arXiv:2301.03489v1 [cs.CY] 9 Jan 2023

揭示并缓解网约车定价中的偏差，促进公平政策制定*

Nripsuta Ani Saxena **, [1]，Wenbin Zhang[2], 和 Cyrus Shahabi[1]

[1] 南加州大学，美国加利福尼亚州洛杉矶，邮编 90089[2] 美国密歇根州霍顿市密歇根理工大学，邮编 49931

摘要。 网约车服务因其便利性而迅速普及，但近期研究表明，其定价策略可能对部分乘客产生差别影响，例如居住在有色人种居民比例较高或贫困线以下居民较多的弱势社区中的乘客。由于这些社区因缺乏充足的公共交通而往往更依赖网约车服务，因此解决这一不公平问题至关重要。为此，本文通过设计适用的公平性度量及相应的公平定价机制，首次对网约车服务的公平定价开展了全面研究。我们提出的方法通过提供折扣（可由政府补贴），让弱势社区居民能够获得更多且更实惠的乘车服务。基于真实世界芝加哥出租车数据开展的实验验证了我们的理论发现，并为政府制定公平的网约车政策奠定了基础。

关键词： 公平性 · 网约车 · 定价 · 政府政策。

1 引言

自Uber于2009年首次推出以来，网约车公司已发展成为一项全球性服务，并成为我们生活中不可磨灭的一部分 [28,27]。尽管此类服务为消费者提供了更多选择和舒适体验，产生了积极影响，但这些服务的某些方面也可能带来负面的社会影响[17,8,5,32]。例如，研究表明，网约车服务的定价算法可能导致差别影响：有色人种比例较高、贫困程度较高且居民年龄较低的社区，与较高的车费显著相关 [17]。这种差别影响尤其令人担忧，因为它进一步损害了那些原本就更依赖这些服务的社区：这些社区由于公共交通连接不畅而更依赖网约车，并且长期以来在诸多方面遭受系统性边缘化。

* 会议前版本：非最终稿。

** 联系邮箱：nsaxena@usc.edu。

全国各地的市政府已经认识到，弱势居民可获得的交通出行选择存在差距，而网约车应用加剧了这种不公平。芝加哥市甚至已将减少所有居民在出行方面的不公平列为其发展计划 *On To 2050* [19] 中促进包容性增长的五项核心要素之一。许多地方政府正在制定政策，以帮助解决这一不公现象 [10,3,25,14]。遗憾的是，这些政策都无法缓解网约车公司定价机制所造成的额外差别影响。尽管这些政策朝着正确方向迈出了一步，但仅凭这些政策仍然不够，因为弱势社区缺乏充足且高效的公共交通覆盖 [7,11,22,33]，这导致其更加依赖网约车 [4]。事实上，研究人员分析了美国和欧洲六座大城市的Uber行程数据，发现20%至40%的网约车行程没有可行的公共交通替代方案 [4]。本研究旨在为政府提供一种技术，以帮助应对加剧的不公平。我们提出的定价机制中的折扣可由政府以补贴形式承担，从而减少对弱势社区造成的差别影响，这与针对低收入居民必要服务的其他补贴举措相类似 [10,3,25,14]。

尽管网约车服务定价机制的研究正蓬勃发展，但大多数研究着眼于提高司机与乘客匹配的效率，以及在供给不足时为行程定价 [24], ，或优化定价以调节平台上的需求 [31]。鲜有研究探讨如何通过公平定价来减少差别影响。值得注意的是，研究发现，美国主要机场至酒店的 Uber 行程车费与酒店客房价格显著相关 [5]，但这些研究并未试图解决网约车中的这种价格歧视。研究也考察了网约车平台上司机能否获得公平收益，但未考虑乘客的公平价格 [23]。在 [13],中，研究探讨了这样一种假设情境：网约车服务利用客户的个人数据，为每位客户确定定制价格。虽然这有助于解决一种形式的歧视，但并未明确解决当前定价技术造成的差别影响。

尽管揭示并缓解网约车应用定价机制的差别影响十分重要且必要，但要以此为政府制定公平的网约车政策提供基础，仍面临独特的挑战。首先，可负担性存在差异。不同社区的平均收入水平可能不同，因此，对一个社区居民来说可负担的价格，对另一个社区的居民而言可能并不可负担。公平、公正的解决方案应根据可负担性为行程定价，确保没有人因价格过高而无法使用网约车。其次，还存在与数据相关的挑战，因为理想情况下对该问题进行整体建模所需的全部数据并不可得。例如，价格弹性——在本文中指行程数量随行程价格变化而变化的程度——是未知的，行程的动态加价水平也是未知的。因此，必须对这些量进行估计，并且有时还需要作出假设。

为应对上述挑战，本文提出了公平定价的新定义及相应的公平定价策略，以解决网约车中的歧视问题。具体而言，本文的主要贡献包括：i）我们运用经济学文献中的价格弹性和消费者剩余概念，探讨不同群体对行程的可负担性，以及定价与实际行程数量之间的关系，从而正式定义网约车中的偏差。ii）我们提出了一种公平定价策略，能够有效减少差别影响，并且可灵活适应多个场景中的不同要求。iii）我们使用来自芝加哥的真实世界网约车数据和人口普查数据开展了多项实验，将我们的定价机制与其他常用定价模型进行比较，从而验证了该方法在现实世界中的实用性。

2 网约车中的偏差

网约车定价存在偏差，会让那些可能本已经济拮据的人支付更高的价格。为解决网约车服务黑箱定价算法造成的这种差别影响，我们将网约车中的偏差视为弱势群体与非弱势群体的平均乘车次数存在显著差异，并考察相应的可负担性。

为正式界定弱势群体与非弱势群体，我们采用芝加哥大都会规划局（CMAP）的分类。该机构将芝加哥地区的人口普查区划分为经济脱节地区（EDAs），这些地区与经济增长和繁荣脱节，并可能正经历投资撤离 [18]。换言之，经济脱节地区是低收入家庭、少数族裔或英语能力有限者高度集中的社区。约三分之一的城市居民生活在经济脱节地区 [18]。为便于阅读，本文余下部分将**经济脱节地区**与弱势社区这两个术语互换使用，也将**非经济脱节地区**与非弱势社区这两个术语互换使用。我们进一步将**经济脱节地区行程**定义为起点或终点位于经济脱节地区的行程（或乘车）。**非经济脱节地区行程**则指起点和终点均不在经济脱节地区的行程（或乘车）。

3 揭示网约车偏差

3.1 相对乘车能力

与美国平等就业机会委员会（EEOC）用于评估差别影响的 p%-规则 [2] 一致，我们引入相对乘车能力 （R^2）评分，用于量化前文所述的网约车偏差，该偏差表现为不同群体的平均乘车次数存在显著差异。具体而言， p%-规则规定，如果某一群体的选择率低于选择率最高群体的选择率的 p%，则选择率存在实质性差异，可能构成差别影响。相应地， R^2 可按以下方式进行数学定义：

$$R^2 = \frac{\min\{d_1, \ldots, d_i\}}{\max\{\neg d_1, \ldots, \neg d_j\}} \tag{1}$$

其中， d_i 是弱势群体 i居民的平均行程数，而 $\neg d_j$ 是非弱势群体j居民的平均行程数。在不公平的网约车服务中， $\neg d_j$ 通常大于 d_i ， R^2 越高，模型就越公平。

3.2 可负担性

接下来，我们将探讨量化网约车偏差的另一种方式：网约车可负担性的差异。可负担性可以通过消费者剩余[30], 这一经济学概念来衡量，它反映消费者为某种产品或服务支付的意愿与其价格之间的差额。如果产品或服务的价格低于消费者愿意支付的金额，消费者剩余便为正；否则则为负，这种情况可能出现在某些必需品上（例如食品或救命药物）。换句话说，消费者剩余越高，消费者就越容易负担得起该产品或服务。

然而，公开发布的网约车数据中并未提供计算网约车消费者剩余所需的信息，例如查看过报价但未提出请求的人数及其所看到的价格。为绕过这一难题，我们改为计算价格弹性 *(E_p)* ，即产品或服务的需求随其价格变化而变化的程度，据此估算消费者剩余：

$$E_p = \frac{\delta_Q}{\delta_P} \tag{2}$$

其中， δ_Q 是对某项产品或服务的需求量变化百分比， δ_P 是其价格变化百分比。需要说明的是，在下文介绍公平定价机制时，我们也会用价格弹性来估算价格变化可能导致的行程数量变化。

价格弹性通常在商品或服务价格发生变化的点上进行考察。然而，网约车行程并非按固定每英里费率计价，因此，简单地研究不同价格下的行程并不理想，因为两趟行程之间的价格差异可能由多种因素造成。因此，我们研究因网约车平台计算动态加价水平的方式不同而显示不同价格的相似行程。动态加价水平是一种乘数，在需求旺盛或供给不足时，用于乘以并提高行程的估算价格。此外，网约车平台通常会为行程计算连续的动态加价水平，但为了简洁易懂，会向消费者显示离散数值。例如，动态加价水平计算结果为 1.449 的行程，显示的动态加价将是 1.4 倍；而动态加价水平为 1.451 倍时，显示的动态加价将是 1.5 倍。我们利用这些离散化的动态加价不连续点附近的回归不连续设计

动态加价点，估算动态加价水平从 1.2 倍升至 1.3 倍、从 1.3 倍升至 1.4 倍、从 1.4 倍升至 1.5 倍等情况下的价格弹性。我们采用线性概率模型（普通最小二乘回归的一种特殊形式），这种模型在经济学中十分常用，适用于结果变量为二元变量、且自变量可以是二元或连续变量的情形。我们针对每个动态加价不连续点拟合以下线性概率模型回归，其中 行程 表示某个特定请求是否会产生一趟行程，并纳入动态加价不连续点两侧的所有行程：

$$Ride = \beta_0 + (\alpha * i_1 * i_2) + (\beta_1 * i_1) + (\beta_2 * (1 - i_1) * i_2) + (\beta_3 (1 - i_2) * x_1) + (\beta_4 * i_2 * x_1) + \epsilon \tag{3}$$

其中， α 表示不连续点附近行程数量的减少， i_1 是一个决策变量，用于表示该行程的动态加价是否处于动态加价价格不连续点的 0.01 范围内； i_2 是一个决策变量，用于表示该行程的动态加价是否位于价格不连续点右侧（即 其离散化动态加价水平高于离散化动态加价点，因此，连续动态加价水平为 1.451、离散化为 1.5 倍动态加价的行程，其值为 1；而连续动态加价水平为 1.449、离散化为 1.4 倍动态加价的行程，其值为 0）； x_1 为实际（未经离散化的）动态加价值， β为系数， ϵ 为误差。为了在每个跳跃水平上计算这一结果，我们使用位于不连续点两侧的行程。由于我们针对每个动态加价不连续点分别进行计算，因此， α 有助于捕捉动态加价水平导致价格变化所引起的行程数量变化。我们的消费者剩余和价格弹性计算采用了 [6] 在一篇经济学论文中详述的方法，该论文计算了美国四个主要市场中 Uber 的消费者剩余。总而言之，我们首先使用在公式 3 中采用的回归方程，估算不同动态加价水平下的价格弹性，然后利用这些价格弹性计算消费者剩余。

我们在每个价格不连续点处计算的回归不连续设计（公式 3）有助于估算 α，该值表示由于价格变化，在该不连续点发生的行程数量变化。随后，我们可以利用此 α，如下所示计算价格不连续点处的价格弹性：

$$E_p = \frac{\delta_Q}{\delta_P} = \frac{\frac{\alpha}{N_p}}{\delta_P} \tag{4}$$

其中， N_p 是以特定价格 p发生的行程比例。

此时，我们又遇到了一个与数据相关的挑战：网约车数据中没有动态加价水平。为了解决这一问题，我们采用随机抽样一致性（RANSAC）回归来确定动态加价何时生效以及动态加价水平是多少。当数据包含离群值时，RANSAC 可利用数据集中的一部分“内点”（即不是离群值的数据点）来稳健估计模型参数。使用 RANSAC 的直觉如下。根据 Uber 的说法，当同时请求行程的人数异常多时，动态加价就会生效 [26]。换言之，当数量非标准或远高于正常水平的客户试图同时预订行程时，就会出现动态加价。如果所有

以 1.0 倍（无动态加价）进行的行程都是标准行程，即内点，那么采用动态加价的行程就会被视为离群值。一旦我们确定了每次行程生效的动态加价水平（1.0 倍或更高），就可以在回归连续性设计中使用动态加价水平两侧的行程来估计价格弹性。

算法1 可负担性

输入： 某一群体的网约车数据（d）
输出： 该群体的消费者剩余总额
1：计算行程的动态加价水平： $Surge \leftarrow RANSAC(d)$
2：**对于** *surge level* $s \leftarrow 1.1x$ 到 N **执行**
3： 计算 *regression discontinuity design* 围绕 s （公式 3）
4： 计算 $E_p \leftarrow \frac{\frac{\alpha}{N_P}}{\frac{\delta}{P}}$ 在 s（公式 4）
5: **结束循环**
6: **循环** 动态加价为 s 的行程 $\leftarrow 1.0x$ 到 N **执行**
7： 消费者剩余 $\Delta_c \leftarrow \sum_{i=s+0.1x}^{N} E_{p_i} * numTrips_i * \left(\left(\frac{i-s}{s}\right) * 100 \right) * avgFare_s$
8： **结束**

再次说明，消费者剩余是指消费者愿意为某种产品或服务支付的价格与实际被收取的价格之间的差额。算法1概述了其计算步骤。具体而言，我们针对某一群体（经济发展区居民或非经济发展区居民）的网约车数据进行分析，首先估算行程的动态加价水平（第1行）。接着，我们围绕每个连续的动态加价水平进行回归不连续设计，并估算该动态加价水平下的价格弹性（第2至5行）。在考察动态加价为1.0倍（即没有动态加价）的行程时，我们采用下一个动态加价水平1.1倍处的价格弹性，计算如果向看到1.0倍动态加价的客户展示的是1.1倍动态加价，本会发生多少行程。然后，我们将这一潜在行程数量乘以票价差额（本例中为10%）以及实际支付1.0倍动态加价时的平均价格。这就是被收取1.0倍动态加价的乘客在动态加价水平最高至1.1倍范围内的消费者剩余。我们随后对每一对连续的动态加价水平重复这一过程，并将结果加总，得出消费者剩余总估值（第6至8行）。最后，算法1输出该群体行程的消费者剩余。

4 公平定价机制

针对网约车量身定制了公平性指标后，我们现在详细介绍有助于减少弱势社区所受差别影响的公平定价机制。我们着重为 EDA 行程定价，以帮助减轻 EDA 居民所受的差别影响，从而帮助芝加哥市解决居民出行不平等方面的担忧 [19]。此外，为了灵活应对不同情况，我们定义了两种定价机制变体：可变折扣和固定折扣。采用可变折扣时，EDA 行程获得的折扣可能会因一天中的时段、需求等多种因素而有所不同

而采用另一种固定折扣方式时，所有 EDA 行程获得的折扣都相同（例如 15%）。这种双重方案体现了平台和/或政府在政策制定方面的需求。具体而言，平台可能需要采用固定折扣方案，因为这种方案更容易实施，也更直接；另一方面，政府可能认为，根据当时的情况为不同行程提供不同折扣是合理的，而实施此类系统可能带来的复杂性并非主要顾虑。

4.1 可变折扣：FairRide

我们首先提出一种名为 FairRide 的新定价机制。FairRide 背后的理念是，定价机制应考虑不同乘客负担乘车费用的能力差异，以减少对社会中处境最不利群体造成的差别影响。为此，我们提出一种机制，分别为经济发展区区域的行程定价。我们重点关注经济发展区行程，并运行多元回归模型，仅针对起点或终点位于经济发展区的行程（经济发展区行程）确定价格。只考虑经济发展区居民的行程，可以使行程定价与乘客的支付能力相符。在第 4 节中，我们将 FairRide 与文献中常用于定价的机器学习模型进行比较 [9,21,15,16,29,1], ，发现只考虑经济发展区行程时，大多数模型为经济发展区居民带来的行程数量都多于当前定价机制，并能提高相对乘车能力（R^2）；而 FairRide 的表现优于所有这些模型。我们还实现了一个朴素基线，即对所有行程一律给予 5 美元折扣。

4.2 固定折扣：FixedFairRide

第二种公平定价机制称为$FixedFairRide$，它通过求解一个优化问题，在维持或增加收入的同时，以一致的固定折扣（δ）优惠所有 EDA 行程，从而使 EDA 行程数最大化。换言之，我们要确定 EDA 行程的固定折扣额度，以确保采用固定折扣政策。我们的总体目标是确定折扣 EDA 行程的最优值 δ，使相对乘车能力提升，并使 EDA 行程数最大化。其数学形式如下：

$$\eta = \max_{\delta} \sum_{t} \left[\sum_{l} N(\delta * p) \right] \tag{5}$$

其中， η 表示乘车总数， t 表示特定时间段，l 表示 EDA 地点对， δ 表示 EDA 行程的折扣， p 表示 EDA 行程的价格， N 表示该价格下的 EDA 行程数。

现在，我们根据两种可能的情形为该优化问题引入两项约束。在第一种情形下，政府以补贴的形式承担给予 EDA 居民的折扣。因此，平台和司机获得的收入与目前每次行程获得的收入相同

，而 EDA居民则可按低于原始方案价格的折扣价乘车，差额由政府补贴补足。这里无需考虑收入，因为平台收取的价格并未改变。也无需考虑行程对司机而言是否可行，因为司机收到的车费与折扣前完全相同。不过，政府可能希望为其提供的折扣设定上限（例如 30%），这可以作为该情形下的一项约束：

$$0 < \delta < n \tag{6}$$

其中， n 是政府愿意补贴的最高折扣。

在第二种设定中，我们假设平台愿意自行承担折扣，以帮助缓解差别影响。然而，平台可能不希望总收入下降。因此，另一种约束为：

$$Revenue \geq r, \tag{7}$$

$$0 < \delta < 1 \tag{8}$$

其中， r 表示按当前定价获得的总收入，而 δ 必须大于0，这表示EDA 行程的折扣始终为正。此外，我们发现，只需增加大量折扣力度很大的新行程，就能使收入超过 r 。然而，价格低于某个水平时，司机可能会觉得接单不划算，因为收入可能不足以支付成本并赚取利润。为考虑行程对司机而言是否可行，我们对收入设置另一项下限，确保司机不会面临接单却一无所获的风险，因此，设置下限后的约束为：

$$Revenue \geq max(r,\ \eta * p_{min}) \tag{9}$$

其中， r 是按当前定价获得的总收入， η 是行程总数。p_{min} 是行程的最低价格，可根据平台认为对行程重要的因素（如行程时间、距离、需求等）确定。

动态加价水平导致的价格不连续性使其成为一个非凸问题，求解并不简单。因此，我们在改变 δ 的取值时采用网格搜索 [12]。

5 实验评估

在本节中，我们使用芝加哥市的现实世界数据开展实验。据我们所知，我们使用该数据集，是因为它是唯一公开可用、包含行程价格（或车费）信息的数据集，而这些信息对于解决网约车定价中的偏见问题至关重要 [20]。来自在芝加哥市内运营的所有网约车平台的数据在发布前均经过汇总、标准化和匿名化处理。数据未标明某次行程由哪个网约车平台提供服务。对于每次行程，数据集提供行程开始时间（或“上车”时间）和行程结束时间（或“下车”时间），时间均四舍五入至最接近的 15

分钟；行程的上车和下车地点（精确到芝加哥人口普查区或社区区）；行程时长（秒）；以及行程车费（四舍五入至最接近的 2.50 美元）。芝加哥市界以外的上车或下车地点信息不可用。司机和乘客的身份均未标明。我们考察 2021 年 1 月至 10 月的行程，并重点关注未获准作为“拼车”行程的行程。自 2020 年初新冠疫情暴发以来，大多数网约车服务都关闭了拼车选项。

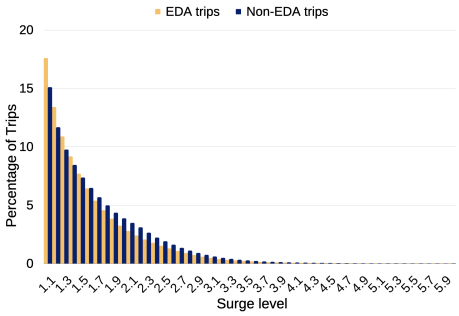
最后，我们将该数据集中的上车点和下车点坐标与芝加哥市发布的空间数据集进行空间连接，该数据集标识了芝加哥的EDA区域，以帮助识别数据中起点或终点位于EDA区域或非EDA区域内的行程。

5.1 深刻的偏差

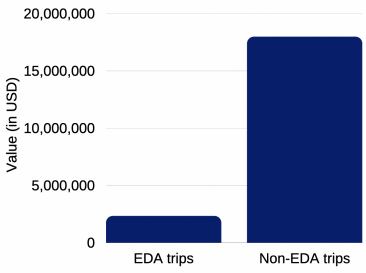
我们使用第4节提出的指标，揭示网约车公司的当前定价机制中固有的歧视。我们发现 R^2 =0.33。如此低的 R^2 表明真实偏差普遍存在，并且可能对弱势社区造成严重的差别影响。我们还将在下文观察到消费者剩余和可负担性方面的显著差异。

接下来，我们通过消费者剩余考察可负担性。我们首先使用RANSAC回归估计动态加价何时生效，以及行程的动态加价水平。图1a展示了总体趋势：通常，随着动态加价倍数增加，发生的行程数量会减少。虽然我们没有动态加价水平的真实值，无法衡量模型预测的准确性，但我们观察到的趋势与 [6],中动态加价水平的趋势一致；该研究由于与Uber合作开展，因此获得了真实值。我们还可以从图1a中得出以下观察结果。首先，与非经济脱节地区行程相比，更大比例的EDA 行程发生在较低的动态加价水平下。其次，当动态加价水平为4.4倍、平均行程价格为62.06美元时，EDA 行程数量降至1,000以下。但对于非经济脱节地区行程而言，直到动态加价水平达到5.9倍、平均价格为97.02美元时，行程数量才降至1,000以下。现在我们已经知道行程的动态加价水平，便使用第3.2节中的方程计算价格弹性和消费者剩余。当动态加价水平非常高（$>$ 8.0倍）且行程数量极少（通常为 $<$100）时，除少数情况外，EDA 行程和非经济脱节地区行程的价格弹性均为负。这表明，随着价格上涨，需求通常会下降，反之亦然。我们发现，非经济脱节地区行程的消费者剩余远高于 EDA 行程。仅考察行程数量合理的动态加价水平（在最高动态加价水平下，我们观察到每个水平的行程数量都少于100，因此不将其纳入以下图表），在我们考察的10个月期间，非经济脱节地区行程的消费者剩余总额为17,972,629.829美元，而 EDA 行程的消费者剩余为2,326,555.876美元（图1b）。将其分别除以各类别所服务的行程总数后，非经济脱节地区行程的平均消费者剩余为67.11美元，EDA 行程的平均消费者剩余为36.76美元。

In 2016, [6]估计，2015年美国全国范围内UberX的消费者剩余为68亿美元，因此按其估算，美国所有城市一天的消费者剩余将达1800万美元。此外，这些计算基于7年前的数据，而现在的数值很可能更高。相比



(a) 动态加价水平下行程的占比 2021 年前十个月。



(b) 行程的消费者剩余。

图 1：动态加价水平与可负担性（Δ_c）。

这些数字，并考虑到芝加哥是美国的一座主要城市，我们的估算很可能低于实际值。因此，在非经济发展区（EDA）区域开始或结束行程的乘客，比在EDA社区开始或结束行程的乘客更能负担得起网约车服务。

5.2 定价机制

FairRide 我们将 FairRide 与用于定价的机器学习模型进行比较[9,21,15,16,29,1]。我们还与一个朴素基线进行比较，该基线是对所有 EDA 行程统一提供 5 美元折扣。

我们发现，使用按区域组织的数据进行训练，会使约 35.7% 的 EDA 行程价格降低，并使我们考察的 2021 年期间起点或终点位于经济发展区的行程数量增加 35.59%。换句话说，随着行程变得更加可负担，另有 1,803,514 次行程的起点或终点位于经济发展区。结果是，借助 $FairRide$，相对乘车能力 (R^2) 提高了 35.6%，达到 0.457；可负担性 (Δ_c) 提高了 22.5%，达到 2,850,552.77 美元。
FixedFairRide 随后，我们使用 p_{min} （一次行程的最低价格）的不同取值测试 $FixedFairRide$，并观察到随着 p_{min} 增加， δ （行程折扣）的值降低（表 2）。

因此，随着每次行程的最低价格提高，折扣就会降低。 p_{min} 的每个取值所对应的行程数量都多于当前定价机制；这些取值也都会带来比$FairRide$ 更多的行程，直至 $p_{min} = 15$，届时行程数量为 6,824,375。EDA 居民的相对乘车能力（R^2）和可负担性（Δ_c，即消费者剩余）均显著提升。在我们实验所用的数据集中，最低票价为 2.50 美元，而行程比例最高（约 14%）的票价约为 10 美元。因此，在现实中，远高于 10 美元的最低行程价格很可能并不实用。

表1：不同模型的行程数量、 R^2和可负担性。

模型	η （变化百分比）	R^2 （变化百分比）	Δ_c （变化百分比）
原始方案	5,066,849	0.337	2,326,555.87
FairRide	**6,870,363**	**0.457（+35.60）**	**2,850,552.77（+22.52）**
Mohd等人 (2020) [16]	6,867,535	0.456 (+35.31)	2,722,461.33（+17.01）
Miao (2017) [15]	6,868,152	0.456 (+35.31)	2,818,648.50 (+21.15)
Wolk (2020) [29]	6,864,923	0.456 (+35.31)	2,534,351.94 (+8.93)
基线 (-$5)	6,618,494	0.440 (+30.56)	2,739,784.93（+17.76）
Gu 等（2020） [9]	6,843,205	0.455 (+35.01)	2,612,239.28 (+12.27)
Rathan 等（2019） [21]	6,385,843	0.424 (+25.81)	2,667,079.84（+14.63）
Alkhatib 等（2013） [1]	6,608,789	0.439 (+30.26)	2,476,128.68 (+6.42)
FixedFairRide($p_{min} = 12$)	**8,142,519**	**0.541 (+60.53)**	**3,201,963.73 (+37.62)**

表 2：折扣 δ 不同的最低行程价格 p_{min}。

p_{min}	5	6	7	8	9	10	11	12	13	14	15
δ	0.77	0.72	0.67	0.62	0.57	0.52	0.47	0.42	0.36	0.30	0.24

6 结论

尽管近年来有不少研究探讨了网约车服务的定价，包括如何为司机提供公平补偿，以及如何通过保险政策防范网约车利用个人信息进行定价，但没有研究关注如何通过更公平的定价来减少对弱势居民的差别影响。我们提出了适用的公平性指标，以揭示网约车定价中的内在偏见，并设计了一种灵活的定价机制，让定价更公平，同时让公共交通服务不足地区的弱势居民能够以更低的费用出行。实证实验揭示了现有网约车定价造成的深刻偏见，并表明，采用我们的机制为行程定价，能够提供更可负担、更公平的网约车服务，从而为政府制定公平的网约车政策提供助力。

参考文献

1. Alkhatib, K., Najadat, H., Hmeidi, I., Shatnawi, M.K.A.：使用 k 近邻（knn）算法进行股票价格预测。《国际商业、人文与技术杂志》**3**(3), 32–44 (2013) 2. Biddle, D.：《不利影响与测验验证：有效且站得住脚的就业测试实务指南》。Routledge (2017) 3. Bus, L.G.：公共交通通票补贴计划（2018年3月），https://pw.lacounty.gov/transit/TransitPass.aspx 4. Cats, O., Kucharski, R., Danda, S.R., Yap, M.：超越二分法：网约车如何与公共交通竞争并形成互补。《Plos one》**17**(1), e0262496 (2022)

5. Chang, Y., Winston, C., Yan, J.：Uber 是否通过价格歧视使出行者受益？(2021) 6. Cohen, P., Hahn, R., Hall, J., Levitt, S., Metcalfe, R.：利用大数据估算消费者剩余：以 Uber 为例。技术报告，美国国家经济研究局 (2016) 7. Garrett, M., Taylor, B.：重新审视公共交通中的社会公平。《伯克利规划杂志》**13**(1) (1999) 8. Ge, Y., Knittel, C.R., MacKenzie, D., Zoepf, S.：交通网络公司中的种族与性别歧视。技术报告，美国国家经济研究局 (2016) 9. Gu, S., Kelly, B., Xiu, D.：通过机器学习进行实证资产定价。《金融研究评论》**33**(5), 2223–2273 (2020) 10. House, T.W.：平价连接计划福利（2021年3月），https://www. whitehouse.gov/getinternet/utm$_$source= healthsherpa 11. Kramer, A.：《负担不起的城市：北美城市的住房与公共交通》。《Cities》**83**, 1–10 (2018) 12. Lerman, P.：通过网格搜索拟合分段回归模型。《皇家统计学会杂志：C辑（应用统计）》**29**(1), 77–84 (1980) 13. Lu, Y., Qi, Y., Qi, S., Li, Y., Song, H., Liu, Y.：拒绝价格歧视：网约车服务价格审计的去中心化自动化激励机制。《IEEE移动计算汇刊》 (2020) 14. Metro, L.：低收入者便捷票价计划（LIFE）（2018年3月），https://www.metro.net/ riding/life/ 15. Miao, H., Ramchander, S., Wang, T., Yang, D.：原油价格预测的影响因素。《能源经济学》**68**, 77–88 (2017) 16. Mohd, T., Jamil, N.S., Johari, N., Abdullah, L., Masrom, S.：房价预测房地产建模技术综述。《描绘东盟商业与社会科学的可持续未来》，第321–338页 (2020) 17. Pandey, A., Caliskan, A.：网约车经济中人工智能偏见对价格歧视算法的差别影响。载于：2021年AAAI/ACM人工智能、伦理与社会会议论文集，第822–833页 (2021) 18. for Planning (CMAP), C.M.A.：经济脱节地区（2018年3月），https://www.cmap.illinois.gov/2050/maps/eda 19. for Planning (CMAP), C.M.A.：On To 2050（2018年3月），https://www.cmap. illinois.gov/2050/principles 20. Quy, T.L., Roy, A., Iosifidis, V., Zhang, W., Ntoutsi, E.：公平性感知机器学习数据集综述。《数据挖掘与知识发现》(2022) 21. Rathan, K., Sai, S.V., Manikanta, T.S.：使用决策树和回归技术预测加密货币价格。载于：2019年第三届电子与信息学趋势国际会议（ICOEI），第190–194页。电气电子工程师学会 (2019) 22. Sanchez, T.W.：贫困、政策与公共交通。《交通研究A辑：政策与实践》**42**(5), 833–841 (2008) 23. S¨uhr, T., Biega, A.J., Zehlike, M., Gummadi, K.P., Chakraborty, A.：双边市场中重复匹配的双边公平：网约车平台案例研究。载于：第25届ACM SIGKDD知识发现与数据挖掘国际会议论文集，第3082–3092页 (2019) 24. Tong, Y., Wang, L., Zhou, Z., Chen, L., Du, B., Ye, J.：空间众包中的动态定价：一种基于匹配的方法。载于：2018年国际数据管理会议论文集，第773–788页 (2018)

25. 美国交通部：Tran serve（2018年3月），https://www.transportation.gov/sites/dot.gov/files/2022$\protect\discretionary{\char\hyphenchar\ font}{}{}$01/FY2022$_$TIGTA$_$Integrity$_$Awareness$_$Training$_$Presentation$_$.01.06.2022.pdf 26. Uber：动态加价（2018年3月），https://www.uber.com/us/en/drive/ driver-app/how-surge-work 27. Wallsten, S.：共享经济的竞争效应：Uber 如何改变出租车行业。技术政策研究所**22**，1–21（2015年）28. Willis, G., Tranos, E.：利用“大数据”了解 Uber 对纽约市出租车行业的影响。《旅行行为与社会》**22**，94–107（2021年）29. Wo-lk, K.：用于短期加密货币价格预测的高级社交媒体情感分析。《专家系统》**37**(2)，e12493（2020）30. Xiao, P., Xiao, R., Liang, Y., Chen, X., Lu, W.：政府补贴计划的影响：超越可负担性的可及性。《管理科学》**66**(7)，3211–3233（2020）31. Yuan, E., Van Hentenryck, P.：用于网约车服务质量的实时定价优化。载于：国际人工智能联合会议。第3742–3748页（2021年）32. Zhang, W., Ntoutsi, E.：FAHT：一种自适应公平感知决策树分类器。载于：国际人工智能联合会议（IJCAI）。第1480–1486页（2019年）33. Zhang, W., Weiss, J.C.：具有删失的纵向公平性。载于：AAAI人工智能会议论文集。第36卷，第12235–12243页（2022年）