小马智行 20260907

小马智行 Robotaxi 商业化拐点分析：单车盈利与技术闭环

摘要

- 小马智行在北广深已实现单车盈利，单车日均运营成本控制在 250 元以内（折旧约占 50%），日均净收入 320-340 元，毛利率达 20%-30% 以上。
- 技术降本路径明确：采用车规级芯片（OrinX）替代高价 H100，硬件成本大幅下降；通过 L4 级自动泊车与充电技术，实现 3 名地勤覆盖 100 台车的极低人力配比。
- 网络效应驱动效率提升：付费里程利用率从 15%-20% 提升至 40%-45%，预计未来可达人类司机 50%-60% 水平；90% 用户通过自有 APP 叫车，流量去中心化趋势明显。
- 世界模型 1.0 与 2.0 形成递归闭环：1.0 生成海量虚拟场景，2.0 利用 AIAgent 提取真实世界概率分布进行校准，自动化替代人工，显著加速技术迭代效率。
- Robotaxi 行业拐点指标：单一公司在单一城市部署 1,000 台车（渗透率约 1%）被视为爆发信号；公司预计 2027 年重卡与轻卡业务将迎来商业化显著发展。
- 竞争壁垒与预期差：车企因缺乏“决策数据”难以形成闭环，L2 经验难以直接迁移至 L4；特斯拉 Cybercab 独立组队验证了 L4 需遵循渐进式验证的客观规律。

Q&A

请介绍一下小马智行 Robotaxi 业务在中国和海外市场的商业模式有何不同？

Robotaxi 的商业模式主要包含三个部分：前端的获客与流量、中端的自动驾驶技术（即“开车”的 AI），以及后端的车辆资产与后勤保障（包括保险、保养、清洁、地勤等）。 在中国市场，我们目前主导模式的前两个环节，即流量获取和自动驾驶技术。而后端资产与后勤部分，我们采用“共建模式”，与地
出租车公司、租赁公司或网约车平台合作，由他们负责车辆资产、折旧、
保养等。随着单车经济模型的利润率提升，未来计划将更多后端业务外包
海外市场，我们的定位是纯粹的技术提供商，只专注于提供自动驾驶技术

人在草木间

具体而言，我们不涉足前端的流量获取和后端的车辆资产管理。海外的合作伙伴会向我们采购车辆，从而承担车辆的采购成本、后续折旧以及清洁、保养、保险等所有后勤保障工作。我们则向其授权自动驾驶技术。目前，我们与 Uber、以及在新加坡、韩国、欧洲和中东地区的合作伙伴，均采用这种合作模式。

小马智行 Robotaxi 业务的单车经济模型是怎样的，尤其是在成本构成、收入水平和盈利能力方面？

我们的单车经济模型中，收入口径为扣除所有折扣或优惠券后的净收入，这也是财务报告中确认的 Robotaxi 收费收入。成本端主要包含七项，其中车辆折旧是最大头，约占总成本的一半，其余成本包括充电、远程协助、地勤、保养、保险和停车费。 目前，在北京、广州和深圳三个城市，我们的 Robotaxi 车队或单车均已实现盈利。每辆车在这三个城市的日均运营总成本（含折旧）控制在 250 元人民币以内。具体来看，折旧成本约为 125 元，这是基于单车约 26.7 万元的成本按 6 年进行摊销计算得出的。其余六项成本合计约为 125 元。 在收入方面，北京、广州、深圳三地的单车日均净收入在 320 元至 340 元之间。因此，单车毛利率根据城市不同，普遍在 20%至 30%多的区间。其中，北京和深圳的盈利水平相对较高，广州因整体打车单价略低而稍逊一筹。

在实现当前单车经济模型的过程中，小马智行在技术层面具备哪些核心优势以支持成本控制？

我们的成本优势根植于技术优势。首先，在计算硬件方面，我们采用价格相对低廉的车规级芯片（如 Orin X）来运行 AI 算法。相较于部分美国友商可能使用价格高出数十倍的非车规级芯片（如 H100），我们在硬件成本上实现了显著节约。考虑到车规级芯片在抗震动、耐高低温及长使用寿命方面的优势，而数据中心芯片在车载环境下寿命可能缩短，这进一步放大了我们在单位时间内的折旧成本优势。 其次，在运营自动化方面，我们实现了 L4 级别的全自动泊车与充电准备。车辆能够自主从公路驶入停车场，通过闸机，并精准停入充电桩旁的停车位，整个过程无需人工干预，仅需地勤人员进行最后的充电枪插拔操作。这与一些需要大量地勤人员手动将车辆开入停车位进行充电的 L4 方案相比，极大地降低了人力成本。基于此技术，我们的人员配比可以达到每 3 名地勤人员负责 100 台车的清洁、保养、维修检查等工作，这些人力成本已计入每日 250 元的总运营成本中。这些技术上的精进是实现极致运营成本控制的关键。此外，网络效应的增强也将进一步提升未来的利润空间。

自动驾驶运营中的网络效应具体体现在哪些方面，这些效应如何提升车辆的运营效率和用户体验？

网络效应体现在多个层面。首先，运营面积的扩大使得能够接收更多长距离订单，这间接提升了单车的里程利用率。其次，车辆能够实现全天候、全时段稳定接单，包括暴雨、昼夜及早晚高峰等复杂场景，这与部分竞品在恶劣天气或高峰时段可

能停运形成对比，避免了因服务不稳定导致的用户流失和里程利用率下降。再者，在路线规划上，系统能选择并执行最短路径，即便该路径路况可能更为复杂，而技术能力不足的玩家可能会选择绕远路以简化驾驶场景，但这会牺牲乘客的时间和出行效率，影响用户体验。这些网络效应，背后是技术能力的支撑，共同促进了新增注册用户的增长。目前，超过 90%的用户通过自有 APP 叫车，对第三方打车平台的依赖度很低，这为构建自有流量平台奠定了基础。此外，车队规模本身也是一种网络效应，车队密度越大，乘客的等待时间就越短，从而提高成单率。综合来看，全天候运营、最短路径行驶、高密度车队以及大范围运营面积等多重网络效应，共同驱动了付费里程利用率的显著提升。该比率已从 2025 年同期的 15%-20%增长至目前的 40%-45%左右。这意味着车辆日均行驶 300 公里中，约有 130-140 公里是可转化为订单的付费里程。未来随着运营面积的进一步扩大和车队密度的增加，预计该比率能够达到甚至超过人类司机 50%-60%的水平，因为系统可以基于大数据进行全局最优调度，确保每辆车高效执行指令，从而最大化商业化收费里程的占比。

请阐述世界模型 1.0 与 2.0 的核心功能、相互关系及其对自动驾驶技术迭代效率的影响？

世界模型 1.0 本质上是一个在云端生成的虚拟世界或“虚拟驾校”，用于训练自动驾驶车辆的 AI 司机。在这个虚拟环境中，系统每周能生成上百亿公里的驾驶场景，尤其是各种长尾和极端场景（corner case），其规模可与特斯拉自创立以来所有车辆真实行驶里程的总和相媲美。AI 司机在这些海量虚拟场景中通过强化学习不断试错、修正，提升驾驶能力，随后将训练成熟的 AI 司机模型“蒸馏”并部署到车端。世界模型 2.0 并非 1.0 的升级版，其核心功能是提升 1.0 生成虚拟场景的“精度”。所谓精度，指的是虚拟场景中事件发生的概率分布与真实世界的匹配程度。例如，真实世界中行人在过马路时边看手机边观察交通灯的概率可能是 80%，如果世界模型 1.0 生成的虚拟场景中此概率仅为 50%，就存在精度差。这种偏差会影响车辆在真实世界中对各类物体的行为预测和最优安全策略的制定。在 2.0 出现之前，这种精度校准工作依赖工程师或质检员的人工观察和反馈，过程机械且效率低下。世界模型 2.0 引入了 AI Agent，能够自动化地分析车辆在真实世界中收集的数据。当车辆遇到长尾场景时，系统能捕捉到该事件的真实概率分布，并记录下车辆自身传感器（如激光雷达、摄像头）的角度、点云精度，以及刹车和动力扭矩等详细参数。这些真实世界的数据被反馈给世界模型 1.0，用于修正其生成特定场景的概率分布，例如将某个事件的概率从 50%精确调整到 40%。通过这种方式，1.0 和 2.0 形成了一个递归式自我改进（Recursive Self-Improvement）的闭环：1.0 生成虚拟世界用于训练，2.0 利用真实世界数据反哺和校准虚拟世界，从而缩小虚拟与现实之间的精度差距。这一闭环自动化了以往的人工校准流程，显著加快了技术的进化和迭代效率。

大语言模型在自动驾驶技术研发中扮演何种角色，它与自动驾驶决策本身的关系是怎样的？

大语言模型在自动驾驶领域的主要作用是作为一种开发工具，辅助工程师提升编写代码的效率。才华横溢的工程师们可以利用大语言模型将他们的思维和逻辑更快地转化为复杂的、高维度的代码，从而加速研发进程。然而，大语言模型本身并不能直接用于车辆的实时驾驶决策。首先，从算力角度看，大语言模型的参数量巨大，远超车端计算单元（例如 4 颗 Orin X 芯片）的处理能力。其次，依赖云端计算和 5G/6G 通信传输决策是不可行的，因为信号传输的延迟无法完全消除，而自动驾驶的所有关键决策必须在车端瞬时完成以确保安全。更本质的是，驾驶能力与语言理解能力是两种不同的技能。一个人即使文化水平不高，也能通过学习掌握驾驶的基本功和与道路参与者的交互方式，这并不需要具备博士级别的知识储备。自动驾驶技术的核心是处理高维度的复杂代码，这些代码需要在有限的算力下高效运行，其逻辑并非基于简单的“if-else”规则，而是更为复杂的系统。因此，大语言模型是提升研发效率的强大辅助工具，但自动驾驶的实现依赖于专门为其设计的、能在车端高效运行的算法模型。

如何看待自动驾驶技术与大语言模型的差异，以及特斯拉 Cybercab 的推出对 L4 技术路线的验证作用？

自动驾驶技术与大语言模型存在本质区别。大语言模型本质上是由海量神经网络参数构成的，例如一个拥有 3 万亿参数的模型，相当于通过 3 万亿个开关组合来逼近真实世界，可以视为海量的“if-else”判断。相比之下，自动驾驶系统是更高维度的计算机代码，能够精炼地运行在算力有限的车端芯片上，并非简单的规则堆砌。 特斯拉近期在美国奥斯汀部署约 40 多台 Cybercab 的举动，恰恰印证了 L4 技术路线的独特性和必要积累。市场普遍关心基于 L2 的 FSD 技术能否进化至 L4，我们认为这非常困难。L2 积累的知识体系难以直接迁移至 L4。为此，特斯拉也组建了独立于其 L2 团队的 Robotaxi 团队来专项研发 Cybercab。其发展路径也遵循了 L4 行业的普遍规律：从几台车起步，逐步扩展到几十台，进行测试、无人化测试，再到无人化收费运营。这表明，L4 的实现无法通过对已售出的存量消费级车辆进行简单的远程升级来一蹴而就，而是需要经历一个严谨的、渐进的验证过程，包括在雨天等更复杂场景下的持续测试。这说明，即便是特斯拉，也必须遵循 L4 技术发展的客观规律，这从侧面验证了 L4 公司多年来积累的经验和发展路径的正确性。车企并非不能研发 L4，但其 L2 阶段的经验积累对 L4 的帮助有限，若要涉足 L4，仍需重新组建团队，并采用一套全新的、包含冗余设计、传感器配置、车载算力和驾驶模型的技术方案。

公司在重卡和轻卡领域的商业化规划是怎样的，其商业模式与 Robotaxi 业务有何异同？

公司预计在 2027 年将迎来重卡和轻卡业务商业化的显著发展。具体而言，重卡业务将聚焦于港口、西北等地的专线运输场景，这些场景比跨省市的干线物流更

容易在政策上实现无人化运营。轻卡业务则主要面向市中心区域，解决商超配送等物流需求。轻卡车型的技术方案可以完全复用 Robotaxi 的成果，因为其车辆底盘的四轮投影面积和体积与乘用车相近，核心差异仅在于承载对象是货物而非乘客。因此，我们可以利用现有的 Robotaxi 运营道路、后勤保障体系、地勤人员以及保洁保养等资源来支持轻卡业务。 在商业模式上，无论是重卡、轻卡还是 Robotaxi，公司都坚持采用纯运营模式，而非销售 L2/L3 级别的车辆。收入模式是根据运营服务来计费，例如重卡按“道路里程-吨”收费，Robotaxi 则按出行服务收费，这与我们现有的 Robotaxi 打车业务模式保持一致。

当前车企与自动驾驶公司在技术路线上似乎趋于融合，即同时采用模仿学习和强化学习，并结合真实世界数据与仿真数据。例如，拥有大量真实数据的车企通过世界模型补足长尾场景，而自动驾驶公司则在强化训练处理长尾能力后，用真实数据补充概率分布。这种技术路径的融合是否合理，特斯拉是否也采用类似方式？车企若同时拥有这两方面优势，能否顺利发展至 L4 阶段？

车企在发展 L4 技术时面临一个核心挑战，即数据无法形成闭环。一个完整的驾驶行为数据链条应包含三个环节：感知（看到什么）、决策（思考什么）和执行（做出什么动作）。Robotaxi 能够完整地采集这三方面的数据：车辆通过 360 度传感器全面感知环境，系统记录下决策过程，并最终执行动作。这三者的数据构成了可供迭代优化的闭环。 然而，车企虽然拥有大量用户驾驶数据，但这些数据是不完整的。车企的系统能够记录传感器“看到”了什么，以及驾驶员最终“做出”了什么动作（例如转动方向盘），但却无法获知驾驶员在做出动作前的“思考”过程。人类驾驶员不会向机器解释其每一个决策背后的心路历程。这种决策数据的缺失，导致车企的数据链条存在断点，无法形成有效的闭环，从而极大地制约了其技术迭代的效率。因此，要真正实现 Robotaxi，必须使用 Robotaxi 自身在真实运营和虚拟仿真中产生的数据，形成一个能够持续进行查缺补漏的递归循环。

从全球范围看，Robotaxi 在海外和国内的商业化落地，哪个区域的发展速度可能会更快？

目前来看，中国和美国无疑是 Robotaxi 发展最快、也最早起步的两个市场。这两个国家拥有本土的优秀企业在持续推动研发、测试和运营，并且在政策层面也具备先发优势。 在其他海外市场，欧洲的个别国家和中东地区也展现出较快的发展潜力。需要指出的是，欧洲并非全境开放，而是特定几个国家进度较快。在中东，卡塔尔和阿联酋的发展速度则相对领先。预计很快会有关于我们与包括 Uber 在内的合作伙伴在这些地区合作的进一步公开信息。

自动驾驶行业是否存在一个非线性的增长拐点？考虑到该行业受到政策、基建等物理世界因素的约束，与网络效应强、边际成本低的纯软件行业不同，当技术成熟后，行业是否会迎来爆发式增长？

Robotaxi 行业是否存在一个非线性的增长时刻？如果存在，预计何时到来，会在哪些区域率先发生？

行业内非线性的增长时刻预计将会出现，尽管无法百分之百保证具体在哪个市场会突然爆发，但判断下来，这种时刻大概率会发生在中国、美国、欧洲和中东这几个区域。

从数据层面看，单一城市运营车辆达到怎样的量级可以被视为行业发展的拐点？

单一公司在单一城市部署至少 1,000 台车，可以被视为一个数字上的拐点。届时，该公司在该城市的渗透率可能已经达到 1%。

如何看待滴滴等原有运营公司转型进入 Robotaxi 领域所带来的竞争关系，特别是考虑到它们掌握的流量入口优势？

对于这类公司而言，最大的制约因素在于技术。其技术何时能达到成熟并面向市场的水平，目前尚无法判断。前端的流量入口并非行业的关键瓶颈。以小马智行为例，车队规模从 2025 年的 100 台增长到接近 1,000 台的过程中，每投放一个批次的 100 台车，都能吸引 15 万至 20 万的新增用户，这表明车辆本身在路面上行驶就是一种有效的移动广告。目前在未投入任何市场营销费用的情况下，仅凭每个城市几百台车的运营，就已获得至少 150 万的注册用户。因此，业务增长并不依赖流量平台，也不信奉通过价格补贴来吸引用户。核心竞争力在于技术和服务本身：标准化的车辆保养、保洁和检修确保了清洁舒适的乘车环境；自动驾驶系统提供了比人类司机更具可预见性的驾驶体验，避免了因司机驾驶风格差异带来的不适。只要能实现足够大的运营覆盖范围、足够短的接驾时间、高效的点对点行车路径，并能应对早晚高峰和各种天气状况，Robotaxi 相较于普通快车就具备天然优势。因此，平台入口的加持并非发展的必要条件。至于这些公司自身能否成功研发，核心还是要看实际成果，至少要在单一城市部署百台级别的车队并跑出规模效应，才能证明其技术已站稳脚跟并具备向上迭代的基础。

Robotaxi 行业的核心壁垒是什么？未来的竞争格局会趋向于头部集中还是相对分散？

Robotaxi 行业的壁垒是多方面的，并非单一因素决定。任何成功的商业模式都像一个木桶，每个木板都不能是短板，且构建每一块长板都存在困难，这样才能

形成长期壁垒。具体到本行业，主要壁垒包括以下几点：第一，技术壁垒；第二，运营和后勤保障能力，这关系到成本控制、打车体验以及能否持续获客并让用户愿意支付溢价；第三，充足的资金实力，以支持在不同国家市场与当地车企合作，满足当地车型要求；第四，量产能力，即能否与车企共同定义和实现车辆的冗余系统、电源管理系统和线控系统等，使其符合标准。这些环节都构成了很高的门槛。只有真正着手实践，才能意识到需要解决的复杂问题。对于许多大型或已有成熟产业的公司而言，其内部文化或 KPI 考核机制可能不适合进行长期的、有沉没成本的投入来补齐所有短板。而对于创业型公司来说，实现 Robotaxi 是唯一目标，因此在执行上可能更快、更具优势。

从 Robotaxi 技术拓展至 Robotruck 业务，技术的复用性有多高？Robotruck 的应用场景、市场空间和落地节奏是怎样的？公司是否会考虑进入无人矿卡这类封闭场景？

卡车业务主要部署在场景更大的专线或干线物流。全国干线物流约有 700-800 万辆卡车，而专线物流主要集中在矿产资源丰富的西北部，约有 180 万台。专线的优势在于距离较短，从几公里到几十公里不等；干线距离更长，需要跨越多个城市，因此在政策突破上可能需要更长的周期。从商业化角度看，更大规模的无人化会先在专线物流领域实现，而干线物流的推广则需等待相关政策的完善。公司不会进入矿区内的无人矿卡业务，因为这不符合公司的优势定位。公司的优势在于技术难度高但市场空间广阔的公开道路干线和专线场景。与矿区业务可能的合作模式是，由其他车辆在矿区内完成作业，将矿产运输至矿区门口，再由我们的自动驾驶卡车在公开的专线道路上接力，运送至火车站、发电厂等下一个交通枢-枢纽。除了西北部的专线场景，公司也在拓展中国沿海地区与港口相连的短途专线物流，例如近期在深圳妈湾港开展的物流业务，并且有更多与港口连接的物流项目正在洽谈中，预计未来几个月会有新的进展公布。

从公司当前发展阶段来看，最紧迫需要提升或最缺乏的资源是什么？

当前的重点工作主要有几个方面：首先是搭建并完善整个安全体系；其次是与车企商讨未来两代车型的开发计划；再者是人才团队的建设，例如需要引进处理 AI 认证等方面的专业人才。在算力和数据方面，目前并不存在短缺。公司的技术路径并非依赖采集大量真实路跑数据，而是更侧重于利用虚拟数据进行仿真测试，再通过真实世界运营进行校准，形成闭环。目前运营的近千台车队以及世界模型 1.0 和 2.0 的闭环已经证明了这条技术路径的有效性。因此，当前没有特别紧迫的难题，更多的是一步一个脚印，踏实地将各项工作做好。

Robotaxi 行业似乎存在一个特殊现象，即个别企业的负面事件可能会导致整个行业发展停滞，例如监管暂停发放牌照。这种情况未来是否会有所改观？

监管机构已经注意到了这一现象。本行业的一个特点是，与普遍不涉及人身物理安全的大模型应用不同，Robotaxi 直接关系到人身安全。因此，一家企业的表

现好坏，市场和监管机构能够迅速检验出来。这也使得行业的竞争格局非常清晰，哪些玩家能够长期存续下来是显而易见的。

从 AI 应用的角度看，对 Robotaxi 行业有什么展望？

Robotaxi 是 AI 应用落地的重要一环。从技术、用户体验、后勤保障、安全体系、运营到生产制造等各个方面，该行业都已相对成熟，目前正处于规模化扩张（scale up）的阶段。建议可以多多关注这一 AI 应用领域的发展。