BEA Working Paper Series, WP2024-7

Consumer Learning and Price Index Bias: How Diffusion of Product Quality Knowledge Impacts Measures of Price Change

Author Daniel Ripperger-Suhler, U.S. Bureau of Economic Analysis[1]

Contact daniel.ripperger-suhler@bea.gov

Date November 2024

Abstract There is a general consensus that the bias associated with the entry of new merchants has nontrivial implications for measuring inflation. However, quantifying the bias empirically has proven difficult in part because little is known about how much of the price differences in goods sold by new versus old merchants represents a pure price difference (inflation) or differences in the quality of the attendant services (quality differences). In the public transportation industry, measurement of quality is complicated by the accompanying technological change rideshare services represented. As with any completely new technology, consumers faced considerable uncertainty around the quality of rideshare services. Consequently, consumers' perceived quality of rideshare services changes over time, which makes the calculation of constant-quality price indexes even more challenging. This paper explores a new method for accounting for this bias by separately identifying changes in product price and quality over time. I estimate multiple hedonic models to recover quality adjustment factors for quality-adjusted unit value price indexes. One of these models utilizes measures of time-varying product quality that are derived from a structural demand model of endogenous consumer learning that explicitly models the diffusion of knowledge about the quality of rideshare services. I compare the measurement of quality and pure price differences across modes of transportation, and the implications they have for constant-quality price indexes and, consequently, the measurement of inflation.

Keywords Price indexes, inflation

JEL Code C43, D12, D83, E31, L15, L91, O33

1. I would like to thank Ana Aizcorbe for many helpful discussions. I am also grateful to Kevin Fox (discussant at the International Association for Research in Income and Wealth (IARIW) 2024 conference) for his insightful comments, and other conference participants at IARIW 2024.

1. Introduction

There are many sectors for which the measurement of price change is particularly difficult due to both quality differences across goods and pricing differences across merchants. Since the goods in these industries are not identical, a unit value index is not appropriate. However, treating the goods as strictly distinct results in pricing differences across merchants being incorrectly subsumed into quality measures, and subsequently removed from the price index. This problem is complicated further when there are unobservable quality characteristics that impact price, because any measure of price change will be biased; even if price differences across merchants are nonexistent.

How do these issues affect price measurement in the public transportation industry, especially in the presence of rideshare services? First and foremost, unobserved characteristics are particularly problematic for price measurement of any service, because their quality can be difficult to quantify with the available variables. Consequently, there is likely a larger bias than is the case for goods with a clearly defined set of quality characteristics. Second, pricing strategies across merchants are not only present, but are well-known to be time-varying, because of the practice of surge pricing by rideshare companies. In addition to surge pricing, both taxis and rideshare services have changed their underlying pricing schedules on a number of occasions. Thus, even if surge pricing weren't employed, these merchants would still exhibit time-varying price differences. Finally, the creation of rideshare services was such a revolutionary event that consumers and the companies alike were faced with the task of learning about the quality of the service. This learning process has led to changes in the quality of rideshare services that impact the measurement of trip prices. Thus, in order to appropriately measure price change in this industry, the econometrician must account for observable and unobservable characteristics that change over time.

Several papers have addressed these difficulties in price measurement in some capacity. However, none are well suited to deal with the problems in the transportation industry mentioned in the previous paragraph. For example, Erickson and Pakes (2011), develops a method for addressing bias due to omitted variable bias. Their approach focuses on accounting for unobserved differences in goods that lead them to enter or exit the market. Unfortunately, the transportation industry does not exhibit a high degree of turnover, yet it experiences perpetual changes in quality of the extant services throughout their lifecycle. Thus, the method in Erickson and Pakes (2011) is not applicable. Another paper that tackles these problems, which focuses on identifying pure price differences across outlets, is Greenlees and McClelland (2011). They estimate a hedonic model with outlet fixed effects, which makes it possible to recover average price differences across outlets over the sample period. This method is particularly useful in quantifying price differences across outlets for identical goods, which is their focus. However, it is not ideal when there are unobserved characteristics that differ across outlets, because it is not clear if the fixed effects are picking up mode-level differences in quality or pricing strategies. As mentioned before, the transportation industry is plagued by unobservable quality due to a lack of relevant characteristics. Additionally, time-varying differences in pricing strategies between taxis and rideshare companies imply that the time invariance of transportation mode fixed effects will introduce a bias by averaging these differences over the entire sample period. Consequently, the fixed effects approach is not appropriate for price measurement in the current setting.

Therefore, I develop an approach using a structural demand model with consumer learning in order to estimate quality adjustment factors that identify unobserved quality as it evolves over time. Without the structural model, it is not possible to separate the impact of unobservable quality on price from changes in pricing strategies in the hedonic price equation. The quality adjustment factors are used to compute a quality-adjusted unit value (QAUV) price index that appropriately measures price change in the subsector of the transportation industry comprised of taxis and rideshare companies. I evaluate this new method by examining the New York City (NYC) market for taxis and rideshare services. I find that the unobserved quality terms estimated in the consumer learning model are important predictors of price, which translates into quality adjustment factors that have a meaningful impact on the QAUV price indexes. In some cases, the compound annual growth rates (CAGRs) computed from the consumer learning model were as much as 6.5% different from quality adjustments made through standard hedonic techniques.

The paper proceeds as follows: Section 2 discusses the taxi industry, the construction of the data, and some reduced-form evidence. Then, I introduce the structural model in Section 3. The estimation procedure and identification are covered in Section 4, results are presented in Section 5, and the paper concludes in Section 6.

2. Industry Background and Data Description

2.1. The Taxi and Rideshare Market in NYC

New York City has the highest population density of any city in the United States at nearly 30,000 per square mile, with Manhattan's population density at just over 2.5 times that of the city as a whole.[2][3] Aside from the monetary implications of car ownership in a densely populated urban area, there is simply

[2] https://www.census.gov/quickfacts/fact/table/newyorkcitynewyork/PST045223

[3] https://data.census.gov/profile/Manhattan_borough,_New_York_County,_New_York?g=060XX00US3606144919

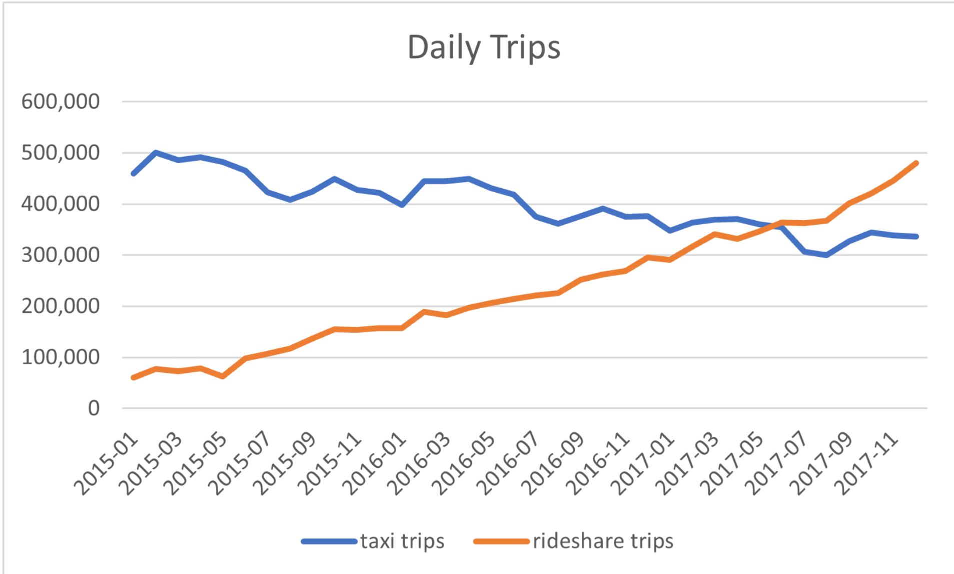

Figure 1. Daily Trips

not enough space for all residents to own a car. This fact is reflected in New York City's statistics, with only 36% of surveyed residents claiming to own a car.[4] As a result, the public transportation system in NYC is one of the most comprehensive in the world out of necessity. There are four main public transportation options in the city: the subway system, the city bus network, taxi cabs, and for-hire vehicles (FHVs). The last category is comprised of rideshare companies and other similar services. The focus of the current paper is on the competition between taxi cab and rideshare companies.

2.1.1. Rise of the Rideshare Company

Uber entered the NYC market in May 2011, and Lyft followed three years later in July 2014. The impact of entry by rideshare companies was felt by all incumbent public transportation options, but the effect on their closest substitute, taxis, was the most drastic. The average number of daily trips for yellow cab taxis fell from 463,701 in November 2010 to 336,737 in November 2016, while the average number of daily trips for Lyft and Uber in November 2016 had reached 269,536. The implication here is that Lyft and Uber did not just reach a new group of consumers, but also attracted riders that would have otherwise chosen to take a taxi. Figure 1 shows the average daily trips by month for taxis and rideshare services from 2015 to 2017 for the entire NYC market.[5] The taxi counts include trips for both yellow and green taxis, while the rideshare counts are comprised of Lyft and Uber trips. Although both Lyft and Uber entered the NYC market prior to 2015, the NYC TLC does not have publicly available data

[4] https://www.nyc.gov/html/dot/downloads/pdf/nycdot-citywide-mobility-survey-report-2018.pdf

[5] The data are compiled by the author from NYC OpenData and TLC monthly trip records.

for rideshare services prior to 2015.

2.2. Data and Descriptive Statistics

This section describes the data used in estimation and provides descriptive and reduced-form evidence.

2.2.1. Data

The raw data are compiled from two sources by Aizcorbe and Chen (2022). One is the NYC Taxi and Limousine Commission (TLC), which has publicly available, trip-level data for yellow and green taxis. The data include pickup and dropoff times and locations as well as detailed fare information including base fares, taxes, tolls, tips, and more. While the TLC also publishes trip-level data for all FHVs, price information is not included. The TLC data are supplemented by a sample of rideshare consumer email receipts obtained from Rakuten Intelligence.[6] The TLC data categorize pickup and dropoff locations by NYC taxi zones, while the Rakuten data use nine-digit ZIP Codes. In order to match the data, geo-processing techniques are used to map ZIP Codes to taxi zones.[7] After matching the data, the observational unit is a trip within a 15-minute period for each mode at the route level, where a route is defined as a directional NYC taxi zone pair. This is the dataset used in Aizcorbe and Chen (2022), which I further augment by restricting the service types to taxis, yellow and green cabs combined, basic Lyft rides, and UberX rides. I ignore specialty rideshare options, such as, Lyft Line, Uber Pool, and Uber Black. Aside from these services being clearly different from standard taxi cabs, they have different pricing schedules and much fewer observations than the basic Lyft and Uber services. I also exclude specialty FHV services like central dispatch facilities, livery, and limousine companies.[8] The data are grouped into seven time blocks that are a combination of those defined in Cohen et al. (2016) and Lam and Liu (2017). Trips are separated due to differences, such as, surge frequency, distance, and duration, which are indicative of these being fundamentally distinct markets with different consumer behavior.[9] These time blocks are listed in Table 1. Within each time block, I aggregate up to the mode-week level across all routes. Finally, I focus on the time period from the beginning of 2015 to September 2016.[10] The final sample contains 1,785 mode-week-level observations.

[6]Rakuten Intelligence was acquired by NielsenIQ in September 2021.

[7]See Aizcorbe and Chen (2022) for more information about the processing of the data.

[8]See `https://www.nyc.gov/site/tlc/businesses/for-hire-vehicles.page` for a brief description of the different FHV categories.

[9]Trip characteristics by time block can be found in Table 7 in Section B of the appendix.

[10]In September 2016, Uber faced price gouging accusations after surge pricing went into effect in the aftermath of the 2016 New York and New Jersey bombings. In response to these accusations, Uber changed their surge pricing algorithm.

Time Block	Day of Week	Time of Day
Morning Rush	Monday-Friday	5:00am to 9:00am
Weekday Day	Monday-Friday	9:00am to 5:00pm
Evening Rush	Monday-Friday	5:00pm to 7:00pm
Weekday Evening	Monday-Friday	7:00pm to 11:00pm
Weekend Day	Saturday-Sunday	6:00am to 5:00pm
Weekend Evening	Saturday-Sunday	5:00pm to 11:00pm
Bar Hours	Thursday-Saturday	11:00pm to 11:59pm
	Friday-Sunday	12:00am to 3:00am

Table 1. Time Blocks

As mentioned earlier, the daily average trips and market share for Lyft and Uber are increasing from 2015 to 2017, the implication being that substantial learning about rideshare services is still taking place, and the quality of these modes of transportation is still evolving. However, it is not guaranteed that the pattern in the aggregate data across all NYC boroughs, shown in Figure 1, will also prevail within each time block for trips within Manhattan. For example, the total number of daily trips for rideshare services in Manhattan does not overtake that of taxis like it does for the entire NYC market, as shown in Figure 1.[11] The pattern of increasing market shares over time within each time block is illustrated in Figure 2, which shows the share of weekly trips for Lyft and Uber relative to the share of taxi trips for each time block. In each individual graph, the relative shares for both rideshare services are increasing over time on average. Thus, the aggregate trend seen in Figure 1 is also present within each time block for each rideshare company. Of course, the growth of market shares on its own does not guarantee that a price index will be biased. The larger implication of increasing shares is that either consumers are still learning about rideshare companies, rideshare companies are reducing prices to attract consumers, or both. Each case presents obstacles to price measurement that were mentioned in Section 1. Thus, increasing shares within time blocks highlights the importance of obtaining time-varying mode quality measures that are separate from mode-level pricing strategy differences.

[11] See Figure 7 in appendix C for the Manhattan analog to Figure 1 aggregated up to the week level.

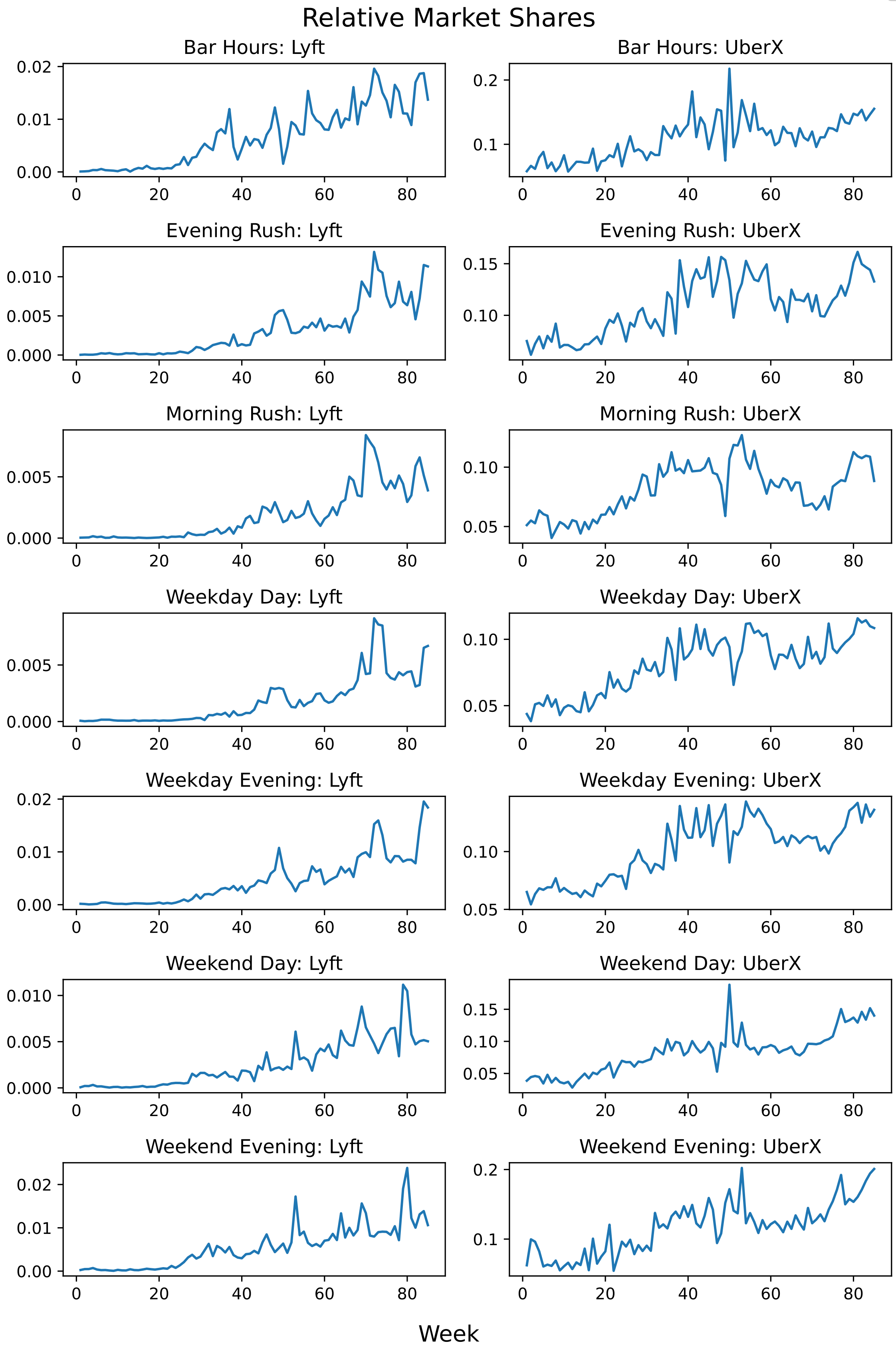

Figure 2. Relative Market Shares

	Mode	Mean	Std. Error	Min	Max	Median
Adjusted Multiplier	Lyft	1.232	0.245	1.000	2.513	1.152
	Taxi	1.000	0.000	1.000	1.000	1.000
	UberX	1.197	0.180	1.000	2.216	1.151
Distance	Lyft	2.300	0.655	0.810	7.959	2.241
	Taxi	1.408	0.084	1.254	1.578	1.399
	UberX	1.896	0.164	1.520	2.985	1.874
Duration (Minutes)	Lyft	16.153	4.264	6.686	49.885	15.634
	Taxi	9.644	1.454	7.310	14.388	9.218
	UberX	13.910	2.250	9.768	21.877	13.456
Inside Share	Lyft	0.003	0.004	6.88e-6	0.020	0.002
	Taxi	0.911	0.028	0.820	0.973	0.913
	UberX	0.086	0.025	0.027	0.166	0.084
Price	Lyft	14.306	3.776	6.063	40.220	13.696
	Taxi	10.148	0.663	8.752	12.029	10.008
	UberX	13.335	2.572	8.768	27.292	12.825
Surge Share	Lyft	0.031	0.079	0.000	1.000	0.010
	Taxi	0.001	0.001	1.99e-6	0.004	0.001
	UberX	0.002	0.002	0.000	0.007	0.002

Table 2. Summary Statistics by Mode

2.2.2. Descriptive Statistics

In Table 2, I show summary statistics for trip characteristics by mode. The distance and duration of a trip are the two most important price characteristics, as trip prices are largely based on nonlinear pricing schedules consisting of a base fee plus the distance and duration. Table 2 shows that trips using rideshare companies are typically longer than taxi trips, in both distance and duration. However, Lyft is the clear outlier in terms of distance and duration. These differences are not inherently quality related, and as such, must be accounted for when computing price indexes. The removal of these differences alone produces heterogeneity-adjusted price estimates (see Silver, 2009). The adjusted multiplier for each rideshare option is the average multiplier applied to a trip during the time block in a given week. Of course, this variable is always one for taxis, as they do not engage in surge pricing. The same is not true for the surge share variable, because it is the share of trips on a particular mode that occurred during a 30-minute window that was flagged as having surge pricing.[12] In other words, it captures how often a mode operates within a market that is experiencing high demand. Finally, the inside share variable is the quantity share of each mode within the market for rides. In other words, it is the inside share used in the discrete choice demand model that is discussed in Section 3.

[12]The number of trips is used in estimation, but the share is presented to give a clearer picture of the number of trips in relation to total trips on each mode.

2.3. Reduced-Form Evidence

In this section, I estimate the baseline hedonic model, which contains the distance, duration, number of surge trips, and the adjusted multiplier in addition to the time dummies. The trip characteristics are the first three variables mentioned, while the adjusted multiplier is used as a proxy for a function of observable differences in pricing strategies across modes. I use this specification to examine the residuals to determine if there is any pattern that could cause the price index to be biased. In particular, if they are correlated with time, it would indicate that there are unobserved characteristics that could influence any price index computed using the regression estimates. Since I am focusing on the residuals, the parameter estimates from the hedonic model are presented in Section 5 alongside the results from the structural model to facilitate a comparison.

In Figure 3, the residuals from the seven separate time block regressions are plotted against time. There are two main points of interest. First, the residuals appear to be correlated with time for every mode across all time blocks. Thus, there are likely unobserved characteristics that impact the prices of each mode and, consequently, the associated price index. Second, within each time block, the relationship between the residuals and time for each mode appears to be different. This feature is particularly important, because it points to a differential impact of the unobserved characteristics on the prices of each mode over time. Thus, for any approach to be successful in producing a price index free of bias, it must separately identify changes in quality from differences in pricing over time and across modes. As previously stated, the method proposed in this paper is designed to address the issue of unobserved characteristics and is discussed in Sections 3 and 4.

3. Consumer Learning Discrete Choice Model

In this section, I present the structural demand model with consumer learning that is estimated in order to compute the time-varying quality adjustment measures used to separate unobserved quality changes from pure price differences. First, I introduce the Bayesian updating process that models how consumers learn about the quality of each transportation mode. Then, I discuss the interpretation of perceived quality and Bayesian updating in the context of price index quality adjustments. Finally, I turn to the expected utility of consumers, the discrete choice demand model, and the associated choice probabilities.

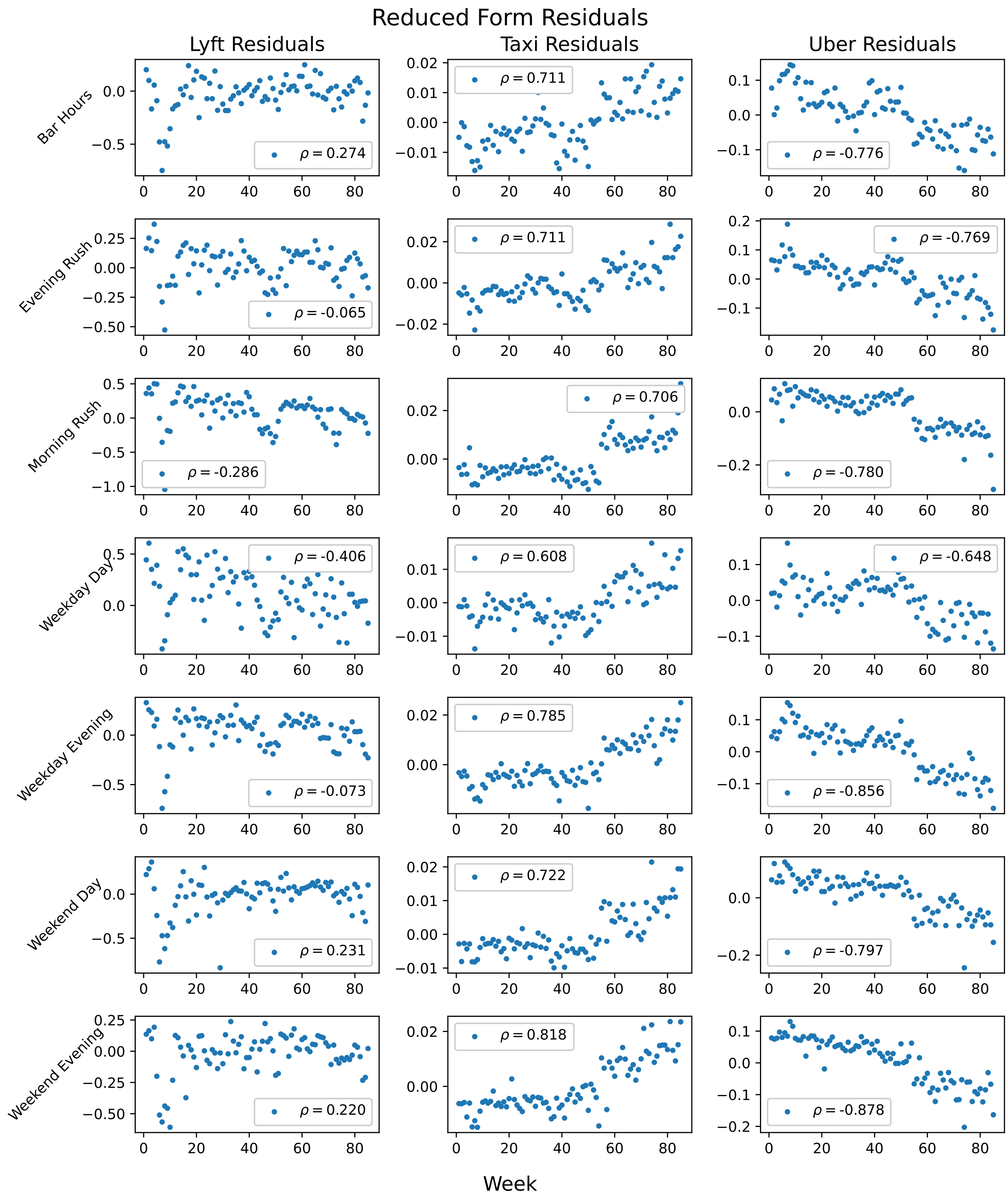

Figure 3. Reduced Form Residuals

Consumers are myopic in that each period, they maximize their current utility based upon their information set, but they are not forward looking. This implies a static discrete choice demand framework that is augmented by consumers' ability to make inferences based on past information. I opt for this approach rather than allowing consumers to be forward-looking for two reasons. First, forward-looking utility maximization is not likely to be common in this subsector of the transportation industry. Second, the dynamic optimization problem with aggregated consumer learning is computationally intensive to the point of near infeasibility.[13] Furthermore, since I do not have the individual-level purchasing histories used in many studies (e.g. Erdem and Keane, 1996; Roberts and Urban, 1988), I use the modeling approach in Ching (2010b). Rather than expected utility being determined by consumer-specific perception of quality, individual choices are made based on the public perception of a given mode of transportation. However, the mode-level perceived quality is computed by aggregating the signals each consumer receives from their own experience each time they take a trip. This approach implies that an information aggregator takes the consumer-specific experience signals each period and updates the public perception of quality using a Bayesian mechanism. In Ching (2010b), patients are the consumers, and physicians assume the role of information aggregator. In each period some fraction of patients report their experiences to their physician who, in turn, updates public perception based on the experiences of their patients. In the context of the market for rides, social media platforms can be viewed as the information aggregators. Each period, some fraction of riders communicate their experiences through social media, from which riders obtain the updated public perception of transportation mode quality. Of course, the updating process through social media is extremely complicated and could also be modelled. However, the additional complexity is unlikely to significantly affect the core results of the paper, and the requisite data are either unavailable or difficult to obtain.

3.1. Consumer Learning

Each trip consumers take, whether it is by taxi or rideshare service, is an experience good. That is to say, each consumer i experiences a ride quality of λ_{ijt}^{E} for a trip using mode j in week t. This observed quality is not necessarily the same as the true mean quality of mode j, which is λ_j. The difference in ride quality across consumers could be due to a variety of factors, such as different drivers, wait times, or rider idiosyncrasies.

In the model, consumers receive an experience signal that informs them about the quality of their chosen

[13] This modeling decision is common in the literature on discrete choice demand with consumer learning. Very few papers estimate models with either consumers or producers solving dynamic optimization problems (see Ching, 2010a; Osborne, 2011), and this author does not know of any example where both consumers and producers are forward-looking.

mode of transportation. The experience signal is:

$$\lambda_{ijt}^{E} = \lambda_j + \varepsilon_{ijt}^{E}, \quad \varepsilon_{ijt}^{E} \sim N(0, \sigma_E^2) \tag{1}$$

where ε_{ijt}^{E} is the rider-mode-week specific signal noise that is normally distributed with mean zero and variance σ_E^2. Since this is an experience signal, the information a rider obtains in period t does not affect their decision in that period. Instead, it impacts their perception of quality for the following period. Since taxis have been operating in NYC for over a century, I assume that the true quality of present-day taxi rides is known to consumers. Consequently, the prior over quality for taxis is equal to the perceived quality, which is equal to the true quality throughout the entire sample. There is no uncertainty about taxi quality, which implies no updating of the expected quality each period. The mode-level experience signal is the mean of all signals received by consumers that took a trip using a given mode. This signal is given by:

$$\overline{\lambda}_{jt}^{E} = \frac{1}{\kappa q_{jt}} \sum_{i=1}^{\kappa q_{jt}} \lambda_{ijt}^{E} \sim N\left(\lambda_j, \frac{\sigma_E^2}{\kappa q_{jt}}\right) \tag{2}$$

where q_{jt} is the number of trips using mode j at time t, and κ represents the fraction of the experience signals that are revealed to all riders. In the example where social media platforms act as information aggregators, this fraction can be thought of as a measure of social network closeness, how many experience signals are shared among consumers, or simply the fraction of riders that use social media platforms.

Under Bayesian updating, the consumer priors for the mean and variance of ride quality must be specified. I choose a normal prior, which is given by:[14]

$$\lambda_j \sim N(\lambda_{j0}, \sigma_{j0}^2) \tag{3}$$

The implication for the model is that in the first period, before receiving an experience signal, consumers believe that the true quality of mode j follows the distribution in (3). Through information aggregation and Bayesian updating, after a single period, the perceived mean and variance of quality in the second period are:

$$\lambda_{j2} = \frac{\sigma_{j1}^2}{\sigma_{j1}^2 \kappa q_{j1} + \sigma_E^2} \kappa q_{j1} \overline{\lambda}_{j1}^{E} + \frac{\sigma_E^2}{\sigma_{j1}^2 \kappa q_{j1} + \sigma_E^2} \lambda_{j0} \quad \text{and} \quad \sigma_{j2}^2 = \frac{1}{\frac{1}{\sigma_{j1}^2} + \frac{\kappa q_{j1}}{\sigma_E^2}} \tag{4}$$

Since experience signals do not impact choices in the same period they are received, the Bayesian posterior mean and variance are delayed by a single period.[15] The expression in (4) can be generalized

[14] A normal prior is common assumption in consumer learning literature, because the posterior distribution is also normal.
[15] If the reader wishes to check that the first period posterior is equivalent to the prior, then the single period lags can be substituted into (4) to verify.

to multiple periods, and rewritten in terms of the prior variance, which gives the following posterior mean and variance of perceived quality:

$$\lambda_{jt} = \frac{\sigma_{j0}^2}{\sigma_{j0}^2 \kappa Q_{jt} + \sigma_E^2} \sum_{\tau=1}^{t-1} \kappa q_{j\tau} \overline{\lambda}_{j\tau}^E + \frac{\sigma_E^2}{\sigma_{j0}^2 \kappa Q_{jt} + \sigma_E^2} \lambda_{j0} \tag{5}$$

$$\sigma_{jt}^2 = \frac{1}{\frac{1}{\sigma_{j0}^2} + \frac{\kappa Q_{jt}}{\sigma_E^2}}$$

where $Q_{jt} = \sum_{\tau=1}^{t-1} q_{j\tau}$ is the total number of trips up to period t for mode j. Also note that this formulation is for periods $t = 2, \ldots, T$, and for $t = 1$, we simply have the prior distribution.

Before introducing the consumer demand model, it is important to frame the Bayesian updating mechanism in the context of price indexes and quality adjustments. From the description of the mechanism above, one may think that the measure to be used in the quality adjustment factor is the true mode quality, λ_j. However, both rideshare services experienced substantial quality increases during the sample period due to continual improvements in various areas, such as their GPS accuracy and size of their driver pool. Therefore, it is more accurate to view the fixed value of λ_j as the quality toward which mode j converges over the course of the sample period. Thus, the prior for quality, λ_{j0}, can be viewed as the quality of mode j in the initial period, while λ_{jT} represents quality in the final period. The Bayesian updating mechanism computes the path of perceived quality between the initial and terminal periods using prices and market shares without restricting the path to a particular shape. Consequently, the path of perceived quality provides a flexible estimate for the change in mode quality. I discuss the variation used to identify perceived quality in more detail in Section 4.2.

3.2. Consumer Demand

I follow Ching (2010b) in modeling consumer utility. Each period, a consumer i chooses a mode of transportation, j, that maximizes their current period expected utility, $E[U_{ijt}|\mathcal{I}_t]$, where $\mathcal{I}_t$ is the information set common to all consumers that is available at time t. The indirect utility from a good j is given by:

$$U_{ijt} = \omega \lambda_{ijt}^E - \omega r \left(\lambda_{ijt}^E\right)^2 + \alpha_i p_{jt} + \mathbf{X}_{jt}\boldsymbol{\beta} + \xi_{jt} + \varepsilon_{ijt} \tag{6}$$

where λ_{ijt}^E is the individual experience signal, ω is the individual's value of the experience signal, and r is the risk coefficient. The price of a trip is p_{jt}, $\mathbf{X}_{jt}$ is a set of trip characteristics, and α and $\boldsymbol{\beta}$ are their respective coefficients. The mode-week-level demand shock unobserved by the econometrician is $\xi_{jt} \sim N(0, \sigma_{\xi}^2)$, where the variance is a parameter to estimate. The idiosyncratic error term, ε_{ijt}, follows a Type-I Extreme Value distribution with variance $\frac{\pi\mu^2}{6}$. As discussed in Nevo (1993), the demand shocks could be unobserved product quality characteristics or taste changes. In a demand model with consumer learning, the perceived quality term captures the unobserved product quality characteristics as they change over time, so that ξ_{jt} can be interpreted as taste changes. Since these taste changes influence demand directly, they cannot be viewed as pure price differences even though they are not inherently related to trip characteristics. Furthermore, because they are modeled as transitory shocks, their impact on price changes over time should be negligible.

Consumers are assumed to have CARA preferences over the uncertain portion of their utility (i.e., the quality signals), and linear preferences over the sub-utility that they observe with certainty.[16] Thus, the utility consumer i expects when choosing mode j at time t is:

$$\begin{aligned} E\left[U_{ijt}|\mathcal{I}_t\right] &= \omega E\left[\lambda_{ijt}^E|\mathcal{I}_t\right] - \omega r E\left[\lambda_{ijt}^E|\mathcal{I}_t\right]^2 - \omega r E\left[\left(\lambda_{ijt}^E - E\left[\lambda_{ijt}^E|\mathcal{I}_t\right]\right)^2|\mathcal{I}_t\right] \\ &+ \alpha_i p_{jt} + \mathbf{X}_{jt}\boldsymbol{\beta} + \xi_{jt} + \varepsilon_{ijt} \end{aligned} \tag{7}$$

The first line of (7) is the stochastic portion of the consumer's expected utility, while the second line is the deterministic portion. Note that the error terms, ξ_{jt} and ε_{ijt}, are deterministic for the consumer, but not the econometrician. Consequently, they enter the linear portion of the consumer's expected utility function. The third term in (7) can be simplified further to $\omega r\left(\sigma_E^2 + \sigma_{jt}^2\right)$. Since the signal is equal to the true quality in expectation, we can replace it in the utility function. This gives the following expected utility:

$$\begin{aligned} E\left[U_{ijt}|\mathcal{I}_t\right] &= \omega E\left[\lambda_j|\mathcal{I}_t\right] - \omega r E\left[\lambda_j|\mathcal{I}_t\right]^2 - \omega r\left(\sigma_E^2 + \sigma_{jt}^2\right) \\ &+ \alpha_i p_{jt} + \mathbf{X}_{jt}\boldsymbol{\beta} + \xi_{jt} + \varepsilon_{ijt} = V_{jt}^* + \varepsilon_{ijt} \end{aligned} \tag{8}$$

Finally, the expected utility for the outside option is allowed to change over time to capture potential quality improvements, but consumers do not receive explicit quality signals. The expected utility is given by:

$$E\left[U_{i0t}|\mathcal{I}_t\right] = \phi_0 + \phi_t t + \varepsilon_{i0t} = V_{0t}^* + \varepsilon_{i0t} \tag{9}$$

where $j = 0$ indicates the choice of the outside option, and t is the time trend used to capture quality changes over time as well as other changes that affect utility. In the context of the taxi and rideshare market in NYC, the outside option can be viewed as the subway or public bus system.

[16] Under CARA preferences, $\omega > 0$ permits the familiar interpretation of r, where $r > 0$, $r < 0$, or $r = 0$ imply riders are risk-averse, risk-seeking, or risk-neutral, respectively.

Since the idiosyncratic taste parameter is distributed Type-I Extreme Value, the familiar functional form for quantity shares obtains. The shares for mode j at time t is given by the following:

$$s_{jt}^{q} = Pr(j|p, ...) = \frac{\exp\{V_{jt}^{*}\}}{\exp\{V_{0t}^{*}\} + \sum_{k} \exp\{V_{kt}^{*}\}} \tag{10}$$

where s_{jt}^{q} is used to denote quantity shares in order to differentiate between quantity and expenditure shares.

4. Identification and Estimation

In this section, I discuss the estimation and identification of the structural model and the procedure followed to estimate the hedonic price equation used to compute the consumer learning price index.

4.1. Estimation

The estimation of the consumer learning price index follows a two-step procedure. In the first step, the structural demand model is estimated in order to recover the structural terms that represent unobserved quality. The second step estimates the baseline hedonic price equation with an additional function of these structural parameters, which is denoted g_{jt} for a given mode and week. This process is done separately for each of the seven time blocks.

4.1.1. Structural Demand Model

The standard issue with estimating demand models is the endogeneity of price. Given that the mode-level unobservables, $E\left[\lambda_j|\mathcal{I}_t\right]$ and ξ_{jt}, are likely correlated with prices, failing to account for this correlation will result in biased parameter estimates. The endogeneity issue is not a new one, with many techniques developed to deal with the problem appropriately; most notably Berry et al. (1995). However, due to consumer learning, the aggregate mode-level unobservables are serially correlated and non-stationary. Consequently, it is both computationally infeasible and not necessarily possible to estimate the model via GMM. Therefore, I follow the method used in Ching (2010b), which uses simulated maximum likelihood to estimate the joint distribution of prices and quantities. Due to computational limitations, estimating a full supply-side oligopoly model with forward-looking firms is eschewed in favor of a simple, hedonic pricing equation given by the following:

$$\ln p_{jt} = \ln\left(h_j\left(t, \mathbf{Z}_{jt}, E\left[\lambda_j|\mathcal{I}_t\right], \sigma^2_{jt}, \xi_{jt}; \boldsymbol{\theta}_p\right)\right) + \nu_{jt}, \quad \nu_{jt} \sim N(0, \sigma^2_\nu) \tag{11}$$

where $\boldsymbol{\theta}_p$ is the set of pricing equation parameters, and ν_{jt} is the prediction error. It is important to distinguish between the pricing policy function in (11) and the second step hedonic price equation in (16). As discussed above, the purpose of the pricing policy function is for identification of the price coefficient on the demand side of the structural model. Although one could use the estimates from the pricing policy function to compute the hedonic price index, there is one major drawback to this approach. The time dummies from the standard hedonic model must be incorporated into the pricing policy function. When the time window of the sample is relatively large, these extra parameters drastically increase the computational burden of the model. This issue becomes particularly pronounced when the prediction error, ν_{jt}, is very small and causes optimization issues. Thus, I include a time trend in the pricing policy function and introduce the time dummies in the second step of the estimation procedure.

The demand-side equation computes quantity demanded, q_{jt}, using the total market size M_t and the shares from (10). The quantity demanded follows a multinomial distribution with sampling errors, η_{jt}. The large sample size for each market permits the assumption that the multinomial distribution approximates the multivariate normal. The estimating equation is given by the following:

$$q_{jt} = M_t s^q_{jt}\left(p_{jt}, \mathbf{X}_{jt}, E\left[\lambda_j|\mathcal{I}_t\right], \sigma^2_{jt}, \xi_{jt}; \boldsymbol{\theta}_q\right) + \eta_{jt} \tag{12}$$

where $\boldsymbol{\theta}_q$ is the set of demand equation parameters. The sampling errors from the demand equation are distributed $\eta_t \sim N\left(\mathbf{0}, \Sigma_{\eta_t}\right)$, where the variance-covariance matrix of the sampling errors is:

$$\Sigma_{\eta_t} = M_t \begin{pmatrix} s^q_{1t}(1 - s^q_{1t}) & -s^q_{1t}s^q_{2t} & -s^q_{1t}s^q_{3t} \\ -s^q_{2t}s^q_{1t} & s^q_{2t}(1 - s^q_{2t}) & -s^q_{2t}s^q_{3t} \\ -s^q_{3t}s^q_{1t} & -s^q_{3t}s^q_{2t} & s^q_{3t}(1 - s^q_{3t}) \end{pmatrix} \tag{13}$$

The joint likelihood of observing a pair of quantity-price vectors $(\mathbf{q}_t, \mathbf{p}_t)$ is given by the product of the conditional likelihoods:

$$\ell\left(\mathbf{q}_t, \mathbf{p}_t|\boldsymbol{\chi}_t, E\left[\lambda_j|\mathcal{I}_t\right], \xi_t; \boldsymbol{\theta}\right) = f_q\left(\mathbf{q}_t|\mathbf{p}_t, \boldsymbol{\chi}_t, E\left[\lambda_j|\mathcal{I}_t\right], \xi_t; \boldsymbol{\theta}_q\right) f_p\left(\mathbf{p}_t|\boldsymbol{\chi}_t, E\left[\lambda_j|\mathcal{I}_t\right], \xi_t; \boldsymbol{\theta}_p\right) \tag{14}$$

where $\boldsymbol{\chi}_t$ contains all observable covariates except for prices. It is important to note that although the variance of perceived quality updates every period, it is entirely determined by observed covariates and estimated coefficients. One can check (5) to see that it is not a function of the experience signal noise ε^E_{ijt}. Thus, when the sampling errors are very small, as noted in Ching (2010b), the unobservables that explain most of the difference between the model and the data are $E\left[\lambda_j|\mathcal{I}_t\right]$ and ξ_{jt}. The likelihood function for a given time block is:

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{t=1}^{T} \ell\left(\mathbf{q}_t, \mathbf{p}_t | \boldsymbol{\chi}_t, E\left[\lambda_j | \mathcal{I}_t\right], \xi_t; \boldsymbol{\theta}\right) \tag{15}$$

which is the likelihood of jointly observing the sequence of price and quantity vectors from time 1 to time T. The Bayesian learning process causes the perceived quality term to be autocorrelated, which makes integrating out the unobservables computationally intractable. Thus, the model is estimated by simulated maximum likelihood where S_E simulated draws of the signal noise, ε^E, and S_ξ draws of the demand shock, ξ, are used to compute the likelihood function in (15). I discuss the simulation procedure in more detail in Appendix D.

4.1.2. Consumer Learning Hedonic Price Model

Once the estimates from the structural demand model are obtained, the time-varying unobserved quality function, g_{jt}, can be formulated. The hedonic price model used to compute the consumer learning price index is given below:

$$\begin{aligned} \ln p_{jt} = \delta_0 + \delta_t + \mathbf{Z}_{jt}\boldsymbol{\beta} + \underbrace{\gamma_1 \widehat{E\left[\lambda_j | \mathcal{I}_t\right]} + \gamma_2 \hat{\sigma}_{jt}^2 + \gamma_3 \left(\hat{\xi}_{jt} - \hat{\xi}_t^{taxis}\right)}_{g_{jt}} \\ + \underbrace{\varphi\left(Adjusted\ Multiplier_{jt}\right)}_{f_{jt}} + e_{jt}, \quad e_{jt} \sim N(0, \sigma_e^2) \end{aligned} \tag{16}$$

The sum of the perceived quality and variance estimates and the unobserved demand shocks from the structural demand model are included as the unobserved quality function g_{jt}. Each of these terms is computed by averaging over the simulated error draws, which is discussed in detail in Appendix D. I proxy for the pure price difference function, denoted f_{jt}, using the adjusted surge multiplier, $Adjusted\ Multiplier_{jt}$. Since this term captures changes in pricing by Lyft and Uber during surge periods, it is a clear difference in pricing strategies between the rideshare companies and taxis. Furthermore, the surge pricing strategies of Lyft and Uber are not identical, so pure pricing differences between the two rideshare companies are also captured.

4.2. Identification

Now that I have introduced the econometric model and estimation procedure, I will discuss identification. I ignore the coefficients familiar from standard discrete choice demand models, and instead focus on the learning parameters.

First, the true quality mean for one ride service must be fixed, so I choose to set the true quality mean for taxis equal to zero. Since the true mean for taxis is known at the start of the sample period, this means the prior mean for taxis is also zero as well as the perceived mean and variance. Variation in the time series of the cumulative number of trips and the simulated learning errors, along with the function form of the uncertain portion of the expected utility function can be used to identify the priors and posterior quality means and signal variances, as well as ω and the risk coefficient r. This can be checked by plugging the cumulative number of trips and simulated errors into the stochastic sub-utility. One can also refer to Ching (2010b) for an in-depth discussion of the identification of the learning parameters. From the expression for perceived quality variance in (5), it can be determined that the rate at which public perception of ride service quality converges to the true mean is governed by the ratio, $\frac{\kappa}{\sigma_E^2}$. Thus, for at least one of the two rideshare companies, either κ or σ_E^2 must be fixed. I choose to fix κ for Uber, and also assume that σ_E^2 is the same across companies. Finally, the standard deviation for the product-week demand shock, ξ_{jt}, can be identified, because the perceived quality variance tends to zero in the long run.

Aside from being able to identify the parameters from variation in the data, it is also important to understand how the identified learning parameters constitute unobserved quality rather than pure price changes. In the demand model, the learning parameters capture unobserved differences in market shares conditional on the observed price in the same period. Thus, they influence demand directly rather than indirectly through their impact on price. For example, consider a case in which a consumer is faced with a choice between modes that are identical from the point of view of the econometrician (i.e., same characteristics and prices). The observed choice is determined by differences in the unobservable factors that are captured by the learning parameters and the mode-week-level demand shocks. Since pure price differences are reflected in the observed price, the learning parameters are picking up variation in unobservable quality characteristics that are exogenous to price, such as driver quality and waiting time, that influence the decision made by the consumer. Therefore, these terms should be treated as quality characteristics in the second stage hedonic regression and included in the quality adjustment factor used to compute quality-adjusted prices.

5. Results

In this section, I present the parameter estimates from the reduced-form and structural models. Next, I discuss the estimated learning parameters and the rates of diffusion in more detail. Then, I examine the different methods for controlling for unobserved quality by comparing the quality-adjusted unit value price indexes.

5.1. Parameter Estimates

The model is estimated separately for each time block for both the structural demand and reduced-form hedonic models. First, I present a subset of the demand estimates for each time block. Then, I show a subset of the hedonic pricing equation estimates for the baseline and mode fixed effects reduced-form specifications, as well as the second stage of the consumer learning model.

The estimates from the structural demand model are presented in Table 3. The price coefficients are relatively small across all time blocks, with some being positive, and others not being statistically significant. These results point to the price of a trip not being the main driving force behind consumers' choice of transportation mode. The more interesting results for the focus of the paper are the mode quality estimates for both the prior and true quality, and the variation in coefficients across time blocks. The differences across time blocks offers support for the decision to treat each period as a distinct market with different consumers and different purchasing patterns. For example, the large variances for both the demand shocks and experience signals during bar hours should not be surprising, because these are periods of high, intense demand with consumers that are least likely to make optimal utility-maximizing decisions. The Weekday Evening period exhibits the other extreme, with very small demand shock and experience signal variances that are ultimately not statistically significant.

The negative values for the mode quality priors and true means indicate that after controlling for price and mode characteristics the discrepancy in market shares for the rideshare companies relative to taxis implies large negative values. These negative values are indicative of both the nascency of the rideshare companies during the sample, as well as their viability in Manhattan. It is important to note that these results are specific to the Manhattan market for rides and should not be applied more generally. Despite the negative values, in every time block, the quality for both rideshare companies increases over the course of the sample. As previously mentioned, the change in quality is more important than the levels when using the estimates to compute the price index.

	Bar Hours	Evening Rush	Morning Rush	Weekday Day	Weekday Evening	Weekend Day	Weekend Evening
Price (α)	-0.029	0.005	-0.043	-0.093	-0.008	-0.005	0.034
	(0.025)	(0.002)*	(0.003)**	(0.002)**	(0.006)	(0.007)	(0.008)**
Experience Signal	275.180	3.222	4.145	0.017	4.3e-4	0.041	2.219
Variance (σ^2_E)	(42.316)**	(0.444)**	(0.523)**	(0.005)**	(7.09e-4)	(0.021)*	(0.485)**
Variance of Mode	3.233	0.027	0.055	0.118	0.049	0.040	0.081
Demand Shocks (σ^2_ξ)	(0.168)**	(0.001)**	(0.003)**	(0.004)**	(0.038)	(0.003)**	(0.006)**
Fraction of Signals Revealed (κ)							
Lyft	0.001	0.010	0.759	3.86e-6	1.57e-8	1.4e-4	0.002
	(6.53e-4)*	(0.003)**	(2.453)	(7.88e-7)**	(2.71e-8)	(4.4e-5)**	(0.001)**
Uber	1e-4	1e-4	1e-4	1e-8	1e-4	1e-6	1e-4
	—	—	—	—	—	—	—
Mode Quality Mean Prior (λ_{j0})							
Lyft	-161.135	-12.244	-13.171	-33.047	-22.163	-15.839	-8.290
	(4.969)**	(0.373)**	(0.618)**	(0.499)**	(6.469)**	(2.625)**	(0.606)**
Uber	-81.631	-5.451	-6.054	-17.120	-12.878	-8.725	-4.461
	(2.256)**	(0.142)**	(0.245)**	(0.182)**	(3.752)**	(1.433)**	(0.316)**
Mode Quality True Mean (λ_j)							
Lyft	-108.789	-8.061	-8.477	-22.918	-13.820	-11.384	-6.052
	(2.916)**	(0.213)**	(0.354)**	(0.273)**	(4.093)**	(1.873)**	(0.433)**
Uber	-74.917	-4.705	-4.922	-16.254	-9.563	-6.681	-3.585
	(1.826)**	(0.114)**	(0.189)**	(0.212)**	(2.799)**	(1.109)**	(0.268)**
Mode Quality Prior Variance (σ^2_{j0})							
Lyft	194.263	0.322	0.008	1.471	6.372	0.438	1.022
	(77.408)*	(0.088)**	(0.024)	(0.306)**	(3.607)	(0.203)*	(0.347)**
Uber	68.553	0.393	0.479	1.490	7.54e-6	0.107	0.066
	(13.697)**	(0.059)**	(0.101)**	(0.187)**	(1.23e-5)	(0.047)*	(0.023)**
LLH	-1756	-1775	-1925	-1955	-1808	-1780	-1713

Table 3. Demand Estimates

Standard errors in parentheses. * $p < .05$, ** $p < .01$ Notes: (i) All estimates and standard errors smaller than .001 are abbreviated using scientific notation. (ii) Results are for simulation using $S_E = S_\xi = 50$.

The next set of results, presented in Table 4, are those from the various hedonic specifications: (i) the baseline model, (ii) the mode fixed effects model, and (iii) the consumer learning model. The quality-adjustment factors used to compute the QAUV price indexes are derived from these results.[17] The baseline hedonic model refers to the specification estimated in Section 2.3, the mode fixed effects hedonic model estimates the same specification with the addition of mode dummies for Lyft and Uber, and the consumer learning hedonic model contains the results from estimating (16). Each regression is weighted by expenditure shares and estimated separately for each time block. Recall that the adjusted multiplier is the proxy for the pure price function included in the price index in addition to the time dummies. The coefficient for this variable is relatively consistent across models and time blocks, except for the Morning Rush and Weekday Day time blocks. In both these time blocks, the coefficient is much smaller in the consumer learning model than the other two models. The implication is that there is some unobserved variation in prices over time that is correlated with the adjusted multiplier for which the first two models do not control. This omitted variable bias is an example of the indirect effect unobserved quality can have on the price index when it is not addressed. The mode fixed effects have mixed results, with some time blocks exhibiting statistically significant estimates while others do not. Interestingly, most of the perceived quality and variance coefficients are statistically significant and have the same sign, while the unobserved demand shock is not significant in any of the time blocks. Finally, I include three metrics by which to measure the ability of each model to explain variation in prices. These are the adjusted R^2, the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC). As expected, the adjusted R^2 in the fixed effects model is larger than in the baseline specification. It is also larger for the consumer learning model than the fixed effects model in all but the Weekday Evening time block, which exhibited estimates of perceived quality that were not statistically significant in both the structural demand and hedonic models. The same pattern can be seen in both the AIC and BIC, with the consumer learning model exhibiting lower (more accurate) values than the baseline and fixed effects models for all time blocks except Weekday Evening. Thus, the results indicate that the perceived quality parameters capture unobserved variation that is likely to impact the QAUV price index.

[17] see Section A in the appendix for a complete discussion on the construction of the quality adjustment factors and the QAUV price indexes

	Bar Hours	Evening Rush	Morning Rush	Weekday Day	Weekday Evening	Weekend Day	Weekend Evening
			Baseline Hedonic Model				
Adjusted Multiplier	0.799	0.794	0.760	1.083	0.907	0.723	0.843
	(0.059)**	(0.044)**	(0.053)**	(0.129)**	(0.102)**	(0.050)**	(0.069)**
Adjusted R^2	0.854	0.937	0.956	0.904	0.779	0.914	0.871
AIC	-557.971	-502.499	-391.057	-471.196	-480.308	-517.550	-499.464
BIC	-242.799	-187.327	-75.885	-156.023	-165.136	-202.377	-184.292
			Mode Fixed Effects Hedonic Model				
Mode Fixed Effects (γ_j)							
Lyft	0.019	-0.129	-0.243	-0.503	-0.223	-0.072	-0.122
	(0.057)	(0.054)*	(0.088)**	(0.080)**	(0.078)**	(0.057)	(0.063)
Uber	-0.035	-0.049	-0.028	-0.153	-0.088	-0.011	-0.084
	(0.029)	(0.027)	(0.047)	(0.045)**	(0.045)	(0.027)	(0.032)**
Adjusted Multiplier	0.785	0.773	0.733	1.027	0.876	0.725	0.815
	(0.058)**	(0.045)**	(0.053)**	(0.126)**	(0.102)**	(0.050)**	(0.068)**
Adjusted R^2	0.859	0.938	0.960	0.927	0.789	0.914	0.875
AIC	-565.285	-507.206	-416.249	-540.922	-490.361	-517.451	-506.449
BIC	-243.030	-184.951	-93.994	-218.667	-168.106	-195.196	-184.194
			Consumer Learning Model				
Perceived Quality ($E[\lambda_j \| \mathcal{I}_t]$)	0.002	0.025	0.026	0.015	-0.007	0.015	0.012
	(3.55e-4)**	(0.005)**	(0.008)**	(0.002)**	(0.004)	(0.003)**	(0.008)
Perceived Quality Variance (σ_{jt}^2)	0.003	0.650	0.661	0.221	-0.024	2.367	0.712
	(4.18e-4)**	(0.063)**	(0.067)**	(0.014)**	(0.020)	(0.245)**	(0.196)**
Unobserved Demand Shock ($\xi_{jt} - \xi_t^{taxis}$)	0.001	-2.52e-4	0.005	0.004	-0.516	-0.047	0.016
	(0.002)	(0.012)	(0.014)	(0.161)	(1.724)	(0.123)	(0.023)
Adjusted Multiplier	0.741	0.717	0.524	0.767	0.946	0.738	0.813
	(0.052)**	(0.036)**	(0.050)**	(0.090)**	(0.103)**	(0.041)**	(0.068)**
Adjusted R^2	0.892	0.961	0.972	0.963	0.782	0.945	0.879
AIC	-633.363	-625.549	-506.271	-710.940	-481.657	-631.252	-514.723
BIC	-307.567	-299.753	-180.474	-385.144	-155.861	-305.456	-188.927

Table 4. Hedonic Price Equation Estimates

Standard errors in parentheses. * $p < .05$, ** $p < .01$ Notes: (i) All estimates and standard errors smaller than .001 are abbreviated using scientific notation.

5.2. Consumer Learning and Diffusion

The change in consumers' perception of rideshare service quality is essential to determining the degree of bias in the various QAUV price indexes computed in this paper. As discussed in Section 1, a model with mode fixed effects offers no indication for whether the fixed effect is picking up variation in unobservable quality, pure price differences, or both. As we will see in Section 5.3, how these fixed effects are categorized substantially changes the resulting price index.

The estimated perceived quality for Lyft and Uber is plotted separately for each time block in Figure 4. In all cases, for both rideshare services, the perceived quality is increasing over the course of the sample period. The differences in magnitudes and rates of increase across time blocks and rideshare services indicates that not only are Lyft and Uber distinct in the eyes of consumers, but also that the time blocks represent different markets, each with their own idiosyncrasies. The differences in slopes are particularly important, since the change in perceived quality with respect to time has a direct impact on the price index. Additionally, perceived quality is concave with respect to time towards the end of the sample period, but not throughout the entire sample. This illustrates both the flexibility of the Bayesian updating mechanism in capturing quality improvements, as well as the diminishing marginal improvements that can be made to a product over a longer time frame. The flexibility is most evident in time blocks like the Weekday Day period, where Lyft experiences large quality increases in the middle of the sample, while the quality improvements for Uber are relatively non-smooth in comparison to other time blocks. These features illustrate that the learning mechanism does not place major restrictions on the shape of the perceived quality path, which reduces the likelihood that the consumer learning modeling assumptions will introduce bias into the quality-adjustment estimates.

The perceived signal variance for Lyft and Uber is plotted separately for each time block in Figure 5. As mentioned previously, there has been substantial learning over the course of the sample period, which is evidenced by the fact that the signal variance becomes very close to zero for most time blocks. This can be interpreted as another dimension along which quality improves, because it represents an improvement in the consistency with which a level of quality is provided across drivers within a rideshare company.

Although the actual values of the perceived quality estimates carry very little meaning, the change in these values over time greatly impacts the price index through the second stage hedonic regressions. The CAGR for perceived quality for each rideshare company is presented in Table 5. The most noticeable feature of the growth rates is that quality increase much more for Lyft than Uber. The main reason is that Lyft entered most markets later than Uber, not just the NYC market, which leads to it experiencing faster quality improvements typically seen earlier in the life of a good or service.

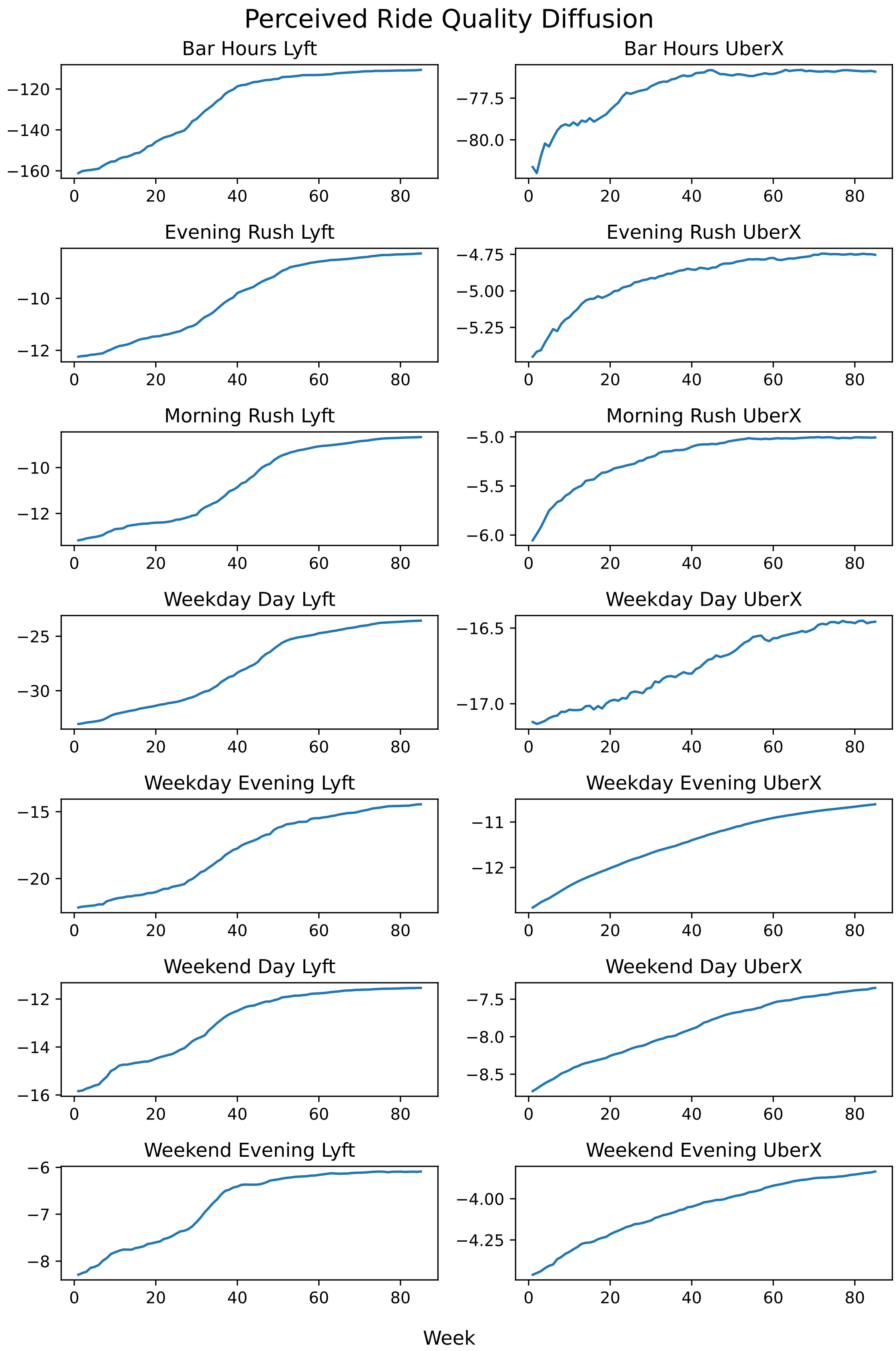

Figure 4. Perceived Quality Diffusion

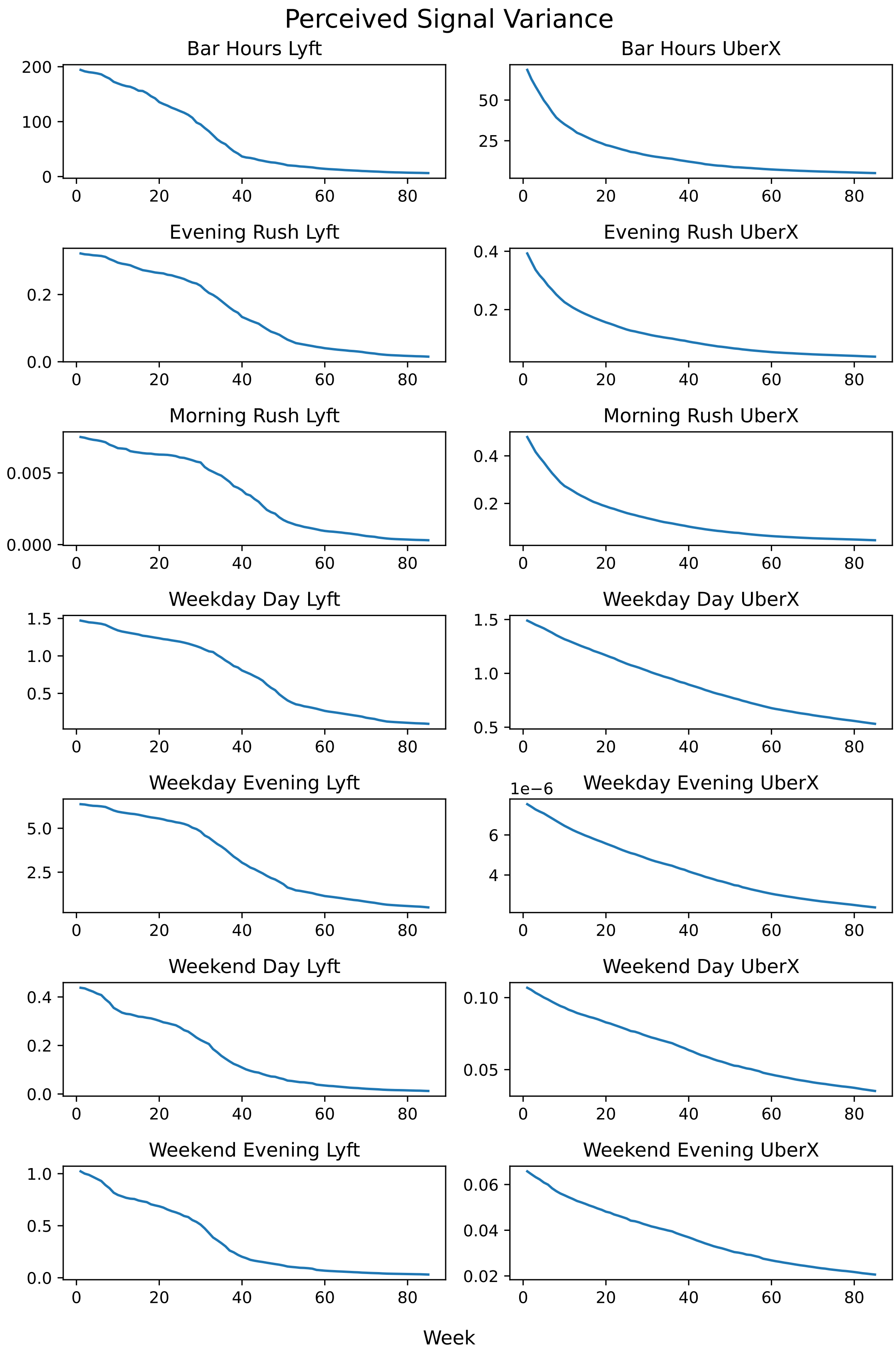

Figure 5. Signal Variance Diffusion

Time Block	Lyft Quality CAGR	Uber Quality CAGR
Bar Hours	18.15	4.23
Evening Rush	18.74	7.65
Morning Rush	19.71	10.26
Weekday Day	16.68	2.35
Weekday Evening	20.05	10.43
Weekend Day	15.84	9.37
Weekend Evening	15.49	8.37

Table 5. CAGR in Perceived Quality

Notes: All values are percent.

5.3. Price Indexes

In this section, I compare the quality-adjusted unit value price indexes computed from the hedonic estimates, as well as a unit value price index. The QAUV Baseline price index uses the estimates from the baseline hedonic model to compute quality adjustment factors. There are two QAUV price indexes computed using the estimates from the mode fixed effects model. The QAUV FE Pure Price index, treats the fixed effects as if they capture only pure price effects, while the other, QAUV FE Quality, assumes they capture only mode quality. These polar assumptions are used to illustrate the large impact they have on the price index, and the need for a more sophisticated approach to separating unobserved quality from pure price differences. Finally, QAUV CL is the price index computed using the structural demand estimates to adjust for mode quality. Figure 6, plots the price indexes separately for each time block, and the CAGR for each index is given in Table 6.

First, the large difference between the unit value price index and QAUV indexes shows that the modes are clearly not the same, and that the quality differences are non-negligible. Second, the stark differences in the mode fixed effects indexes, across all time blocks, highlight the importance of separating pure price terms from unobserved quality terms. In fact, the price index categorizing the mode fixed effects as pure price terms is almost always closer to the baseline QAUV price index than the index that assumes the fixed effects represent quality. Finally, index computed using the estimated quality terms from the structural demand model is distinct from the other price indexes to varying degrees across the time blocks, but it is usually closer to the baseline model and the pure price model than it is to the mode fixed effects quality model. Even more importantly, it tends to be above the index for the pure price model rather than between it and the pure quality model, which highlights the fact that the fixed effects price indexes do not capture all the unobserved variation in quality; namely variation over time.[18]

Time Block	Unit Value	QAUV Baseline	QAUV FE Pure Price	QAUV FE Quality	QAUV CL
Bar Hours	1.48	-1.02	-1.54	-6.01	-0.69
Evening Rush	4.80	-2.04	-2.55	-5.42	-1.37
Morning Rush	-1.20	3.15	2.84	2.00	0.17
Weekday Day	5.33	1.49	0.37	-2.37	0.69
Weekday Evening	2.47	-1.48	-3.26	-6.35	0.17
Weekend Day	2.84	-1.44	-1.57	-5.81	-0.05
Weekend Evening	2.38	-0.51	-2.37	-5.88	-0.82

Table 6. Price Index CAGR Table

Notes: All values are percent change.

Table 6 offers further insight into the differences across price indexes. First, the time blocks have different growth rates within price index method, which implies that rides during one time block can be considered different goods from those in another time block. This should not be too surprising. For example, rides during bar hours and those during morning rush hour serve entirely different purposes, so different price trends should be expected. The growth rates highlight the large difference in the fixed effects price indexes, with the magnitude of the CAGR for the pure quality model being at least twice that of the pure price model in almost every time block. The growths rate for the consumer learning model are relatively small for all time blocks. Again, the morning rush period provides a particularly interesting comparison; with the other models exhibiting relatively large price increases, while the consumer learning model finds almost no price increase.

6. Conclusion

This paper has presented a new framework for estimating quality adjustment factors for QAUV price indexes that separates unobserved quality from pure pricing differences. The method utilizes a Bayesian updating mechanism to capture unobserved variation in demand that is exogenous to prices, which allows

[18] It is also true that the price indexes derived from the polar assumption on the treatment of the fixed effects do not necessarily provide bounds for an index that assumes some portion of each fixed effect is quality and some is pure price difference (i.e., an index that relaxes the polar assumptions).

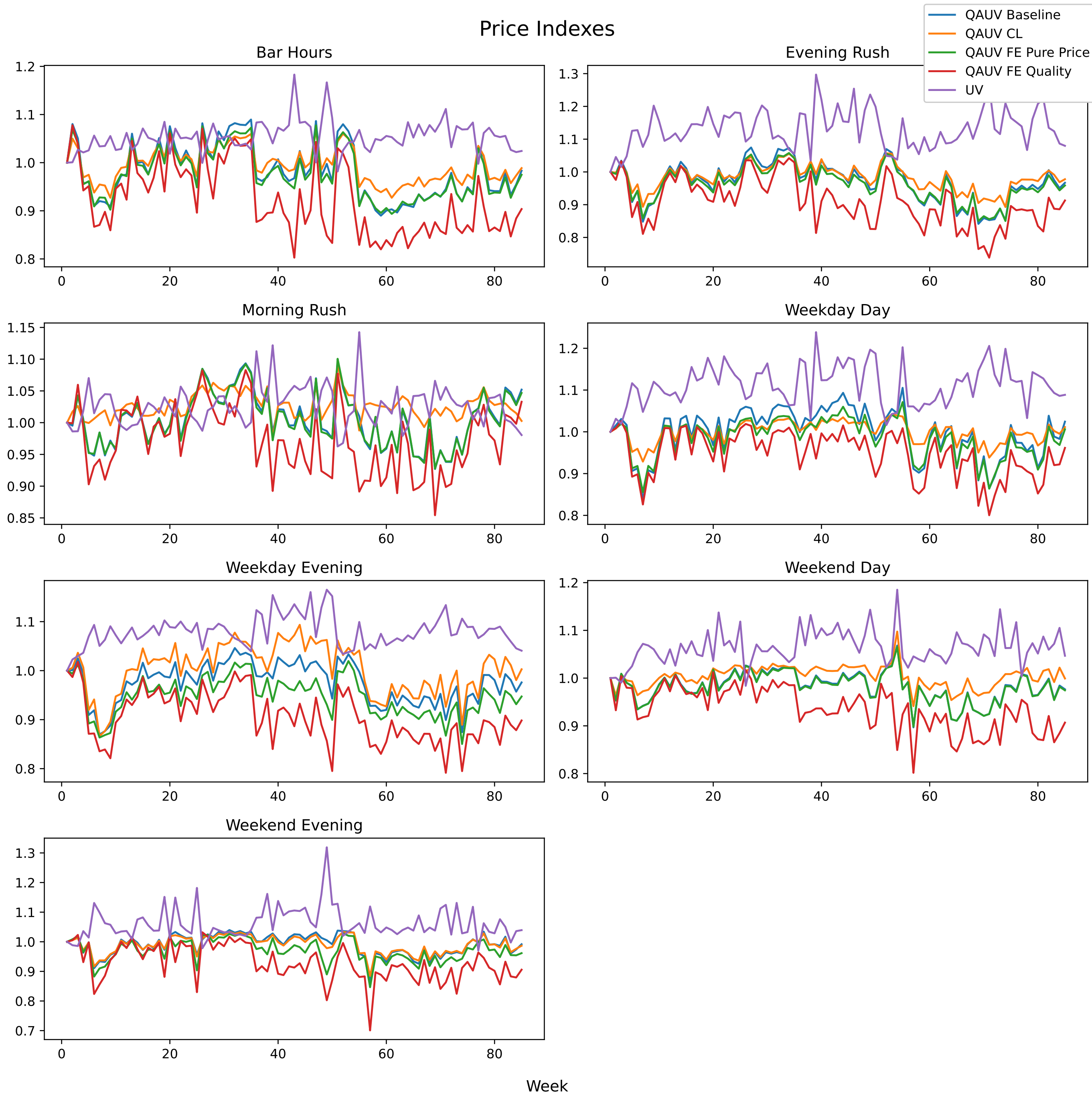

Figure 6. Price Indexes

for the identification of unobserved quality adjustment terms. The approach is assessed by computing price indexes for taxis and rideshare services in Manhattan using a variety of methods, which are then compared to the QAUV consumer learning price index. The price indexes computed using the proposed method exhibited consistent differences, across a variety of distinct time blocks, relative to the existing methods. Most notably, the CAGRs calculated from the consumer learning model indicated differences in price growth estimates that were as large as 6.5%, in some cases, depending on the QAUV method used in the comparison. The implication being that the proposed method is not just algebraically different, but that the difference in numerically meaningful. However, it is important to recognize that the framework will be more useful when there are important unobserved quality characteristics that not only impact price but change over the course of the sample. In settings where these issues are not substantial, the improvement over other methods is likely not large enough to justify the time-consuming estimation procedure. Therefore, a careful assessment of the market for a good or service is essential to determining whether the framework will be helpful.

A. Derivations

Following de Haan and Krsinich (2018), I derive the consumer learning estimate of the quality adjustment factor for mode j at time t relative to a hypothetical mode b. The hypothetical mode is constructed to have constant quality over the entire sample period, which is done by averaging the vector of characteristics over all modes and time periods. Additionally, the unobserved quality function is set to equal that of taxis, which is zero for all periods. Thus, the estimate for the quality adjustment factor for mode j at time t is:

$$\check{\psi}_{j/b,t} = \exp\{\left(\mathbf{Z}_{jt} - \overline{\mathbf{Z}}\right)\check{\boldsymbol{\beta}} + \check{g}_{jt} - \check{g}_{t}^{taxis}\} \tag{17}$$

where $\overline{\mathbf{Z}}$ is the vector of characteristics averaged over all modes and time periods. In the case of the standard hedonic model, the quality adjustment factor is constructed in the same fashion, but there are no unobserved quality functions. The quality-adjusted predicted price is:

$$\check{\check{p}}_{jt} = \frac{p_{jt}}{\check{\psi}_{j/b,t}} = p_{jt}\exp\{\left(\overline{\mathbf{Z}} - \mathbf{Z}_{jt}\right)\check{\boldsymbol{\beta}} + \check{g}_{t}^{taxis} - \check{g}_{jt}\} = \exp\{\check{\delta}_0 + \check{\delta}_t + \overline{\mathbf{Z}}\check{\boldsymbol{\beta}} + \check{f}_{jt} + \check{e}_{jt}\} \tag{18}$$

where the last equality follows from the unobserved functions for taxis being equal to zero. The bilateral quality-adjusted unit value price index is then given by:

$$P_{t-1,t}^{QAUV,CL} = \frac{\sum_j \check{\check{p}}_{jt} q_{jt} \Big/ \sum_j q_{jt}}{\sum_j \check{\check{p}}_{jt-1} q_{jt-1} \Big/ \sum_j q_{jt-1}} = \frac{\sum_j s_{jt}^{q} \check{\check{p}}_{jt}}{\sum_j s_{jt-1}^{q} \check{\check{p}}_{jt-1}} \tag{19}$$

B. Additional Tables

	Time Block	Mean	Std. Error	Min	Max	Median
Adjusted Multiplier	Bar Hours	1.157	0.199	1.000	2.123	1.102
	Evening Rush	1.195	0.251	1.000	2.513	1.098
	Morning Rush	1.250	0.270	1.000	2.299	1.200
	Weekday Day	1.095	0.128	1.000	1.962	1.048
	Weekday Evening	1.102	0.148	1.000	2.002	1.048
	Weekend Day	1.106	0.158	1.000	2.058	1.039
	Weekend Evening	1.096	0.169	1.000	2.222	1.029
Distance	Bar Hours	1.960	0.531	0.952	3.898	1.841
	Evening Rush	1.698	0.504	0.810	6.227	1.703
	Morning Rush	1.949	0.446	1.000	3.755	1.994
	Weekday Day	1.834	0.466	1.288	3.456	1.831
	Weekday Evening	1.790	0.415	1.329	3.453	1.752
	Weekend Day	1.969	0.698	1.320	7.959	1.904
	Weekend Evening	1.875	0.586	1.316	5.400	1.789
Duration (Minutes)	Bar Hours	11.215	3.107	6.686	27.554	11.086
	Evening Rush	14.844	4.022	7.155	32.133	14.973
	Morning Rush	13.095	3.299	7.932	22.223	13.676
	Weekday Day	16.257	4.080	10.142	40.375	16.487
	Weekday Evening	11.751	2.561	7.535	20.653	11.964
	Weekend Day	12.781	4.414	7.310	49.885	12.617
	Weekend Evening	12.704	3.539	7.665	35.986	12.627
Inside Share	Bar Hours	0.333	0.404	6.02e-5	0.950	0.098
	Evening Rush	0.333	0.404	2.9e-5	0.942	0.097
	Morning Rush	0.333	0.421	6.88e-6	0.961	0.076
	Weekday Day	0.333	0.420	2.83e-5	0.963	0.079
	Weekday Evening	0.333	0.407	5.23e-5	0.948	0.096
	Weekend Day	0.333	0.422	2.46e-6	0.973	0.076
	Weekend Evening	0.333	0.402	6.63e-5	0.948	0.104
Surge Share	Bar Hours	0.016	0.037	0.000	0.256	0.003
	Evening Rush	0.012	0.031	0.000	0.241	0.002
	Morning Rush	0.031	0.109	0.000	1.000	0.003
	Weekday Day	0.003	0.008	0.000	0.073	0.001
	Weekday Evening	0.004	0.009	0.000	0.070	0.001
	Weekend Day	0.007	0.018	0.000	0.175	0.001
	Weekend Evening	0.007	0.028	0.000	0.343	0.001

Table 7. Summary Statistics by Time Block

C. Additional Figures

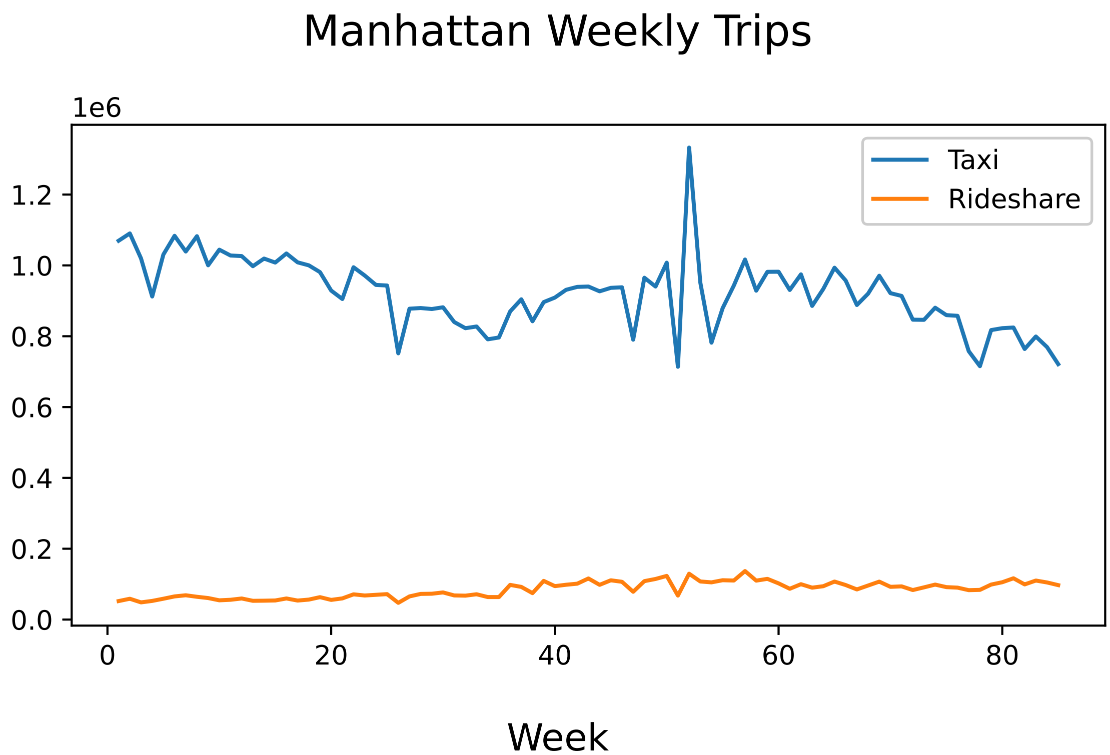

Figure 7. Weekly Trip Counts in Manhattan

D. Simulation Procedure

D.1. Simulation

In order to evaluate the likelihood function, both the signal noise and mode-level demand shocks must be integrated out. Since the mode-level demand shocks are i.i.d., the full expression for the likelihood function in (15) can be written as the following:

$$\mathcal{L}(\boldsymbol{\theta}) = \int \left[\prod_{t=1}^{T} \int \ell\left(\mathbf{q}_t, \mathbf{p}_t | \boldsymbol{\chi}_t, E\left[\boldsymbol{\lambda} | \mathcal{I}_t\right]^s, \boldsymbol{\xi}_t^r; \boldsymbol{\theta}\right) \, dF(\boldsymbol{\xi}_t) \right] \, dF(\boldsymbol{\varepsilon}_t^E) \tag{20}$$

Since it is computationally infeasible to evaluate these integrals numerically, they must be simulated. The simulated likelihood function is obtained by taking S_E draws of the signal noise, ε^E_{jt}, and S_ξ draws of the mode-level demand shocks, ξ_{jt}. Each sequence of signal noise draws, $\left\{\varepsilon^{E,s}_{j\tau}\right\}_{\tau=0}^{T-1}$, where s denotes a single draw, is used to recursively generate a simulated sequence of perceived quality vectors, $\{E\left[\boldsymbol{\lambda}|\mathcal{I}_\tau\right]^s\}_{\tau=1}^{T}$. The draws of mode-level demand shocks are denoted by r, so that the simulated likelihood is given by:

$$\mathcal{L}\left(\boldsymbol{\theta}\right) = \frac{1}{S_E}\sum_{s=1}^{S_E}\left[\prod_{t=1}^{T}\left(\frac{1}{S_\xi}\sum_{r=1}^{S_\xi}\ell\left(\mathbf{q}_t, \mathbf{p}_t|\boldsymbol{\chi}_t, E\left[\boldsymbol{\lambda}|\mathcal{I}_t\right]^s, \boldsymbol{\xi}^r_t; \boldsymbol{\theta}\right)\right)\right] \tag{21}$$

The simulated sequence of perceived quality for a given mode, j, is computed using the formula in (5). This requires computing the mode-level signal for each simulation draw. For a given draw, s, the simulated mode-level signal is given by:

$$\overline{\lambda}^{E,s}_{jt} = \hat{\lambda}_j + \varepsilon^{E,s}_{jt}\frac{\sigma_E}{\sqrt{\kappa_j q_{jt}}} \tag{22}$$

where $\hat{\lambda}_j$ is the estimate of the true mode quality.

D.2. Kernel Smoothing

The second part of the simulation procedure is in regard to the sampling and prediction errors introduced into the estimating equations that form the likelihood function. These errors are η_t for the demand equation, and ν for the pricing policy function. They ensure that the simulated likelihood function is differentiable and assigned positive density for each simulation draw. If the variance of these errors is small enough, derivative-based optimization procedures tend to suffer. When this is an issue, a kernel-smoothed frequency simulator similar to smoothing procedure in McFadden (1989) can be implemented. This method simply multiplies the variance of the sampling errors by a constant, k, which leads to the following simulated likelihood function:

$$\mathcal{L}\left(\boldsymbol{\theta}\right) = \frac{1}{S_E}\sum_{s=1}^{S_E}\left\{\prod_{t=1}^{T}\left[\frac{1}{S_\xi}\sum_{r=1}^{S_\xi}\left(\prod_{j=1}^{J_t}\frac{1}{k^p_j}\mathcal{K}\left(\frac{\ln p^{sr}_{jt} - \ln p_{jt}}{k^p_j}\right)\right)\cdot\left(|\boldsymbol{K}_t|^{-\frac{1}{2}}\mathcal{K}\left(\boldsymbol{K}_t^{-\frac{1}{2}}\left(\boldsymbol{q}^{sr}_t - \boldsymbol{q}_t\right)\right)\right)\right]\right\} \tag{23}$$

where $k^p_j = k\sigma_\nu$, $\boldsymbol{K}_t = k\Sigma_{\eta_t}$, $\mathcal{K}(\cdot)$ is the Gaussian kernel, and the rs superscript denotes the rs simulated draw. The scaling factors, k, for each time block are presented in Table 8.

Time Block	k
Bar Hours	3
Evening Rush	3
Morning Rush	4
Weekday Day	5
Weekday Evening	5
Weekend Day	5
Weekend Evening	5

Table 8. Scaling Factors for Kernel Smoothing

References

Aizcorbe, Ana and Jeff Chen (2022), "Outlet Substitution Bias Estimates for Ride Sharing and Taxi Rides in New York City." *BEA Working Paper Series*.

Berry, Stephen T., James Levinsohn, and Ariel Pakes (1995), "Automobile prices in market equilibrium." *Econometrica*, 63, 841–890.

Ching, Andrew T. (2010a), "A Dynamic Oligopoly Structural Model for the Prescription Drug Market After Patent Expiration." *International Economic Review*, 51, 1175–1207.

Ching, Andrew T. (2010b), "Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration." *International Journal of Industrial Organization*, 28, 619–638.

Cohen, Peter, Robert Hahn, Jonathan Hall, Steven Levitt, and Robert Metcalfe (2016), "Using Big Data To Estimate Consumer Surplus: the Case of Uber." *NBER Working Paper Series*.

de Haan, Jan and Frances Krsinich (2018), "Time Dummy Hedonic and Quality-Adjusted Unit Value Indexes: Do They Really Differ?" *Review of Income and Wealth*, 64, 757–776.

Erdem, Tülin and Michael P. Keane (1996), "Decision-Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets." *Marketing Science*, 15, 1–20.

Erickson, Tim and Ariel Pakes (2011), "An Experimental Component Index for the CPI: From Annual Computer Data to Monthly Data on Other Goods." *American Economic Review*, 101, 1707–1738.

Greenlees, John S. and Robert McClelland (2011), "New Evidence on Outlet Substitution Effects in Consumer Price Index Data." *The Review of Economics and Statistics*, 93, 632–646.

Lam, Chungsang Tom and Meng Liu (2017), "Measuring Consumer Surplus in the On-Demand Economy: The Case of Ride Sharing." *WP*.

McFadden, Daniel (1989), "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration." *Econometrica*, 57, 995–1026.

Nevo, Aviv (1993), "New Products, Quality Changes, and Welfare Measures Computed from Estimated Demand Systems." *The Review of Economic Statistics*, 85, 266–275.

Osborne, Matthew (2011), "Consumer learning, switching costs, and heterogeneity: A structural examination." *Quantitative Marketing and Economics*, 9, 25–70.

Roberts, John H. and Glen L. Urban (1988), "Modeling Multiattribute Utility, Risk, and Belief Dynamics for New Consumer Durable Brand Choice." *Management Science*, 34, 167–185.

Silver, Mick (2009), "An Index Number Formula Problem: The Aggregation of Broadly Comparable Items." *IMF Working Paper*.