

数据经济学

第六章：数据市场均衡

陈希路

暨南大学经济学院

2026 年秋

章节目录

1 第一节：数据市场均衡理论

2 第二节：数据一级市场定价

3 第三节：数据二级市场定价

4 第四节：数据三级市场定价

市场均衡

是指商品或服务的供给和需求达到平衡的一种状态。

理论基础：

- 核心：供需理论（亚当·斯密）
- 动态性：随供需条件变化，价格变化以重新实现均衡
- 分类：一般均衡 vs. 局部均衡

数据市场均衡的特点：

- 复杂性：涉及确权、隐私、合规、价值链
- 去中心化：以平台为中心，涉及供给方、需求方、平台方

第一节

数据市场均衡理论

数据市场均衡的八大前提

理论

结合数据要素特征与传统经济理论，提出八大前提：**理论：八大前提**

- ① **确权**：产权明晰，交易有据
- ② **脱敏**：保护隐私，提高信任
- ③ **标准化**：统一格式，降低协商成本
- ④ **信息完全**：破解信息悖论与不对称
- ⑤ **竞争性**：参与者地位对等
- ⑥ **无外部性**：交易仅影响双方
- ⑦ **无交易成本**：无手续费、税费
- ⑧ **有效监管**：规范行为，保障权益

数据市场均衡的八大前提

案例：上海数交所

案例：上海数交所

- **确权 + 标准化**：首创数据产品登记凭证，原始数据转为可流通标准化产品
- **信息完全**：提供产品说明书与质量评估，买方不见底层数据即可预估价值
- **无交易成本 + 监管**：引入第三方合规审查，降低面议成本

数据市场均衡的机制与多样性

● 一般市场均衡的基本内涵

- **核心定义**：存在特定的均衡价格，使得市场供给量完全等于需求量
- **价格调节机制**：
 - 超额供给 → 导致价格下降
 - 超额需求 → 导致价格上涨
 - 均衡状态 → 价格停止波动，保持稳定

● 数据市场均衡的特征

- **供需决定**：数据市场的均衡价格由数据的供给与需求共同决定
- **模型多样性**：不同市场结构的供需关系与价格形成机制各异，均衡模型高度多样

不同市场结构的数据市场均衡

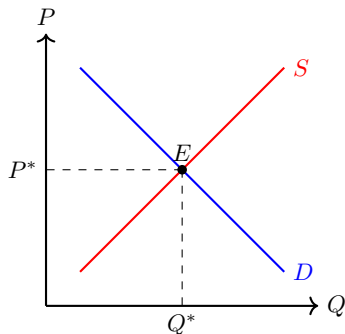
完全竞争市场

- 市场结构特点：

- 单个买者和卖者都无法控制价格，是价格的接受者

- 市场均衡点表现：

- 供给曲线与需求曲线共同决定市场均衡
- 在均衡价格下，市场出清，即需求量完全等于供给量



不同市场结构的数据市场均衡

垄断竞争市场

- 市场结构特点：

- 市场中存在大量的卖方和买方
- 各卖方占一定市场份额，产品差异化，通过价格与非价格竞争争夺市场

- 市场均衡点表现：

- 在短期内，企业可能获得超额利润
- 但在长期来看，新企业的自由进入会导致市场回归到零超额利润的均衡状态

不同市场结构的数据市场均衡

寡头垄断市场 + 彭博终端

理论：寡头垄断均衡

- **结构特点**：少数几家大型数据供应商主导市场，企业间相互依存
- **均衡表现**：因**进入壁垒**高，寡头通常能**长期维持超额利润**
- 与垄断竞争不同：后者因自由进入，长期回归零利润

案例：彭博终端

- **供给壁垒高**：整合全球数百万数据源、建专用闭环网络，仅路孚特等极少数寡头能抗衡
- **需求缺乏弹性**：基金经理错过一秒数据即损失惨重，对价格不敏感
- **均衡表现**：彭博为价格制定者，长期获**高额超额利润**

不同市场结构的数据市场均衡

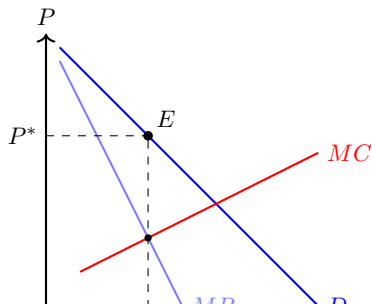
完全垄断市场

- 市场结构特点：

- 市场上只有一个供应商占主导地位

- 市场均衡点表现：

- 完全垄断企业是价格的制定者
- 企业根据边际收益等于边际成本 ($MR = MC$) 的原则确定产量，并在需求曲线上获得更高的垄断定价



不同市场结构的数据市场均衡

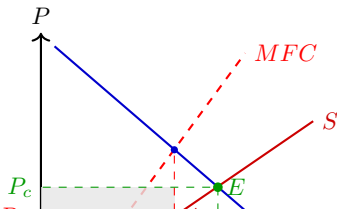
单一买方市场

- 市场结构特点：

- 只有一个买方，但供给方有多个
- 市场供给曲线通常是相对水平的

- 市场均衡点表现：

- 由于买方在市场上的垄断地位，市场需求曲线相对更加垂直
- 买方可以影响价格，使得价格随着需求变化而变化
- 买方将购买量限制在边际价值 = 边际要素成本 ($D = MFC$) 处，按供给曲线付更低价格，获垄断租金



不同市场结构的数据市场均衡

双边市场

- 市场结构特点：

- 市场供给曲线和需求曲线都不是单一的曲线
- 而是由两个群体的需求和供给曲线组成的

- 市场均衡点表现：

- 市场上的均衡点不仅取决于供给和需求的均衡
- 还取决于两个群体之间的网络外部性

自由数据市场的市场均衡机制：模型

假定自由数据交易市场的供需关系如下：

- 供给函数： $Q_s = c + d \times P$
- 需求函数： $Q_d = a - b \times P$

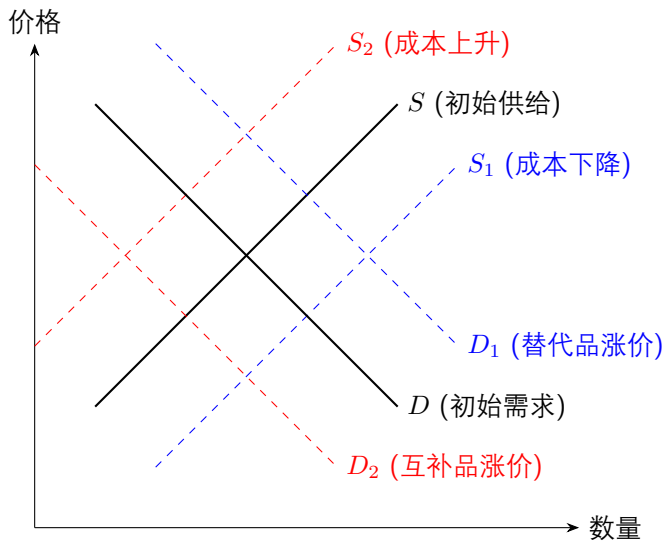
市场均衡点 ($Q_s = Q_d$) 的均衡价格为：

$$P = \frac{a - c}{d + b}$$

其中：

- P ： 价格
- a, c ： 截距项
- b, d ： 弹性系数（价格变动引起数量变动的程度）

自由数据市场的市场均衡机制：曲线移动



自由数据市场的市场均衡机制：动态变动

供给曲线移动

- 向左上 (S_2): 收集成本上升 $\rightarrow$ 均衡价格上升, 数量下降。
- 向右下 (S_1): 开发成本下降 $\rightarrow$ 均衡价格下降, 数量上升

需求曲线移动

- 向右上 (D_1): 替代品价格上升 $\rightarrow$ 均衡价格、数量均上升
- 向左下 (D_2): 互补品价格上升 $\rightarrow$ 均衡价格、数量均下降

案例分析：AI 数据标注市场的动态均衡

案例背景：Scale AI 与自动驾驶/大模型数据

独角兽 Scale AI 早期靠人工标注图像文本供大模型训练，如今引入 AI 预标注与自动化工具。

供需曲线的动态变动分析：

- 阶段一：高成本时期（均衡点在左上方）
 - 纯人工标注开发成本高，供给曲线 S 位置高，训练数据价格高昂、交易量受限
- 阶段二：技术进步引发的动态变动 ($S \rightarrow S_1$)
 - AI 辅助工具介入，大幅降低了数据的单位处理成本
 - 供给曲线向右下方移动，在相同价格下，供应商愿意提供更多数据
 - 新的市场均衡：高质量标注数据的均衡价格下降，总体的均衡数量增加

第二节

数据一级市场定价

数据一级市场定价概述

定价基础

在数据一级市场中，企业从事授权、采集和归集工作，将原始数据转化为可交易资源。

- **核心考量：** 转化的成本成为企业定价主要考虑的因素
- **准则：** 以最小的成本实现利润最大化
- **主要方法：** 会计学成本法

数据一级市场定价策略：成本结构

数据产品特性 + 案例：天眼查

核心视角——供给：生产成本

1. 企业生产函数

- 数据挖掘成本
- 脱敏成本
- 标准化成本
- 运储成本
- 隐私风险

2. 数据产品特性（案例：天眼查）

- **固定成本高**：初期巨资构建爬虫集群、清洗模型、知识图谱
- **复制边际成本 ≈ 0** ：多服务一个 API 客户成本几乎为零

成本法定价的优劣与方法

理论 + 案例：AI 数据标注平台

理论：成本法

- **优势**：客观、易操作，不受价格波动影响
- **劣势**：基于历史成本，往往会**低估**数据价值
- **方法**：直接成本法、间接成本法、标准成本法、作业成本法

案例：AI 数据标注平台

- **直接成本法**：按标注员计件工资算人力支出，构成数据包保底价
- **作业成本法**：拆为采集 → 预处理 → 标注 → 验收四环节，核算算力与管理成本加总报价

数据一级市场定价的实践应用

案例：加拿大统计局

采用成本法估计数据资产的价值。

估算步骤：

- 1 在国家职业分类中筛选与数据、数据库、数据科学相关的职业
- 2 假设各职业耗费在生产数据上的时间占比，作为直接劳动力成本
- 3 计算附加成本：
 - 间接成本：设定为总工资成本的 50%
 - 资本服务加成：设定为 3%
- 4 加总计算得出投入价值

加拿大 2018 年数据资产估算结果

- **投入价值**：295 亿至 400 亿美元
 - 数据：94 亿至 142 亿
 - 数据库：80 亿至 116 亿
 - 数据科学：120 亿至 142 亿
- **存量价值**：1570 亿至 2170 亿美元
 - 占知识产权净资产存量的 **68.9% 至 95.2%**

第三节

数据二级市场定价

数据二级市场定价概述

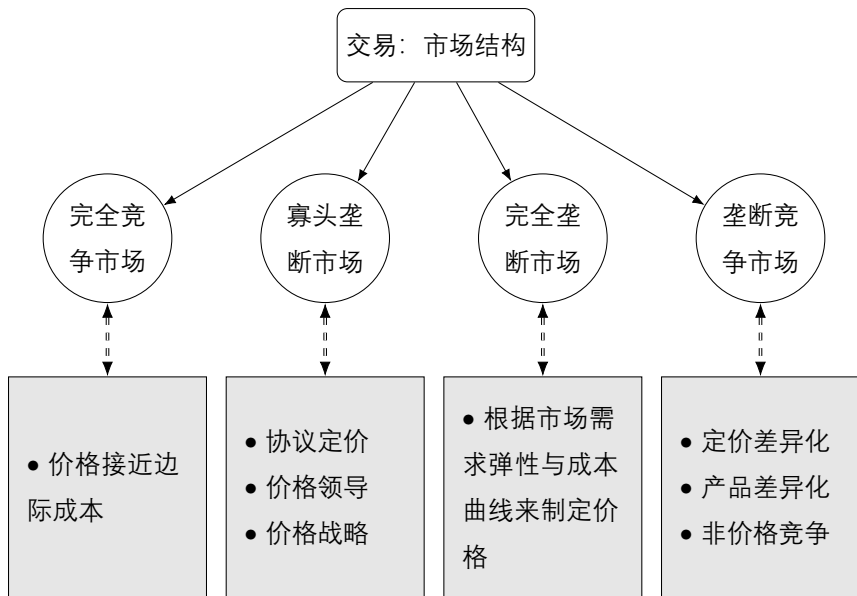
市场定位

二级市场是数据要素市场，介于一级市场（授权/采集）和三级市场（产品/服务）之间。

定价考量的特殊性：

- 相对一级市场：需额外考虑要素市场结构。
- 相对三级市场：需充分考虑数据获取成本。

数据二级市场定价策略：结构决定策略



案例分析：API 中介平台的差异化定价

案例背景：API 数据中介平台的垄断竞争策略

聚合数据、阿里云市场等平台向上游采购天气、交通、工商数据，打包成标准 API 出售。

理论与现实的映射：

- **充分考虑获取成本**：定价底线锚定上游（一级市场）采购授权固定成本，叠加接口维护边际成本
- **产品差异化**：同一天气数据，按时效、并发量、可用性分为体验版、开发者版、企业高可用版
- **定价差异化（阶梯定价）**：按调用次数（1 万/10 万/不限量包年）非线性定价，调用量越大单次价格越低

定价策略的深化与现状

理论 + 案例：Wind

理论

- **博弈论视角**：考虑买卖双方互动（如 Stackelberg 模型），关注消费者异质性
- **中国市场现状**：定价以**卖方报价**为主，头部企业往往成为**价格领导者**

案例：Wind

- **卖方报价主导**：金融数据壁垒高、转换成本高，Wind 拥绝对话语权
- **价格领导者**：Wind 涨价成行业标杆，iFinD 等对标折价跟随
- **异质性博弈**：对头部券商与小型私募进行价格歧视

数据二级市场定价的实践应用

案例：贵阳大数据交易所

全国第一家大数据交易所，截至 2023 年 3 月交易额达 6.3455 亿元。

三种定价模式：

- ❶ 协议定价：协商确定
- ❷ 固定定价：明码标价
- ❸ 实时定价：取决于样本量和单一样本指标项价值，系统自动浮动

创新工具：数据产品交易价格计算器

- **背景：**国家发改委价格监测中心指导
- **模型：**建立多因子修正成本模型
- **考量因子：**
 - 准确性、质量
 - 稀缺性、时效性
- **功能：**自动化计算建议价格
- **意义：**补全“报价—估价—议价”路径，促进公平贸易

第四节

数据三级市场定价

数据三级市场定价概述

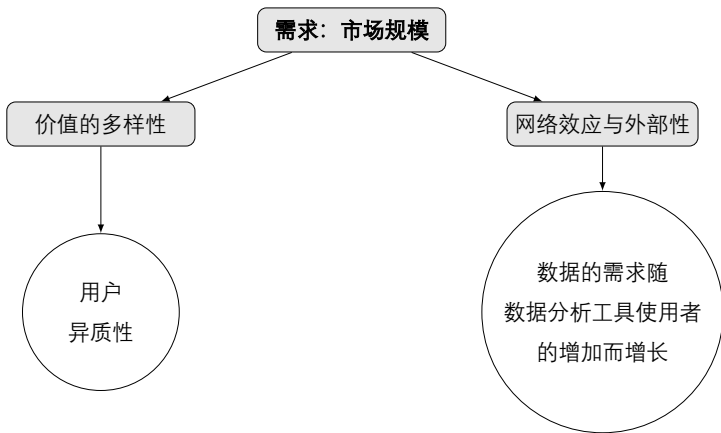
市场特征

三级市场是数据产品/服务的交易市场。与传统市场相似，但具有**极强的买方异质性**和复杂的市场结构。

核心策略：

- 主要是以**需求**为导向
- 需通过多种定价机制促进多主体参与

数据三级市场定价策略：需求驱动



具体定价机制：静态与动态

理论 + 案例：字节巨量引擎

理论：定价机制

- **静态**：固定定价（交易成本低）；差别定价（因买方异质性合理）
- **动态**：自动实时定价；拍卖式（适合一卖多买）；协商定价

案例：字节巨量引擎

- **拍卖式 (RTB)**：广告主实时竞价，出价与预估点击率最高者获展示
- **自动实时**：毫秒级自动撮合，底价随活跃度、时段动态浮动
- **差别定价**：不同行业对流量价值预期不同，成交价差异巨大

具体定价机制：信息熵定价法

理论 + 案例：Orbital Insight

理论：香农信息论 (Shannon, 1948)

- 信息熵表示排除冗余后的平均信息量
- 信息量相同时，事件不确定性越大，价值越高
- $H = - \sum_{i=1}^k p_i \ln p_i$

案例：Orbital Insight

- AI 分析卫星图像（统计停车场、监测原油储罐）
- 消除不确定性：财报发布前卫星数据越准，排除冗余越多，定价越高
- 买方异质性：同一张卫星照片，对评估供应链风险的企业可高价成交

具体定价机制：信息熵定价法

评价与性质

信息熵定价法的评价：

- **优点：**充分考虑稀缺性，强调信息量
- **局限：**操作难度大，适用范围有限
- **性质补充：**本质上也是一种动态定价法，且不能完全代表数据质量

数据三级市场定价的实践应用

- **主要模式：** 智能定制化服务（通过谈判和合作完成）
- **典型案例：** 某互联网平台为跨国快餐企业提供选址服务
- **合作细节：**
 - **供给：** 利用定位数据分析分布密度、消费能力
 - **需求：** 快餐企业获得市场洞察和选址依据
 - **定价考量：** 数据价值、分析复杂度、人力成本、竞争情况

本章要点总结

数据市场均衡的核心逻辑

数据市场均衡是供需作用的动态结果；数据具非竞争性、可复制性，均衡更依赖制度支撑。

1. 均衡的八大前提：

- **制度保障**：数据确权（基础）、数据脱敏（安全）、有效监管（秩序）
- **交易效率**：标准化（降低成本）、无交易成本（理想状态）
- **市场质量**：信息完全（消除悖论）、竞争性（地位对等）、无外部性

2. 动态调节机制：

- 超额供给 → 价格下降；超额需求 → 价格上升
- 技术进步推动供给增加（ S 右移），替代品涨价拉动需求增加（ D 右移）

本章要点总结

市场层次	核心逻辑	主要定价策略与工具
一级市场	成本导向	会计学成本法（直接/间接/作业成本）、加拿大统计局估算法
二级市场	结构导向	价格领先制（如 Wind）、边际成本定价（完全竞争）、博弈定价（Stackelberg 模型）
三级市场	需求导向	智能定制化服务、动态定价（RTB 拍卖、实时算法）、信息熵定价法

学习建议：从“结构”看“定价”

- 一级看投入：理解数据的“资源化”过程。
- 二级看对手：关注企业在要素市场中的市场势力。
- 三级看场景：关注买方异质性如何驱动价值的最终释放。

谢谢大家！