看淡（underperform）	预计行业弱于整体市场表现。

备注：评级标准为以报告日后的 6~12 个月内，证券相对于市场基准指数的涨跌幅表现，其中 A 股基准指数为沪深 300 指数、港股基准指数为恒生指数、新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）、美股基准指数为标普 500 或纳斯达克综合指数。我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议；投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者应阅读整篇报告，以获取比较完整的观点与信息，不应仅仅依靠投资评级来推断结论。

分析、估值方法的局限性说明

本报告所包含的分析基于各种假设，不同假设可能导致分析结果出现重大不同。本报告采用的各种估值方法及模型均有其局限性，估值结果不保证所涉及证券能够在该价格交易。

法律声明

开源证券股份有限公司是经中国证监会批准设立的证券经营机构，已具备证券投资咨询业务资格。

本报告仅供开源证券股份有限公司（以下简称“本公司”）的机构或个人客户（以下简称“客户”）使用。本公司不会因接收人收到本报告而视其为客户。本报告是发送给开源证券客户的，属于商业秘密材料，只有开源证券客户才能参考或使用，如接收人并非开源证券客户，请及时退回并删除。

本报告是基于本公司认为可靠的已公开信息，但本公司不保证该等信息的准确性或完整性。本报告所载的资料、工具、意见及推测只提供给客户作参考之用，并非作为或被视为出售或购买证券或其他金融工具的邀请或向人做出邀请。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。客户应当考虑到本公司可能存在可能影响本报告客观性的利益冲突，不应视本报告为做出投资决策的唯一因素。本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。本公司未确保本报告充分考虑到个别客户特殊的投资目标、财务状况或需要。本公司建议客户应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议。在任何情况下，本公司不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。若本报告的接收人非本公司的客户，应在基于本报告做出任何投资决定或就本报告要求任何解释前咨询独立投资顾问。

本报告可能附带其它网站的地址或超级链接，对于可能涉及的开源证券网站以外的地址或超级链接，开源证券不对其内容负责。本报告提供这些地址或超级链接的目的纯粹是为了客户使用方便，链接网站的内容不构成本报告的任何部分，客户需自行承担浏览这些网站的费用或风险。

开源证券在法律允许的情况下可参与、投资或持有本报告涉及的证券或进行证券交易，或向本报告涉及的公司提供或争取提供包括投资银行业务在内的服务或业务支持。开源证券可能与本报告涉及的公司之间存在业务关系，并无需事先或在获得业务关系后通知客户。

本报告的版权归本公司所有。本公司对本报告保留一切权利。除非另有书面显示，否则本报告中的所有材料的版权均属本公司。未经本公司事先书面授权，本报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

开源证券研究所

上海

地址：上海市浦东新区世纪大道1788号陆家嘴金控广场1号楼10层
邮编：200120
邮箱：research@kysec.cn

深圳

地址：深圳市福田区金田路2030号卓越世纪中心1号楼45层
邮编：518000
邮箱：research@kysec.cn

北京

地址：北京市西城区西直门外大街18号金贸大厦C2座9层
邮编：100044
邮箱：research@kysec.cn

西安

地址：西安市高新区锦业路1号都市之门B座5层
邮编：710065
邮箱：research@kysec.cn