Price discrimination and Bias in the Ridesharing Industry

Amar Dholakia, Erwin-Jeffrey Komguem, Hyun Woo Lee
University of Toronto
STA2052: Statistics, Ethics and Law
November 15, 2021

Abstract

Price discrimination, simply put, is the differential pricing of identical goods to different persons. Despite the name, price discrimination does not necessarily imply the practice is discriminatory against specific protected classes. Specifically, we investigate the existence of price discrimination in the Boston, MA Uber rideshare market, and discuss its impact on different neighbourhoods taking into context their demographics. Our central arguments rest on whether price discrimination adversely affects protected classes, and if this can be considered discriminatory, by Moreau's definition of discrimination. Furthermore, we reconstruct an approximation of Uber's dynamic pricing algorithm using linear mixed models, and investigate any potential algorithmic bias in prediction, in terms of differential pricing.

1. Introduction

Ridesharing services like Uber and Lyft employ 'dynamic pricing algorithms' that vary prices of similar or identical rides based on present driving conditions, and demand-supply tradeoff. Per Uber, the dynamic algorithm adjusts price using both historic and current data on trip time, distance, and weather conditions that may increase rider demand (Uber, 2021). By temporarily inflating prices, and subsequently driver rates, the algorithm hopes to incentivize more drivers to offer their services to meet the demand. In summary, dynamic pricing algorithms aim to maintain a supply-demand equilibrium to ensure relatively continuous service in a variety of adverse or high-demand conditions.

Dynamic pricing algorithms potentially introduce an avenue for price discrimination. In short, price discrimination is the differential charging of price to different individuals for similar goods or services. Price discrimination need not be discriminatory in the legal sense. Per the Ontario Human Rights Code,, as outlined in "Freedom From Discrimination", while the Services section states that all individuals have equal opportunity to goods and services regardless of demographics, it does not explicitly state that they should all benefit from equal price (Government of Ontario, 1990).

From the consumer perspective, price discrimination can be argued to be ethical and even moral. The simplest example of price discrimination is the offering of student discounts for inelastic goods, such as clothing, or mandatory paid course material. Students who would normally be priced out of the market and those who would ordinarily be inclined to buy the good are now incentivized or even enabled to do so. In theory, consumers who are willing to be paid more for certain goods can be charged more to effectively subsidize consumers who are less willing or able to pay for the same good.

This was empirically demonstrated by Ian Fillmore, who modelled how the U.S. college system uses applicant demographic information to gauge willingness to pay, redistributing tuition of wealthier students as financial aid for less wealthy students. (2016). A final empirical example follows that of prescription drug price differentials between countries. Lichtenberg (2010) found that drug prices were generally positively correlated with drug availability and usage between countries, identifying price discrimination. He argues that price discrimination allows fewer wealthy markets access to goods and services that improves population health outcomes with increased use.

Conversely, price discrimination while not legally discriminatory, may inflict adverse effects to certain protected groups. Drawing from another health industry example Kaplan and O'Neill (2020) illustrate the deleterious effects of price discrimination of private insurance and hospital fees. It is suggested that hospitals benefit tremendously when providing highly inelastic medical services to clients with private coverage, a group that is disproportionately wealthy and white. While protected classes, such as minorities or the impoverished, are not directly discriminated against, they are priced out and the provision of quality healthcare is ultimately segregated. As noted by Lotters (2005), groups with poorer health outcomes tend to suffer from lower productivity, earn on average less, and remain cyclically priced out of the private health market (Tompa, 2002).

Based on our discussion above, we would like to investigate whether the Uber dynamic pricing algorithm imposes price discrimination. We emphasize appraisal on candidate predictive algorithms; specifically, we ask whether there are differentials in the mean predicted price of ridesharing based on geodemographical information, and whether the model performs equally well on different groups (neighborhoods). We evaluate recent publicly available models and frameworks that aim to quantify bias in rideshare pricing algorithms. Specifically, we compare the mean price prediction and prediction variability in a naïve regression algorithm, presented by Todd Schneider (Schneider, 2016), and an improved linear mixed model that incorporates neighborhood and vehicle type information.

We find that in both the historical data and future predictions, patrons hailing rides from lower-income neighborhoods tend to pay higher fares compared to those hailing rides from higher-income areas. We note that both models predict with more error in lower-income neighborhoods than higher-income, quantified by root mean square error (RMSE). These results suggest the existence of price discrimination in the Boston rideshare market.

These results lead us to a discussion in which we present literature and arguments that support that price discrimination in our scenario is discriminatory against patrons in adversely affected neighborhoods. Our model with Uber's pricing data depicts evidence in violation of normatively extraneous features, which must be protected against. We suggest a counter argument that shows user-level discrimination cannot occur, and that neighborhood-level selection is not actually discriminatory, but beneficial to patrons in adversely affected neighborhoods. Finally, we suggest that the dynamic pricing algorithm in question is biased due to the biased nature of the data that was used to train it. We explore some possible solutions to price discrimination and machine learning bias, which brings us to understand that the only way to truly mitigate AI bias is to enforce algorithmic transparency.

2. Methodology

Data collection
The dataset consisted of records of ridesharing in the greater city of Boston from November 26, 2018 to December 18, 2018. We are primarily interested in the features that potentially influence price, including, but not limited to, price surge multiplier, date, timestamp (categorized into day, evening, night), distance, source, destination, and vehicle type. Aside from ride specific features, we were also motivated to analyze if geodemographical features have an influence in price. These include income, crime rate, population, age as divided per neighborhoods in Boston.

Data Processing

We were principally interested in the Uber rides, although this analysis can easily be extended to include Lyft or other ridesharing services. We filtered out any non-Uber rides, including any rides designated as 'taxicab'. Furthermore, we did not include weather information, although the analysis could be extended to include this.

To adjust for supply and demand differentials, we standardized the true price by dividing it by the surge multiplier variable. We named this variable the 'adjusted price' and later used it as our principal response variable fit and predicted in our models. Our dataset did not contain information about the duration and/or expected time of ride, which is an important predictor in the Uber pricing algorithm, and the model published by Todd Schneider. We were unable to model expected time of ride and is an inevitable limitation of our replication of his model.

Modelling

We fit and compared two models the naïve OLS published by Todd Schneider (Schneider, 2016) and our own improved model. The naïve OLS model can be represented as:

- $Y_i = \beta_0 + x_i\beta_1 + \epsilon_i$, where:
    - Y_i is the fare price in USD\$, where $Y_i \sim N(0, \sigma_i^2)$
    - β_1 is the distance in miles
    - ϵ_i is the error term, where $\epsilon_i \sim N(0, \sigma_i^2)$

Our improved model can be represented as:

- $boxcox(Y_{ij}) = \beta_0 + x_{ij}\beta_{1j} + x_{ij}\beta_{j2} + U_j + \epsilon_i$, where
    - $boxcox(Y_i)$ is the surge-adjusted, boxcox fare price in USD\$, where $boxcox(Y_i) \sim N(0, \sigma_i^2)$
    - β_{1j} is the distance in miles
    - β_{2j} is the factor neighbourhood (Back Bay contrast)
    - U_j is the mean random effect of the factor vehicle type, with $U_j \sim N(0, \sigma_{uj}^2)$
    - ϵ_i is the error term, where $\epsilon_i \sim N(0, \sigma_i^2)$

While we could not replicate the Naïve OLS as presented by the author, we are confident that if we were able to model the original model, we would report similar findings, consistent with our analysis, due to the sheer simplicity of the model.

For all models, we assume that the response is normally distributed, even if this assumption is violated (as was the case in the Naïve OLS). For the Linear Mixed Model, we apply the boxcox transformation to ensure normality, with a lambda of ~ -3.89. We applied the inverse boxcox transformation on our model coefficients and predicted response for interpretation on the natural scale.

We arrived at our best approximation of the 'best' specified model through backwards selection. Primarily, we were interested in the associations between the predictors: distance travelled during ride (km), neighborhood (of pickup), vehicle type (Standard Uber/Uber SUV/etc…); and the response: boxcox adjusted price. Our original naïve OLS model attributed the largest signal to the vehicle type; we were more interested in the effect neighborhood and distance had on price and so we included neighborhood as a random effect, instead.

Prediction

Ultimately, the dynamic pricing algorithm is a predictive model that aims to predict an appropriate price given current rider-internal and external (ex. Weather, crime rate) data, as well as current supply and demand. As such, we wanted to evaluate both the mean predicted price by neighborhood and the average root-mean-square-error (*RMSE*) in prediction by neighborhood, for all models.

We evaluate predictive efficacy through a train-test split. We first randomly shuffled our data and assigned 3/4 of it to a train split, and the other 1/4 to a test split. We trained all of our models on the train split, then predicted on the test split using our fitted models. We calculated the root-squared-error of the test labels and corresponding model predictions, then took the groupwise (neighborhood and vehicle) means of predicted price, 'true' price, and RMSE. As noted prior, vehicle type explains a lot of the variation and only after adding it a grouping variable were we able to determine insightful relationships between predicted and true price and RMSE between neighborhoods.

To reverse engineer the dynamic pricing algorithm, we attempt a clustering analysis. In doing so, we hope to be able to group Uber trips of similar characteristics together, revealing whether like-trips are equally dispersed throughout the city of Boston, or if like-trips are unique to neighborhoods. The characteristics that we consider are the price of the ride, the distance travelled, and the type of vehicle used. We use a k-means classifier in order to group like rides together, where we set the cluster parameter, k, to be equal to the number of neighborhoods present in our dataset. We hypothesized that there would be two extreme scenarios we may encounter: clusters constituting of an even distribution of rides from each neighborhood, and clusters constituting primarily of rides from 1 or 2 neighborhoods. The former would suggest no difference in the types of rides across different neighborhoods, whereas the latter would imply that rides from specific neighborhoods are clearly differentiable from rides from other parts of the city.

3. Results

We compare the model fits of the naïve OLS, and our updated LMM. The effects are summarized as below:

Model	Variable	Coefficient Value	Confidence Interval
OLS	Distance	2.4411	(2.418, 2.464)
LMM	Distance	1.06348	(1.0634, 1.0634)
LMM	Back Bay (contrast)	1	N/A
LMM	Beacon Hill	1.00255	(1.002, 1.004)
LMM	Boston University	0.992197	(0.991, 0.993)
LMM	Fenway	0.998385	(0.987, 0.999)
LMM	Financial District	0.986659	(0.9855, 1.004)
LMM	Haymarket Square	1.00294	(1.002, 1.0145)
LMM	North End	1.01383	(1.001, 1.0124)

LMM	North Station	0.999273	(0.9951, 0.9993)
LMM	Northeastern University	0.993831	(0.9931, 1.003)
LMM	South Station	1.00219	(1.001, 1.0124)
LMM	Theatre District	1.01103	(1.003, 1.0103)
LMM	Base_UberX (contrast)	0.901303	
LMM	Black	1.21168	
LMM	Black SUV	2.5396	
LMM	UberPool	0.875913	
LMM	UberXL	1.0494	
LMM	WAV	0.901302	

Table 1.1 – Summary of the estimated effects for each of the models with 95% confidence intervals for fixed effects.

The units of the categorical variables relate to the natural scale of the response ($USD), whereas distance is in miles. The only estimated effect in the OLS was distance, whereas the fixed effect neighborhood and random effect vehicle type were estimated in the LMM, denoted by nb and veh, respectively.

We only modelled the relationship between price (non-adjusted) and distance to emulate the naïve OLS model. We reported that the effect of distance was significant with a coefficient value of 2.44, suggesting that with one unit increase in the trip distance, the price is expected to increase by 2.44 USD.

Model diagnostics confirmed that the model was far too simplistic and did not sufficiently explain the variation in price, nor were the appropriate adjustments made to correct the model. The figure in the Appendix indicate that the covariate distance alone does not adequately describe the covariance structure. Furthermore, both the response and residuals fail normality assumptions, which undermines the validity of our previous inference of the model coefficients. We address this issue in three steps; first, we survey additional predictors, including time, vehicle type and neighborhood (our variable of interest with respect to identifying potential price discrimination, Secondly, we apply a boxcox transformation to our response, price, in hopes to improve to normality assumptions. Finally, we fit a linear mixed model, setting vehicle type as a random effect. We justify this choice by noting that our dataset did not specify repeated trips for the same Uber; all rides were treated independently.

We assumed, however, that an Uber driver (driving the same vehicle) motivated to maximize their profits given a scarce and set time period would make multiple trips. This assumption necessitated treating rides and vehicles as non-independent, although we cannot necessarily discern repeated trips from our data. While this may apply for all categorical variables, we are still interested in inference of the neighborhood covariate (including confidence intervals), hence it was deliberately left as a fixed effect. Undoubtedly this may affect the validity and performance of the model, but it is an important tradeoff we make to investigate our initial query of price discrimination.

The coefficient summary of the LMM is also summarized in Table 1.1. Firstly, we find that the signal of distance is considerably more diminished than that reported in the OLS model, with an

estimate of ~1.06. Upon considering the other estimates, this however makes sense, as the vehicle type is a much larger signal than the other covariates. We see that, on average, the larger or more luxurious Ubers are much pricier: UberX, Uber Black, and Uber Black SUV are 1.04x, 1.21x, and 2.54x pricier than the base Uber, respectively. In contrast, accessibility vehicles (WAV) and shared rides (UberPool) are approximately equal and lower in price with estimated mean effects of 1 and 0.901 respectively.

Interestingly, we see small but significant differences in the neighborhood fixed effect estimates, relative to the neighborhood Back Bay (base contrast), except for North Station. We note that neighborhoods with large young populations such as Boston University, Northeastern University and Fenway tend to pay less on average for trips compared to other neighborhoods. Wealthier neighborhoods near north-west Boston, such as North End and the Theatre District tend to pay more per trip, on average.

Clustering
We notice that rides from the Fenway area and the Financial District area are differentiable from other rides in Boston. While this does not directly imply that these neighborhoods are priced differently, it informs us that the ride characteristics (price, distance travelled, vehicle used) for rides in these areas are different from that of other rides in the city of Boston. This now allows us to narrow down the focal point of this study to specific, impacted neighborhoods of the Greater Boston Area.
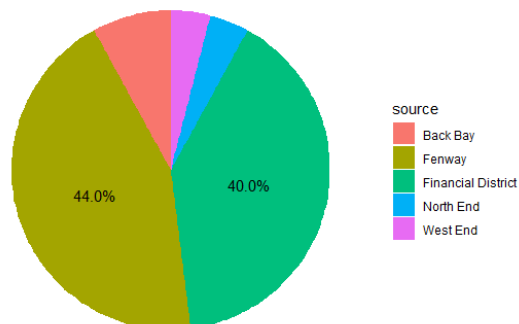

k-means group 8
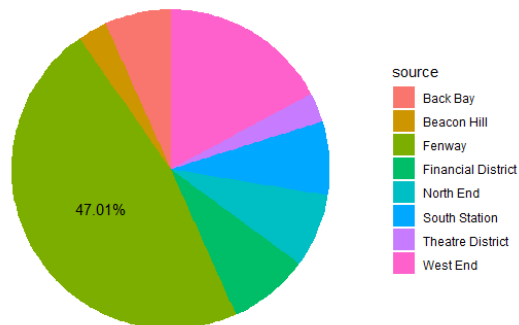
k-means group 2

Fig 2.2 – Clusters 2 and 8 from clustering price, distance, vehicle type. The rides in Fenway are clearly differentiable from other rides in Boston, which suggest clients from different neighbourhoods pay similar prices, despite some neighbourhoods being worse off.

Price Discrimination

The principal aim of this paper was to determine whether price discrimination existed in the Boston Uber rideshare market, specifically between different neighborhoods. We take an algorithm-focused approach, investigating differentials in price prediction between neighborhoods (with blocking of vehicle type) that may suggestion empirical price discrimination. We also evaluate the predictive validity of our candidate model, comparing RMSE between neighborhood-vehicle pairings. We postulate that an algorithm is not biased if there is no significant difference in the RMSE between groupings and biased otherwise.
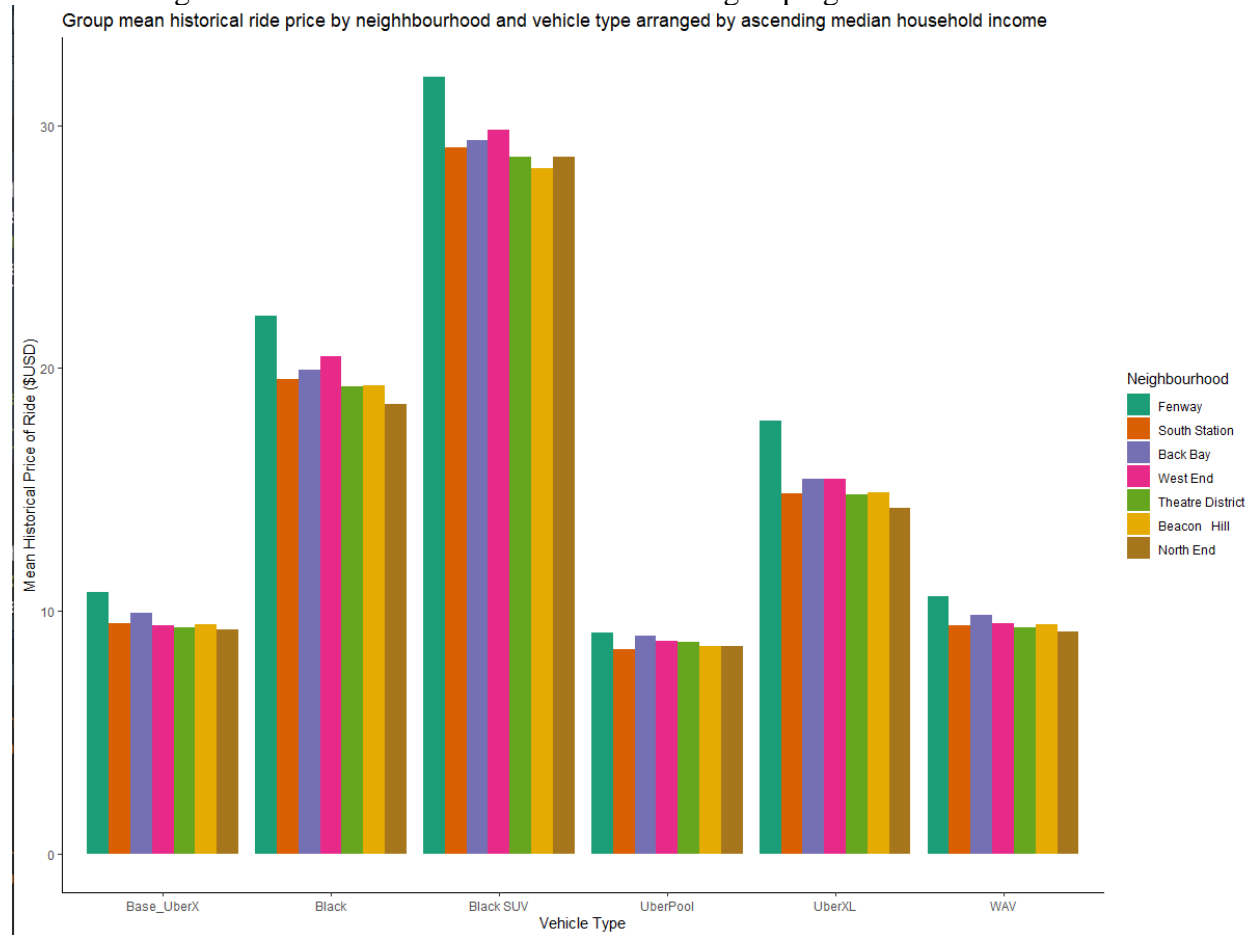

Fig. 3.1 – Comparison of mean historical price (USD$) by neighbourhood of pickup, isolating for vehicle type. Neighbourhoods are arranged by ascending median household income to illustrate that pickups from lower income neighbourhoods tend to cost more on average.
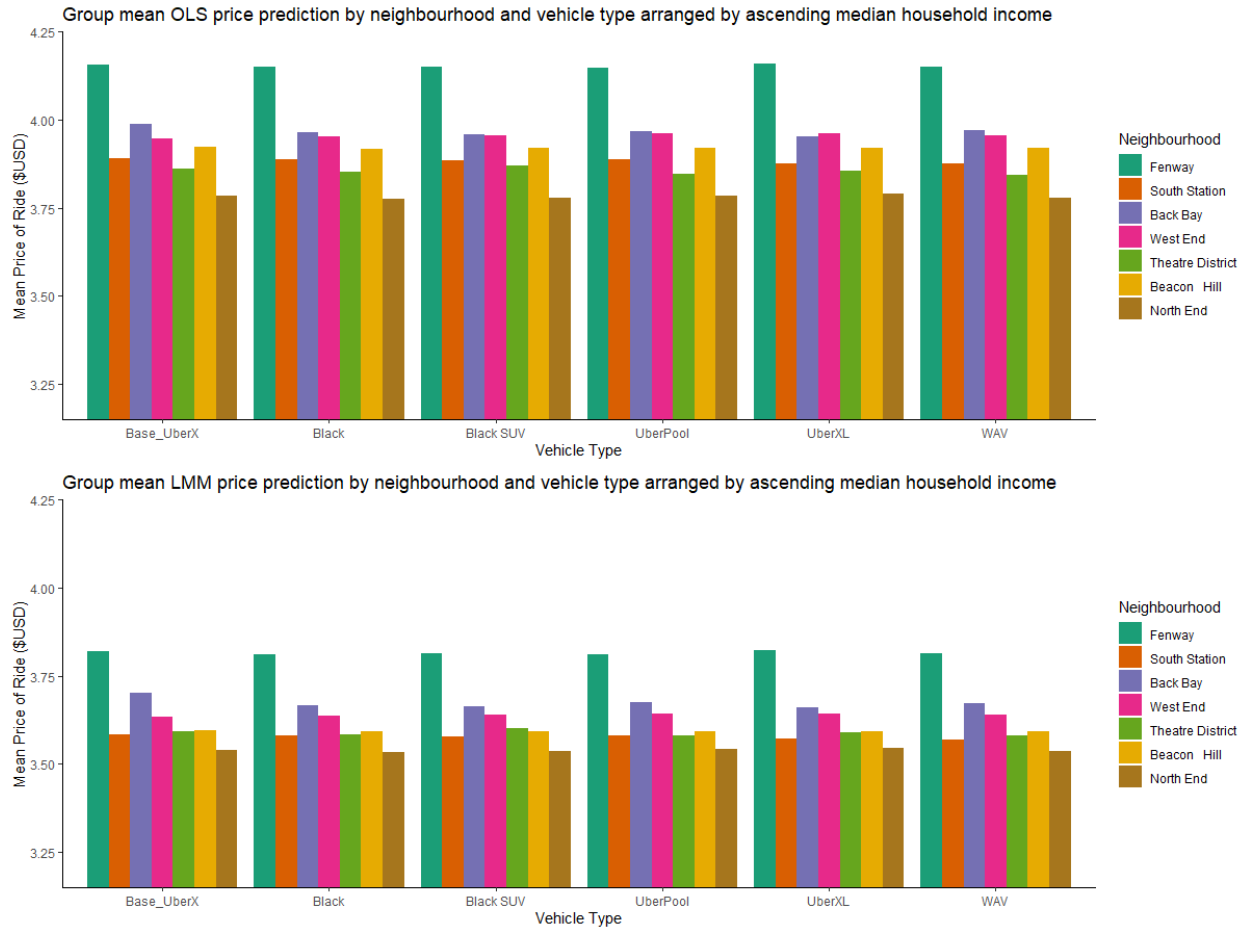
Fig. R.3.2 – Comparison of future predicted price (USD$) by neighbourhood of pickup, isolating for vehicle type. The results for OLS are shown on **top**, whereas results for LMM are shown on **bottom**. Neighbourhoods are arranged by ascending median household income to illustrate that pickups from lower income neighbourhoods tend to cost more on average.
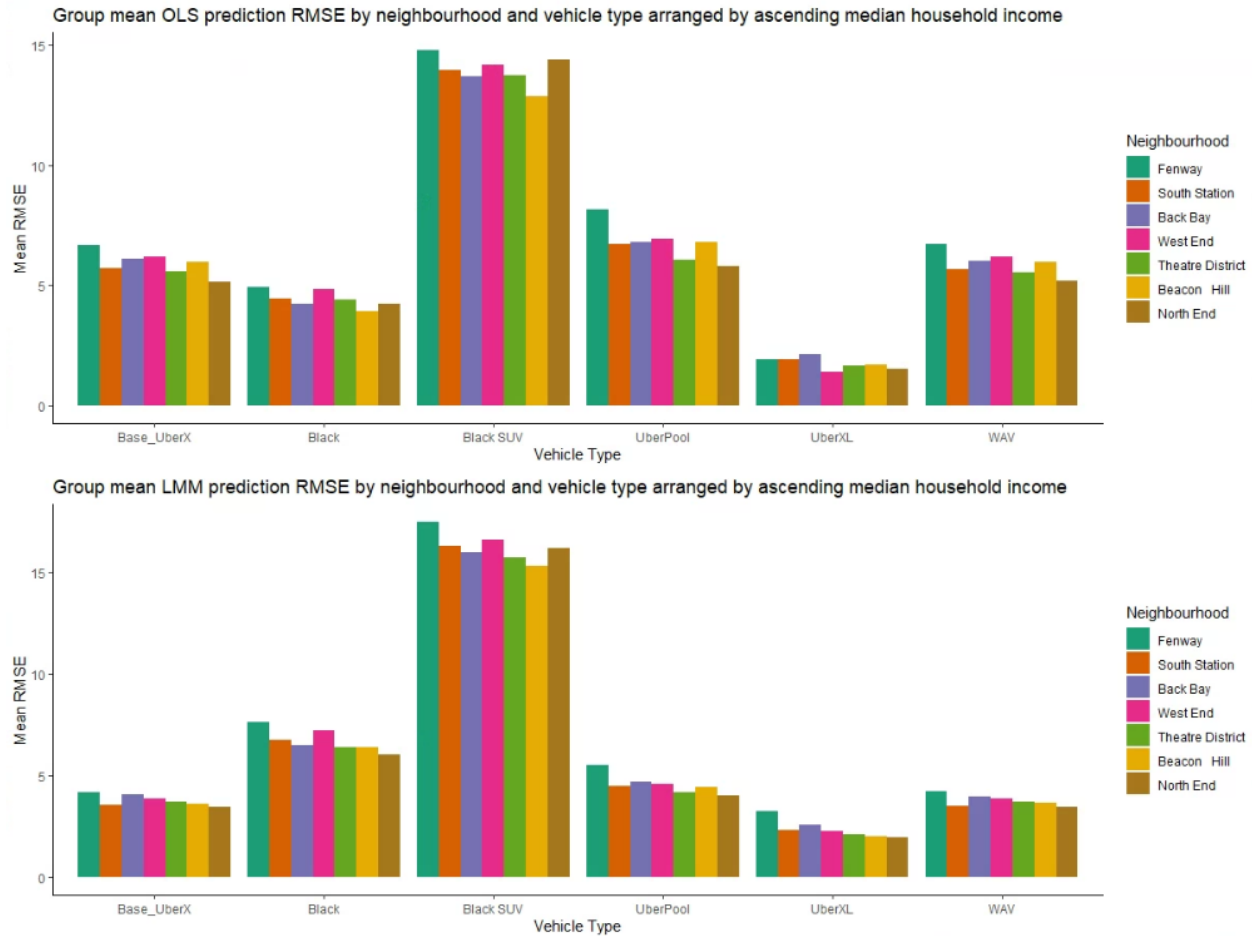
Fig. 3.3 – Comparison of group-wise mean RMSE (USD$) by neighbourhood of pickup, isolating for vehicle type. The results for OLS are shown on **top**, whereas results for LMM are shown on **bottom**. Neighbourhoods are arranged by ascending median household income to illustrate that pickups from lower income neighbourhoods tend to cost more on average.

Fig 3.1 compares the historical prices between neighborhoods after controlling for distance and vehicle type. The neighborhoods are color coded within each vehicle strata and arranged from lowest to highest average median household income. Immediately, we see that the least wealthy neighborhood (Fenway) reports the highest average historical ride price, and the wealthiest neighborhood reports the lowest average historical ride price, consistent for all vehicle types.

Fig 3.2 compares 'future' predicted prices of the naïve OLS and LMM models, between neighborhoods after controlling for distance and vehicle type. Similar to Fig 3.2, neighborhoods are color coded within each vehicle strata, and arranged from lowest to highest average median household income. The results illustrated in Fig 3.3 closely parallel those seen in Fig 3.2 – we see that for the worst-off neighborhoods, the model predicts higher prices for similar rides, compared to better-off neighborhoods. Furthermore, we note that RMSE on average tend to be much higher for less wealthy neighborhoods compared to wealthier neighborhoods, as indicated by the error bars, consistent between vehicle types.

Between models, we observed that the naïve model reported on average higher RMSE and mean ride price for a given group, compared to the better calibrated LMM, suggesting that the OLS model charges on average more for the same ride. When comparing the groupwise RMSEs between models, we demonstrate that while a poorly fit, overly simplistic model is just as prone to error as our model, but would overcharge patrons at all group levels.

4. Discussion

From the above analysis, it is shown that while controlling for the covariates (type of vehicle and distance), riders from neighborhoods with low average income are paying more than riders from neighborhoods with high average income. The discussion then follows to be price discrimination against riders from low average income area. As defined by Patrick Joyce and Thomas E. Merz, 'price discrimination is the practice of charging different prices to different customers for the same goods or services in a competitive market' (Joyce & Merz 1985). The service in our context is ridesharing. It is the act of picking up a passenger and delivering the said passenger to a predetermined location as agreed by both the rider and the driver. There is clear price discrimination in the case of Uber's ridesharing program. Low-income neighborhoods are paying more for the same service than high income neighborhoods.

Let's circle back to the definition of discrimination. Discrimination is the violation of deliberate freedoms to make decisions about how to live insulated from pressures stemming from extraneous traits (Moreau 2010). For the purchase of the same service, which is the physical ride from your source to destination, while controlling for distance and other factors, people are paying different prices depending on their location. No certain group of people are benefiting from this dynamic pricing algorithm practice. Rather, a specific group of people are at a targeted disadvantage based because they requested the service in a low-income neighborhood.

Now naively speaking, let's assume that the pick-up location of the rider is their place of residence. Or rather, the dynamic pricing algorithm assumes this to be true. Note that normatively extraneous features are traits that we believe a person should not have to factor into their deliberations as costs (Moreau 2010). Each of us is entitled to some deliberative freedoms. In the case of Uber's dynamic pricing algorithms, that deliberative freedom is disregarded when riders have to pay more for the same service based on their location or neighborhood. The level of service does not change because you are in a high-income neighborhood. This location feature is a normatively extraneous feature, then there must be a way to guard this against such discrimination.

This form of discrimination is further amplified in a scholarly paper by Caliskan and Pandey, where they analyzed ridesharing data from the city of Chicago. The authors used Iterative Effect Size Bias (IESB) to quantify bias score on price per distance based on ethnicity, education, average house price and citizenship status. As a result, there was a significant difference in ride-hailing prices. When passengers were picked up or dropped off in neighborhoods with low percentage of 1) people with high school education or less, 2) below average house price, 3) high non-white population.

Additionally, there were studies conducted to see immediate discrimination against other normatively extraneous features such as individual's name. Knittel reported that African

American are at 35% increased wait time for ridesharing services when compared to White American. Furthermore, African American sounding names are 3 times more likely of the ride being cancelled than White sounding names (Knittel et, al 2016). Names are considered to be normatively extraneous features and thus should be protected against; however, it is clearly not regarded in Uber's dynamic pricing model.

This form of discrimination against low-income neighborhood will seek to encourage riders to search for alternate form of transportation. Whether that be public transportation or walking. Thereby discouraging riders from ordering the service. This essentially reduces the market size and prices them out of the market. It may be beneficial to the drivers by maximizing their profit, but it is putting another layer of barrier that may restrict service to certain groups of people.

The model that we fit on the data shows bias in dynamic ride-hailing based on the demographic of the neighborhoods. It's crucial that unintended consequences like racial disparities and other normatively extraneous features are identified and accounted for.

One may counter argue that the existence of price discrimination in the Boston rideshare market, is not necessarily bad in this scenario, nor is it necessarily discriminatory. In fact, the data may suggest that both riders and drivers benefit from individualized pricing, in the form of rider demand being met and drivers maximizing expected profit.

With respect to Moreau's definition of discrimination (Moreau, 2010), extraneous features do not explicitly constrain riders from any neighbourhood from hailing rides, nor are drivers constrained from driving in personally preferred neighbourhoods. This is due to Uber's upfront pricing, where parties agree on transparent, fixed fares before the ride, combined with blinding to extraneous features. These work in tandem to eliminate user-level discrimination.

On the demand-side, no consumer is explicitly barred from ordering an Uber. Consumers are not allowed to discriminate against drivers and are simply assigned the closest ride. Riders cannot select on, and thus cannot discriminate against drivers' extraneous features.

On the supply-side, we follow similar reasoning; drivers also cannot choose passengers based on their extraneous features. From the literature, it is safe to assume suppliers are motivated to maximize their long-term expected profits (Asghari, 2016). In fact, assuming time is scarce, and location is relatively constrained, drivers should aim to complete as many (short) rides as possible to maximize profit. Because user-level extraneous factors are otherwise protected and do not play a role in selecting passengers, they cannot weigh user features as costs against profit.

While we suggest that discrimination on user-level covariates is not possible, we have yet to discuss discrimination on group parameters, such as neighbourhood demographics. This is entirely viable if group-level parameters are generally well-known amongst drivers and aggregately extraneous.

Particularly, on the supply-side, we should account for potential driver self-selection causing neighbourhood-level shortage supplies. This self-selection may be motivated by negative prejudices and perceptions towards certain neighbourhoods, resulting in selective withholding of service, and thus neighbourhood-level discrimination exclusive of user-level discrimination.
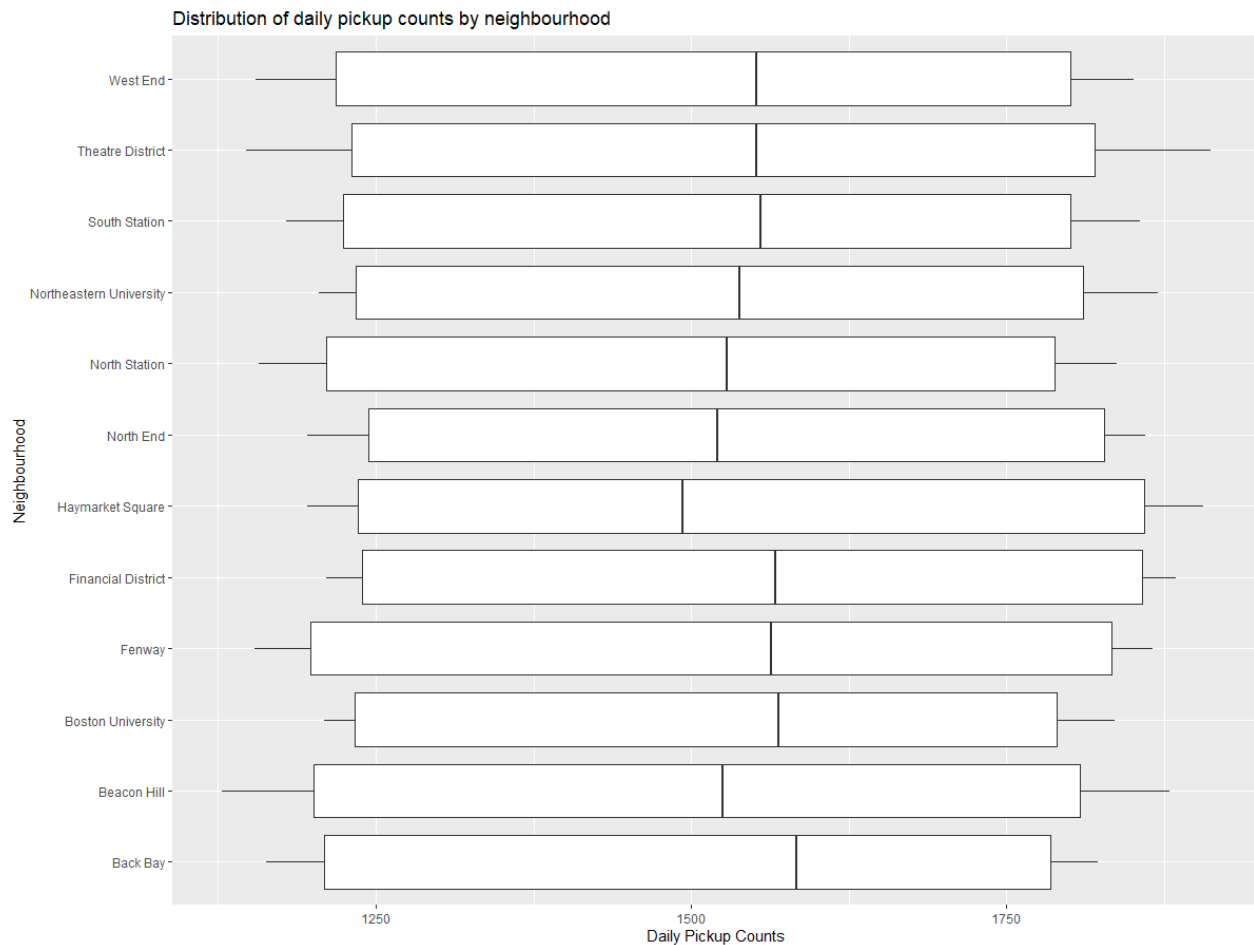
Fig. 4 – Distribution of daily Uber pickup counts by neighbourhood.

Interestingly, our data counters this aforementioned 'intuitive' argument. When we visualize the daily pickup counts by neighbourhood, as in Fig. 4., those that are worse off socioeconomically, such as Fenway and Boston University, feature some of the highest ride counts per capita, compared to wealthier ones such as North End. This suggests that despite higher mean prices, users from adversely affected neighbourhoods are not particularly dissuaded from seeking out Ubers, and perhaps against expectation, seek out more rides per capita.

The willingness for patrons from low-income communities to pay and ride more may be explained by the benefits incurred by these groups. In her 2018 dissertation, Brown highlights how in Los Angeles, Ubers are used in low-income communities as a safe, affordable, and accessible alternative to owning a car. Recalling that neighbourhood is often a proxy for race, we cite another paper by Brown (2019), who reports that the discrepancy of waiting times and rates of trip cancellation for taxicabs between white and black riders is much greater than the same discrepancy seen in Uber ridesharing. Brown attributes this form of discrimination to taxi drivers knowing and acting on riders' extraneous features, compared to those same features being better protected in ridesharing programs, in agreement with our arguments.

Again, price-based discrimination clearly exists in this market, but I assert that it benefits low-income neighbourhoods, as opposed to discriminates against them. User-level discrimination

does not exist if users' extraneous features are not visible. Neighbourhood-level discrimination still may exist; however, recent literature suggests that higher prices in lower-income areas are costs passengers are willing to incur to improve their personal welfare.

Furthermore, our pickup instances suggest that demand is still being met across neighbourhoods and riders in adverse neighbourhoods are not necessarily being constrained due to supply shortages. In return, drivers who are motivated to maximize profits benefit from marked fare differentials. Thus, price discrimination generated by dynamic pricing is in fact *good* and produces a more efficient market that mutually benefits both riders and drivers.

Despite the arguments, dynamic pricing algorithms still perpetuate inequity. At the surface, dynamic pricing is all about optimizing supply and demand – but because of the way that cities are often segregated by age, race, or income, we tend to notice a bias that is unintentionally split by neighborhood demographics. One of the biggest caveats of AI and machine learning techniques is that the products built from these techniques are only as good as the data that they were trained on. When machine learning is applied to social data, the algorithms learn the statistical regularities of the historical injustices and social biases embedded in these data sets. A machine trained on biased data will inevitably result in a biased machine, and this is exactly what we're seeing with dynamic pricing models in the ridesharing industry. Because societal biases and price discrimination already existed in our data, bias was introduced in our approximation of Uber's dynamic pricing model when we included neighborhoods, assumed to be the location of residence, as a covariate in the model. Recall that neighborhoods of residence are protected attributes (or proxies of), and therefore any blatant pricing differences based on neighborhoods qualifies as unfair price discrimination.

Besides AI/machine learning bias, we also encountered label bias in our data. Ridesharing is a relatively new industry, and in its early days only a small demographic of the general population interacted with this type of service. Overtime, the dynamic pricing algorithms were trained using this demographic of individuals, but as ridesharing has now become more of a social norm, more of the general population is looking to get involved with ridesharing services. Unfortunately, the data that was used to train the product (the dynamic pricing algorithm) is not representative of the entire population to which the product is being marketed. As a result, we see a product that is being tailored more to the accessibility and the needs of a subset of the population using the product. This label bias is a bias that the dynamic pricing algorithm will gradually unlearn, as it continues to learn from a more diverse set of data; but the fact of the matter is that its foundation was built on bias, and it will never be able to completely overturn that bias. Some examples of demographics being affected by label bias are individuals who were initially priced out of the market, and were unable to afford the services until recently, as well as older generations who were not keen on the idea of ridesharing but eventually warmed up to it as the ridesharing industry engulfed the taxi industry which had been dominating the market for years.

Earlier we saw that low-income areas were being charged the highest prices for rides, and this was justified by the amount of demand these areas generated; low-income areas, such as Fenway (includes Boston University and Northeastern University), had some of the highest demand per capita in Boston. The problem with this ideology is that when an algorithm is biased the bias gets amplified overtime since the algorithm is continuously learning from biased decisions. When the

dynamic pricing algorithm's bias continues to propagate away from fair pricing, the supply of drivers is continuously shifted to higher demand areas, and understandably so. As a driver, given 2 opportunities that require the same amount of time and effort, it is almost certain that the driver will choose the opportunity that will generate the higher return. This cycle is what continues to amplify bias over time, as drivers prefer to serve high-demand areas because it maximizes their profits.

We've discussed about a few of the inequities identified in the dataset, so we will now explore some possible remedies for these. First, we revisit Knittel's paper which addresses race and gender inequalities in Uber wait times and cancellation rates. Here, the author suggests that discrimination can be avoided by removing the driver's access to view any of the rider's protected attributes. Knittel states that "removing names from trip bookings may alleviate the immediate problem" and suggests a 4-digit pin verification system rather than a name verification for Uber pick-ups. However, the author follows by warning readers that "[this solution] could introduce other pathways of discrimination for unequal treatment of passengers" (Knittel, 2016). Where this proposed solution fails is that the response of a driver who is now being "forced" to drive a rider he would have otherwise cancelled on is somewhat unpredictable. A likely scenario may be that the driver would still cancel on the rider, but later than he normally would have. An even more likely scenario would be that the driver would still driver the rider, but give them a bad star rating, to no fault of the rider's. This would then make it more difficult for the rider the secure Uber rides in the future. Bimpikis' paper addresses price discrimination based on race, which as previously discussed, is strongly correlated to price discrimination based on neighborhood demographic. The author suggests that we can correct the price inequity by redistributing the supply of drivers to lower demand areas. If the drivers can be incentivized to work in lower demand areas, for example by having the rideshare company pay out the difference in revenue that the drivers would have made if working in a high demand area instead, price inequity would be removed. The author also notes that "if prices are uniform across the different neighborhoods of a city, profits are maximized", which seems like an advantageous solution for all involved (Bimpikis, 2019). Ultimately, we will not be able to mitigate AI bias without algorithmic transparency, so black box algorithms will need to be made transparent in order to ensure a minimal amount of bias.

5. Conclusions and Limitations

Likely the main limitation of our work was our final candidate model. While we certainly improved upon the statistical validity of publicly available models, we fall short in predictive validity, as evidenced by Fig.3.3. It was kindly suggested to us to include other neighbourhood demographic data, such as walk scores, and public transport metrics that might influence a rider's decision to request an Uber. Undoubtedly, future work would address these issues, namely by fitting the covariates and a more complex covariance structure. We may be interested in more or an ensemble of modern predictive techniques, such as gradient boosted regression, or neural nets.

Despite a poor (but improved) model, we still feel our results are still demonstrative of price discrimination and historical label bias inducing algorithmic bias. Furthermore, even a simple investigation on empirical prices by neighbourhood would have yielded extremely similar

discussion, where we consider whether price discrimination is beneficial or detrimental to adversely affected groups.

Despite fitting an 'incomplete' model, we are satisfied with our discussion and would be surprised to find Uber's black box algorithm does similarly reflect bias, as supported by the literature, unless specifically controlled against. Ideally, we would like to test this hypothesis using Uber's actual predictive algorithms, however, this would not be possible unless they are publicly released (which is unlikely). This again reinforces the need for transparency and explainability for these 'black box' algorithms to effectively combat bias in AI.

6. References

[1] Asghari, M., Deng, D., Shahabi, C., Demiryurek, U., & Li, Y. (2016, October). Price-aware real-time ride-sharing at scale: an auction-based approach. In *Proceedings of the 24th ACM SIGSPATIAL international conference on advances in geographic information systems* (pp. 1-10).

[2] Bimpikis, K. et al. (2019). Spatial Pricing In Ride Sharing Networks. University of Toronto: Informs Pubs Online. https://pubsonline.informs.org/doi/10.1287/opre.2018.1800

[3] Brown, A. E. (2019). Prevalence and mechanisms of discrimination: Evidence from the ride-hail and taxi industries. *Journal of Planning Education and Research*, 0739456X19871687.

[4] Brown, A. E. (2018). *Ridehail revolution: Ridehail travel and equity in Los Angeles.* University of California, Los Angeles.

[5] Fillmore, I. (2016). Price discrimination and public policy in the us college market. *Employment Research Newsletter*, *23*(2), 2.

[6] Joyce, P., & Merz, T. E. (1985). Price Discrimination in Academic Journals. The Library Quarterly: Information, Community, Policy, 55(3), 273–283. http://www.jstor.org/stable/4307858

[7] Knittel, C R. et al. (2016). Racial and Gender Discrimination in Transportation Network Companies. National Bureau of Economic Research. https://www.nber.org/papers/w22776

[8] Lötters, F., Meerding, W. J., & Burdorf, A. (2005). Reduced productivity after sickness absence due to musculoskeletal disorders and its relation to health outcomes. *Scandinavian journal of work, environment & health*, 367-374.

[9] Pandey, A. & Caliskan, A. (2020). Disparate Impact of Artificial Intelligence Bias in Ridehailing Economy's Price Discrimination Algorithms. Cornell University. https://arxiv.org/pdf/2006.04599.pdf

[10] Schneider, T. (2016, April 5). Taxi, uber, and lyft usage in New York City. toddwschneider.com. Retrieved December 14, 2021, from https://toddwschneider.com/posts/taxi-uber-lyft-usage-new-york-city/.

[11] Tompa, E. (2002). The Impact of Health on Productivity: Empirical. *The review of economic performance and social progress 2002: towards a social understanding of productivity*, *45*.

[12] How uber's dynamic pricing model works. Uber Blog. (2021, December 13). Retrieved December 14, 2021, from https://www.uber.com/en-GB/blog/uber-dynamic-pricing/.

[13] Law document english view. Ontario.ca. (2018, November 19). Retrieved December 14, 2021, from https://www.ontario.ca/laws/statute/90h19.
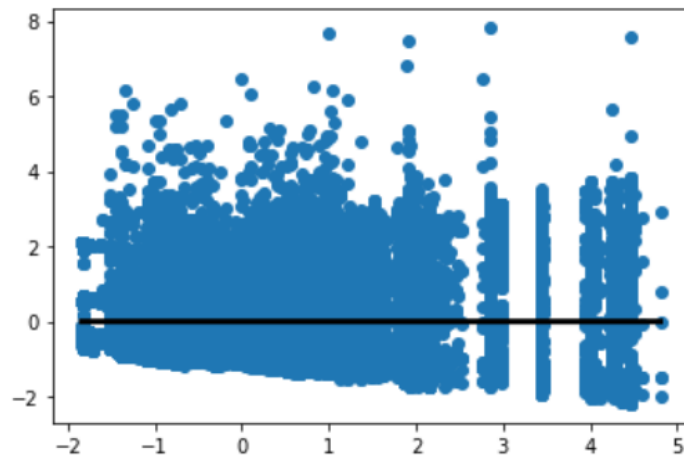
Appendices
Model Diagnostics: OLS

Fig A.1.1 – Plotted residuals over fitted values for OLS. Clearly, residuals are top-heavy and variance may be non-constant
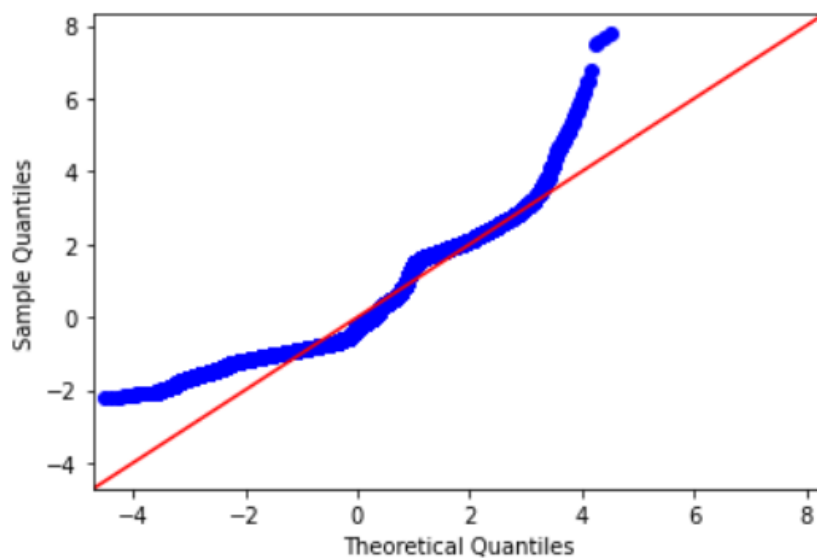

Fig A.1.2 – QQ-plot of residuals for OLS. Note that there is a strong right-skew, and normality is clearly violated.
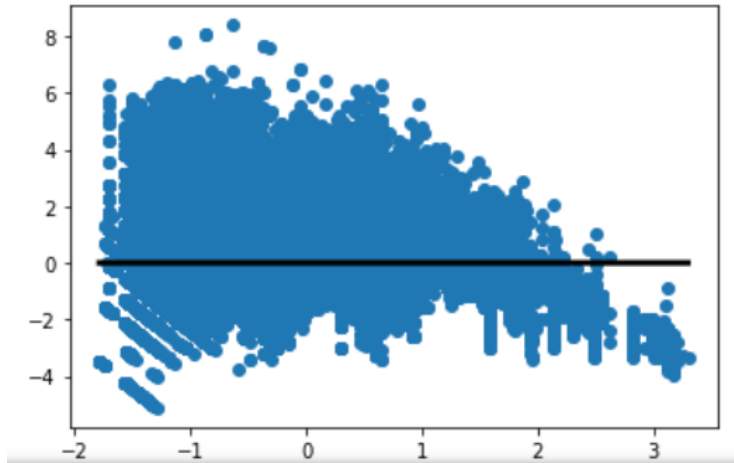
Fig A.1.3 - Plotted residuals over fitted values for LMM. Adding more information did not improve heteroscedasticity, however, top-heaviness is improved, especially for costlier rides. Scale is on boxcox intervals.
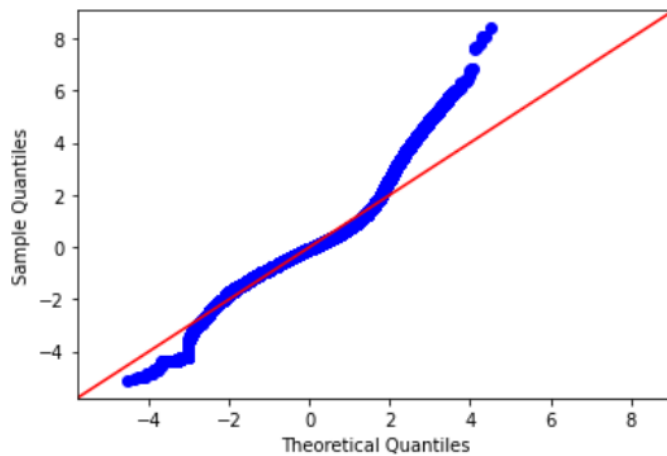

Fig A.1.4 - QQ-plot of residuals for LMM. Note that there is a strong right-skew, and normality is clearly violated, however, after applying the box-cox transformation, skew seems to be diminished on both ends.