Taxi Fare Mechanism: New York City Yellow Cab Trips

Hyoung Suk Shim*

January 27, 2017

Abstract

I develop a taxi fare mechanism model that explains drivers’ taxi trip route choice behavior with respect to a regulated taxi fare rule. I define the equilibrium condition under which drivers choose optimal routes, those which make a taxi fare mechanism Pareto efficient. Then I draw a theoretical prediction from the model that heterogeneous trip characteristics can cause inefficiency of the taxi fare mechanism for a trip relative to one that is Pareto efficient with the same optimal route, if the regulated fare increases monotonically with trip distance. I test the prediction empirically using 379 million New York City yellow cab trip records. In the results, under the given metered fare rule, trips exiting Manhattan display an efficient taxi fare mechanism but trips entering Manhattan show an inefficient mechanism, which means drivers are likely to make trip longer than optimal and demand for taxi trip to Manhattan is not maximized.

1 Introduction

Expanding taxicab service access in a metropolitan area is a newly emerged policy issue for taxi regulatory agencies, especially in New York City. According to the New York City Taxi & Limousine Commission (TLC), “95% of yellow taxi pick-ups occur in Manhattan below 96th Street and at John F. Kennedy International Airport (JFK) and LaGuardia airport.”

By the yellow cab market failure, I refer to taxi access being limited to very specific areas such as Manhattan below 96th Street and at JFK and LaGuardia airports.

To begin addressing this market failure, a careful examination of its identifying causes of yellow cab service failure, 95% pick up occurrence only in Manhattan and the city airports, is necessary. The ideal standard economic model of for taxi trips for such an examination

---

*School of Business, College of Staten Island, The City University of New York, New York, email: Hyoungsuk.Shim@csi.cuny.edu

1In his 2011 State of the City address, Mayor Michael Bloomberg noted the inequitable distributions of taxi service in New York City: about “the yellow taxi GPS data that the TLC has collected shows, 95% of yellow taxi pick-ups occur in Manhattan below 96th Street and at JFK and LaGuardia airports.” And he proposed the Five Borough taxi plan, which included the launch of the green cab taxi, to alleviate the problem of accessibility. See the background on the Boro Taxi Program at the TLC’s website.
needs to take must take the following stylized facts into account: i) market entry and pricing are regulated, ii) drivers cannot deny trip requests from passengers at a certain area, iii) passengers have less information about traffic condition and trip routes than drivers, iv) passengers’ preference over different trip location may not be homogenous, and v) drivers care about the possibility of finding subsequent passengers when a trip ends.

I develop a taxi fare mechanism model that explains drivers’ trip route choice behavior with respect to a fare rule given by the taxi regulatory agency. Given in the model description is that, for any taxi trip request, the driver can choose the trip route, which ultimately determines the duration and distance of the trip. The passenger then pays the taxi fare under the rule devised by the taxi regulatory agency. The driver’s route choice partially determines the taxi fare but it is private information, where the passengers and the regulatory agency have incomplete information about how individual drivers choose a route. The regulated fare rule, designed to achieve the regulatory agency’s objective, is expected to lead the driver to choose the optimal route that minimizes trip time and distance. From the taxi fare mechanism model, I define the equilibrium condition under which drivers choose optimal routes, those which make a taxi fare mechanism Pareto efficient, and show that Pareto efficiency of a taxi fare mechanism can be examined by looking at how much drivers’ desired route distances deviate from the optimal route distances for a given fare rule, which is the regulatory agency’s policy instrument.

I then show that Pareto efficiency of the taxi fare mechanism can differ by a heterogeneous trip characteristic, specifically the drivers’ differing anticipation of finding subsequent passengers depending on the trip destination. A superior taxi trip market, defined in this study as one for which drivers anticipate the highest probability of finding passengers for the subsequent trip, plays a crucial role in imposing the heterogeneous trip characteristic. The theoretical prediction is that drivers are more likely to extend trip distance longer than optimal for the trip to a STM, and this possibly shrinks the taxi trip market by reducing demand for taxi trip, if the passengers are aware of the drivers’ behavior. This inefficient taxi fare mechanism is caused by i) the monotonic increasing fare rule, and ii) the participation constraint that the drivers cannot deny trip requests from passengers.

The empirical analysis bears out in the yellow cab trip date the theoretical prediction—a monotonic increasing fare rule may yield an inefficient taxi fare mechanism for trips with the same optimal route but different direction. The metered fare rule is observed to be inefficient for trips within the city area, and entering to Manhattan from the other boroughs, where Manhattan is a superior taxi trip market (STM). On the other hand, the metered fare rule is observed to be efficient for trips exiting Manhattan to the other boroughs. Furthermore, the TLC’s fare rule for the entire taxi trip and within the city trips are both observed that
it is inefficient. This inefficient fare rule can be a cause and consequence of the inequitable yellow taxi cab failure distribution where that 95% of trips had occurred only in Manhattan below 96th Street and at the city airports in the city.

For the empirical analysis, I perform generalized method of moments (GMM) estimation for an econometric model that can examine whether a given fare rule yields a Pareto efficient taxi fare mechanism for a specific group of taxi trips that share the same trip characteristic. New York City yellow cab trip data. TLC collects yellow cab trip records that contains geographic coordinates of trip origin and destination, itemized taxi fares, date and time at trip begin and end, total trip time and distance. The data has 379 million yellow cab trip records collected from January 2008 to November 2010.

Mechanism design is the ideal approach to analyze the taxi trip service market. It demonstrates strategic interactions between a principal and agents for implementing a social choice function. The principal devises rules for the agents’ participation and monetary compensation to implement the social choice function that is an objective of an aggregate entity in a market such as consumer demand or welfare. In the taxi trip market, the taxi regulatory agency is the principal, drivers are the agents, and the passengers’ taxi trip demand is the social choice function. The regulatory agency compels drivers to comply passengers’ trip requests as a participation constraint, and regulates fares as a price control. For each taxi trip, drivers choose the trip routes and receive the trip fare as a monetary compensation from the passengers. As Holmström and Myerson (1983) stated, the taxi fare mechanism can be defined as an allocation that consists of the fare rule and the demand for the taxi trip.2

Using mechanism design, we can examine the Pareto efficiency of the rules that govern the agents with private information, and in the same way, Pareto efficiency of the taxi fare mechanism can be examined. We can further examine whether the efficiency of the taxi fare mechanism holds for different trips with different probabilities of finding subsequent passengers at the destination. A Pareto efficient mechanism is one that implements the social choice function by leading the agents to reveal their types truthfully, where the agents’ type is hidden information to the principal. Maskin and Sjöström (2002) defines the implementation as “the set of equilibrium outcomes equals the set of optimal outcomes identified by the social choice rule”. In the taxi trip market, I posit that demand for a taxi trip is implemented by

---

2The exact quotation of Holmström and Myerson (1983) is “a mechanism is any specification of how the economic decisions are determined as an individual’s information”. The taxi fare mechanism is a specification of how taxi trip demand is determined by drivers’ route choice, which is private information, for a given taxi operation rule. Since a taxi fare is primarily determined by trip time and distance, drivers may have an incentive to adjust time and distance of the trip by choosing a different route than the optimal, which the regulatory agency and the passengers expect to be chosen. The drivers’ route choice is hidden information to the regulatory agency and the passengers until the trip terminates its destination.
the taxi fare mechanism, if the fare rule leads the drivers to choose an optimal route that maximizes the demand by minimizing time and distance of the trip.

This is the first analysis of the economic effects of taxi fare rules and the associated efficiency of the market allocation for individual trips. A number of studies analyze the NYC yellow cab but their unit of observations are either cabdrivers’ shifts or drivers’ profile records, which are an aggregate level of individual trip that is this study’s unit of observations. The most recent closely related studies, by Jackson and Schneider (2011), and Schneider (2010), examine the New York City taxi drivers’ moral hazard which motivates drivers to engage in risky driving and criminal activities. The unit of observation of these studies, however, is the individual driver’s legal record, not individual taxi trips. Behavioral economics studies of the New York City taxi drivers’ labor supply are done by Camerer et al (1997), Farber (2008, 2015), and Crawford and Meng (2011), which argue that cabdrivers’ daily labor supply decision violates the neoclassical model’s prediction that supply will monotonically increase with revenue. The unit of observation here is drivers’ shifts in terms of hours.

The current investigation focuses on assessing the efficiency of taxi fare rules by demonstrating the taxi fare mechanism model, based mainly on Holmström and Myerson (1983), instead of deriving an optimal fare rule and regulation that yields a Pareto efficient allocation. Most of mechanism design application studies are devoted to finding Pareto efficient mechanism for a particular market design. Abdulkadiroğlu and Sönmez (2003), the seminal paper on school choice, and its subsequent research propose the best student-school matching mechanism based on the Gale-Shapley algorithm for different school choice rules. Recent studies on mechanism design applications focus on matching market and propose efficient matching mechanisms. Roth et al (2007) proposes an efficient kidney exchange rule between incompatible patient-donor pairs. Hitsch et al (2010) examines online dating market and finds the empirical evidence that the successful couple matchings are mostly achieved by the Gale-Shaply algorithm. Sönmez and Switzer (2013) proposes an efficient mechanism for cadets’ military branch choices. However, deriving an optimal mechanism for the taxi trip market is quite challenging, because the taxi regulatory agency’s objective is not only maximizing the market participants’ utilities, but comprises ancillary issues such as preventing criminal activities and accidents, controlling traffic congestion, and balancing taxi service with other modes of transportation.

3Economic analysis on taxi trip market as an industry are done by Douglas (1972), De Vany (1996), Beesley and Glaister (1983), Cairns and Liston-Heyes (1996), Arnott (1996), and Flores-Guri (2003). These studies analyze the taxi trip market with the fare as a unit price of the trip, and discuss whether the fare has to be regulated because of its monopoly pricing due to medallion licensing as an entry control. But these studies treat all drivers, passengers, and trips as homogenous, so none of the important factors such as drivers’ route choice and its incomplete information, heterogeneous trip characteristic, in the taxi fare allocation can be taken into account.
In the next Section, I introduce the taxi fare mechanism model, which incorporates with drivers’ route choice with respect to a given fare rule and the demand for taxi trip. In section 3, the taxi fare mechanism is analyzed to infer the theoretical prediction about its Pareto efficiency of the taxi fare mechanism is discussed. Section 4 describes empirical analysis issues such as econometric models, identifications, and large scale GMM estimations using high performance computing techniques. The GMM estimation results are presented and interpreted in section 5.

2 The Taxi Fare Mechanism Model

I propose a taxi fare mechanism model and establish the taxi fare allocation, consists of the fare rule and the consumer’s demand for taxi trip. Then I define the expected utilities of the drivers and of the taxi regulatory agency in order to discuss their optimization behaviors. Finally, I review the taxi fare rule and taxi trip demand model, proposed by Peters et al (2011) to establish the validity of its use in the taxi fare mechanism.

2.1 Taxi Fare Mechanism

Consider an origin-destination (OD) taxi trip, where passengers desire to go from a specific location to another specific location in a metropolitan area. Suppose that there is a number of drivers $I$ operating taxicabs in that area. A taxi regulatory agency governs the taxicab operation, it establishes market participation rules such as the taxi medallion system in New York City, and regulates taxi trip fares. The $I$ taxi drivers are thus either medallion owners, or agents who lease the medallion from the medallion owners. Denote driver $i$’s type as $\theta_i$, and we define a type vector $\theta = (\theta_1, \theta_2, \ldots, \theta_I)$ for all driver $i$, $(i = 1, \ldots, I)$. We assume that driver $i$’s hidden type $\theta_i$ is the distance of a trip route that driver $i$ chooses for a particular OD trip.

In other words, the trip distance $\theta_i$ of the OD trip is a single input for driver $i$’s utility maximization because the distance $\theta_i$ can vary according to the route chosen, for any given OD trip. Driver $i$ can increase or decrease $\theta_i$ by altering her route choice, in order to maximize her utility for each OD trip.

Now I construct a probability model for the distance $\theta_i$. Assume that there are finitely many OD trips in a metropolitan area. The distance $\theta_i$ is drawn from a set of all possible routes’ distance, denoted $\Theta_i$, for each given OD trip, The distance $\theta_i$ is therefore a random variable with a probability distribution function (p.d.f.) $\phi(\theta_i)$, and the set $\Theta_i$, a sample space of $\theta_i$, belongs to a collection of sets of trip distance, denoted $\mathcal{R}$, for all possible OD trips.
in the metropolitan area. This collection $\mathcal{R}$ is a $\sigma$-algebra defined on the road network of the area.\footnote{Following the same logic, the distance vector $\theta = (\theta_1, \ldots, \theta_I)$ belongs to the set $\Theta = \Theta_1 \times \Theta_2 \times \ldots \times \Theta_I$, and the set $\Theta$ contains the distances of all possible routes for a given OD trip. The distance vector $\theta$ is a vector of elements drawn from the $\sigma$-algebra $\mathcal{R}$, which is a mathematical representation of the road network in the metropolitan area, and therefore, the distance vector $\theta$ is a vector of well-defined random variables.} The p.d.f. $\phi(\theta_i)$, sample space $\Theta_i$, and the $\sigma$-algebra $\mathcal{R}$ construct a probability space $(\phi(\cdot), \Theta_i, \mathcal{R})$ for the random variable $\theta_i$.\footnote{Note that the pdf $\phi(\theta_i)$ is a probability measure of driver $i$’s anticipation to find subsequent passengers for a given OD trip. In other words, driver $i$ assigns probabilities for all possible routes along the road network $\mathcal{R}$ with the route distance $\theta_i$, so the probability that driver $i$ chooses a route with distance $\theta_i$ is $\phi(\theta_i)$.}

[Figure 1 about here.]

Figure 1 is a NYC road network map showing the probability space $(\phi(\cdot), \Theta_i, \mathcal{R})$ for $\theta_i$ applied in a real world case. The thick red line represents the shortest route for the trip from Times Square to JFK airport. This shortest route is one out of finitely many alternative routes, with the distance of each route being a given constant. The distance $\theta_i$ for an OD trip is therefore randomly drawn from a set of all possible route distances $\Theta_i$, on the road network $\mathcal{R}$. In Figure 1, $\mathcal{R}$ is the NYC road network, $\Theta_i$ is a set of all possible distances for the trip from Times Square to JFK airport, and the distance $\theta_i \in \Theta_i$ is the shortest possible for the trip.

Next I define an allocation of the taxi trip fare, denoted $x(\theta) = (y(\theta), t(\theta))$, that consists of the demand for a taxi trip and the associated fare amount.\footnote{Fudenberg and Tirole (1991), FT hereafter, calls $x(\theta)$ as an allocation, whereas Mas-Colell, Whinston and Green (1995), MWG hereafter, calls it an outcome function. The key difference between the two is that the outcome function has a social choice function as a single input, and the mechanism implements the social choice function if there is a strategy profile that yields the outcome function equal to the social choice function. In FT, on the other hand, the allocation has two inputs, the social choice function and the monetary transfer. The monetary transfer is a compensation for agents’ action from their principal, and it plays a key role to define the implementability condition for the social choice function. FT set up is suitable for the taxi fare mechanism because taxi fare rules can be modeled as the monetary transfer, and this allows the model to examine whether a given fare rule implements taxi trip demand, which is the social choice function in the taxi fare allocation.} Let $y : \Theta \mapsto \mathbb{N}_0$ be the taxi trip demand function that maps from the distance to the non-negative integer, including zero. Let $t : \Theta \mapsto \mathbb{R}_+$ be a monetary transfer from passenger to driver, so $t(\theta)$ is the fare for a taxi trip with respect to the distance $\theta_i$. Assume that the taxi fare $t(\theta)$ is strictly increasing in trip distance $\theta_i$.\footnote{Details of the trip fare $t(\theta)$ and the taxi trip demand $y(\theta)$ are discussed in section 2.2 and 2.3 respectively.} The implication of the allocation $x(\theta) = (y(\theta), t(\theta))$ is that for an OD trip, driver $i$ delivers passengers along a route with distance $\theta_i$, and the passengers pay $t(\theta_i)$ as trip fare. The fare rule $t(\theta)$ is determined by the taxi regulatory agency, and I assume that the agency designs the fare rule in order to implement the demand for the OD trip $y(\theta)$, as Fudenberg and Tirole (1991) stated that “the object of the mechanism built by
The principal is to determine an allocation."

The taxi fare allocation allows us to examine drivers’ route choice behavior with respect to a given fare rule and demand for taxi trip. Consider a Von-Neumann-Morgenstein utility $u_i(x, \theta)$ for player $i$, $i = 0, \ldots, I$. Let player 0 be a taxi regulatory agency, the principal, and the rest of the players are drivers whose licenses are issued by the regulatory agency. The drivers’ utility $u_i(x, \theta)$, $i = 1, \ldots, I$ is strictly increasing in $t(\theta)$ and the regulatory agency’s utility $u_0(x, \theta)$ is decreasing in $t(\theta)$ because $t(\theta)$ is the price for a taxi trip and it decreases the demand $y(\theta)$. In this set up, the regulatory agency chooses a fare rule to maximize its utility, denoted $t^*(\theta)$. This optimal fare rule is expected to maximize taxi trip demand by minimizing monetary transfer, the taxi fare, from passengers to drivers, and providing enough incentive to driver $i$ to choose the optimal route with distance $\theta_i$. The driver $i$’s object is then to maximize the monetary compensation $t^*(\hat{\theta}_i)$, and it can be done by the choice of trip route with distance $\hat{\theta}_i$, since the taxi trip demand $y(\theta)$ is endogenous to the drivers.

With the probability space for $\theta_i$ and the utility functions $u_i(x, \theta)$, we can define drivers’ expected utility:

$$U_i(\theta_i) = E_{\theta_{-i}}[u_i(x(\theta_i, \theta_{-i}), \theta_i)|\theta_i], \text{ for } i = 1, \ldots, I$$

and the regulatory agency’s expected utility is then:

$$U_0(\theta) = E_{\theta}[u_0(x(\theta), \theta)],$$

for given type-contingent allocation $\{x(\theta)\}_{\theta \in \Theta}$. With the expected utilities of the drivers and the regulatory agency, the taxi fare allocation takes a Bayesian game form, where driver $i$ and the regulatory agency choose $\hat{\theta}_i$ and $t^*(\theta)$ respectively to maximize their expected utilities (2.1) and (2.2).

Another important component in mechanism design as a Bayesian game is a message of player $i$, denoted $\mu_i(\theta_i) \in M_i$. The message is player $i$’s announcement to the principal about his type, and the message space $M_i$, a collection of subsets of the announcement is defined by a given mechanism. The message for a type vector $\theta$, denoted $\mu = (\mu_1, \ldots, \mu_I)$, is essentially the same as a strategy profile in a Bayesian game, where agents exercise strategies by announcing their types that maximize their expected utilities. We consider a vector of trip distances, denoted $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_I)$, that is the vector of trip distances that yields a
BNE:

\[ \mu^*(\theta) = (\mu^*_1(\theta_1), \ldots, \mu^*_I(\theta_I)) = \hat{\theta}, \]

where \( \mu^*(\cdot) \) is a BNE strategy profile for a given allocation. So, for an OD trip, \( \theta_i \) implies the shortest route distance and \( \hat{\theta}_i \) is a route distance that driver \( i \) chooses.

### 2.2 The Taxi Regulatory Agency’s Problem

In the taxi fare mechanism, the regulatory agency, as a principal, is assumed to choose trip fare rules to maximize its expected utility subject to the drivers’ participation (individual rationality, IR) and incentive compatibility (IC) constraints. The regulatory agency’s optimization problem then can be summarized as:

\[
\max_{t(\theta)} U_0 = E_0 [u_0(y(\theta), t(\theta), \theta)]
\]

s.t. (IC) \( u_1(y(\theta), t(\theta), \theta) \geq u_1(y(\hat{\theta}), t(\hat{\theta}), \theta), \) for all \( (\theta, \hat{\theta}) \)

(IR) \( u_1(y(\theta), t(\theta), \theta) \geq u, \) for all \( \theta. \)

The regulatory agency is assumed to employ medallion systems and to comply with passengers’ trip requests. This assumption eliminates the IR constraint from the optimization problem. Note that the regulatory agency identifies drivers’ utilities as homogeneous and deterministic, so the dominant strategy IC is employed instead of the Bayesian IC with (2.1).

The main mechanism design problem is finding the optimal allocation that yields truthful implementation of a social choice function in the mechanism, which is strategy-proof or incentive compatible. In this paper, however, I do not derive the optimal allocation for the taxi fare mechanism. Instead, I focus on evaluating Pareto efficiency of a chosen taxi fare mechanism. Hereafter, I let \( t^*(\theta) \) be a proposed fare rule that the taxi regulatory agency expects to be a solution for its optimization problem.

### 2.3 Taxi Trip Demand

Taxi trip demand is the social choice function in the taxi fare mechanism, and it is modeled as \( y : \Theta \mapsto N_0 \), a count demand function of trip distance for an OD trip. The idea of the taxi trip demand is that passengers for an OD trip decide the number of taxi trips with respect to the trip distance \( \theta_i \). The distance \( \theta_i \) is distinct from trip distances of other trip modes such
bus, train, subway, and etc. The taxi demand \( y(\theta_i) \) has thus trip distance as a single input, and the rest of trip demand attributes are assumed to be considered in the passengers’ trip mode choice, so are given in the taxi demand itself. Since \( y(\theta_i) \) is decreasing in the distance \( \theta_i \), moreover, the taxi regulatory agency expects to maximize the demand by the choice of fare rule \( t^*(\theta_i) \), which is expected to lead to driver \( i \) choosing the shortest route for the OD trip.

The count model for travel demand is proposed by Peters et al (2011). The baseline assumption of the model is that individuals choose a number of trips to maximize their utility, subject to their budget and travel time forecast. Utility is a function of consumption and leisure, and the consumption and leisure time are discounted by cost and time of travel, respectively. The set up is thus the disutility of travel in which individuals can achieve maximum utility by minimizing travel time and cost, in order to spend (allocate) money and time for travel to consumption and leisure.

For an OD trip with taxis, the count model for travel demand can be treated as a function of distance. The mode specific travel attributes in this demand model are the unit price of travel, trip time, and distance. The unit price of taxi trip is regulated by a fare rule \( t^*(\theta) \) that is a function of trip distance, and other attributes are constant over travel mode. The trip time forecasting with the random component due to traffic conditions is made up of trip route and distance so that it is a function of trip route distance. Therefore, the count model for taxi trip demand is \( y(\theta_i) \), a composite function of trip route distance. The characteristic of \( y(\theta_i) \) is decreasing in \( \theta_i \) because taxi fare \( t^*(\theta_i) \) is increasing, and trip time forecasting \( h'(\theta_i) \) is strictly decreasing in \( \theta_i \).

The implementation of \( y(\theta_i) \) thus maximizes demand for taxi trips by the choice of fare rule \( t^*(\theta) \), and the optimal route distance \( \theta_i \) is obviously the distance that yields implementation of the demand \( y(\theta) \). Since the fare amount \( t^*(\theta_i) \) is at minimum, moreover, we can define a Pareto optimal taxi fare allocation as the shortest distance route with \( \theta_i \) for an OD trip. Further details about implementable \( y(\theta) \), Pareto optimum, and truthful implementation are discussed in the next section.

---

\(^8\)The conventional taxi travel demand model has waiting time and fare amount as its attributes but this does not seem to be an applicable model set up for the current study because i) the attributes of distance and travel time are as important as fare, ii) waiting time is a part of the entire travel time and the entire travel time is a function of the travel distance, so the demand model with fare and waiting time only a special case of taxi trip demand, and iii) absence of trip distance means the model cannot provide any implications on drivers’ route choices. See Douglas (1972), and the subsequent research on taxi demand.
3 Mechanism Design of Taxi Industry: Analysis

I demonstrate for the taxi fare mechanism model about the equilibrium condition and its Pareto efficiency, which are determined by taxi fare rules. I infer the theoretical prediction that the fare rule, established by the regulatory agency, causes different consequences in the taxi fare mechanism. This analysis considers the following points: i) Drivers have preference over the OD trips because each OD trip has a different probability of finding subsequent passengers; ii) The same fare rule, determined by trip distance, applies for all OD trips; iii) Drivers’ route choice therefore may give different from OD trips with differing probabilities of finding subsequent passengers, even if the OD trips have the same optimal route distance. The prediction asserts that the different drivers’ route choice behavior with respect to the probability of finding subsequent passengers may result in different demand for taxi trips with identical optimal route distance.

3.1 Equilibrium of the Mechanism

First, I define Bayesian Nash equilibrium (BNE) and implementability conditions of the taxi fare mechanism. Let $\hat{\theta}_i$ be a driver $i$’s desired route distance, and $\theta_i$ is its optimal route distance for an OD trip. Recall that $t^*(\theta)$ is the fare rule given by the taxi regulatory agency, and the associated taxi fare allocation is $x^*(\theta_i) = (y(\theta_i), t^*(\theta_i))$. The demand for taxi trip $y(\theta_i)$ is implementable, if the given fare rule $t^*(\theta)$ satisfies the Bayesian incentive compatibility (BIC hereafter) constraint,

$$E_{\theta_{-i}}[u(x(\theta_i, \theta_{-i}), \theta_i)] \geq E_{\theta_{-i}}[u(x(\hat{\theta}_i, \theta_{-i}), \theta_i)],$$  (3.1)

for all $(\theta_i, \hat{\theta}_i) \in [\theta, \bar{\theta}] \times [\theta, \bar{\theta}]$, for each driver $i = 1, \ldots, I$, and for each $\theta_i$, $\hat{\theta}_i$, and $\theta_{-i}$.

In game theoretic interpretation, the truthtelling $\{\hat{\theta}_i = \theta_i\}$ is a BNE of the taxi fare mechanism with $t^*(\theta)$. Recall that the message $\mu_i(\theta_i)$ is the distance $\hat{\theta}_i$, and the vector of messages $\mu(\theta)$ is the distance vector $\hat{\theta}$. The BNE strategy profile $\mu^*(\theta) = (\mu^*_1(\theta_1), \ldots, \mu^*_I(\theta_I))$ is the vector of optimal distances $\theta = (\theta_1, \ldots, \theta_I)$, if the fare rule $t^*(\theta)$ satisfies the BIC (3.1). The BNE $\{\hat{\theta}_i = \theta_i\}$ implies that the taxi fare rule $t^*(\theta)$ gives enough incentive to driver $i$ choosing the optimal route with distance $\theta_i$, for the given OD trip. As defined in section 2.3, the taxi trip demand $y(\theta)$ is maximized at the trip distance $\theta_i$ and it is the implementation of $y(\theta)$. The taxi fare rule $t^*(\theta)$ that yields the truthtelling $\{\hat{\theta}_i = \theta_i\}$ is therefore truthfully implementing the taxi trip demand $y(\theta)$.

The truthful implementation of the taxi demand $y(\theta)$ through the allocation $x^*(\theta)$ is thus
defined. The question may then be asked whether there exists a direct revelation mechanism that is a BNE, which implements \( y(\theta) \) truthfully. The direct revelation mechanism implies, in the taxi fare mechanism, that the taxi regulatory agency regulates drivers’ route choice so that the drivers always choose the trip route that is given by the agency. In other words, the agency orders driver \( i \) to drive an exact trip route with distance \( \theta_i \), every time the driver \( i \) has passenger request an OD trip (???). This direct revelation mechanism is feasible, if the optimal route distance is fixed and time invariant. But in reality, the optimal route for an OD trip is not the same at all times because traffic conditions are not fixed and time invariant. For this reason, I do not consider a direct revelation mechanism in the taxi fare allocation.

Next, I define Pareto efficiency of the taxi fare mechanism. The taxi fare allocation or mechanism \( \{x^*(\theta)\}_{\theta \in \Theta} \) with \( t^*(\theta) \) is ex-ante Pareto efficient, if the fare rule \( t^*(\theta) \) satisfies the BIC (3.1) that yields the truth-telling \( \{\hat{\theta}_i = \theta_i\} \), for each \( i = 1, \ldots, I \). The reasoning for this Pareto efficiency is as follows: Since the taxi fare \( t(\theta) \) is strictly increasing in \( \theta \), the taxi fare amount \( t^*(\theta) \) is minimized at \( \theta_i \). Implementation of the taxi trip demand \( y(\theta) \) means that the fare rule \( t^*(\theta_i) \) maximizes \( y(\theta_i) \), for each driver \( i = 1, \ldots, I \). The taxi fare rule \( t^*(\theta_i) \) is thus a minimum monetary transfer from passengers to driver \( i \), and it maximizes the demand \( y(\theta_i) \). Moreover, the taxi fare allocation \( \{x^*(\theta)\}_{\theta \in \Theta} \) with \( t^*(\theta) \) is ex-ante efficient because driver \( i \)'s route choice decision is unknown for the regulatory agency, passengers, and the other drivers until the trip terminates at its destination.9

### 3.2 The Superior Taxi Trip Market

A superior taxi trip market (STM) is a physical location where a driver can find passengers and commence a taxi trip, at least risk. By assuming that, for an OD trip, the origin has the highest probability of finding subsequent passengers, the probability decreases as the distance from the origin increases. Drivers foresee smaller probabilities of finding subsequent passengers at the destination, as the trip length the origin increases. An OD trip from the STM can be represented by a spatial distribution of potential passengers, and the distribution is defined on distance from the origin. Let \( f(\cdot) \) be a probability (spatial) distribution function of finding subsequent passengers that is defined on \( \theta_i \). With \( f(\theta_i) \), we can provide a mathematical definition of the STM,

---

9Holmström and Myerson (1983)’s definition covers interim and ex-post efficiency allocation. The term ex-ante is more appropriate for the taxi market because a taxi trip is commenced (???) by a passenger who chooses to take taxi, out of the other travel modes such as driving own car and mass transit. The interim and ex-post efficiencies are infeasible, or ineffective for the passengers’ travel mode decision, and ex-ante efficiency argument is more appropriate for the taxi market.
**Definition 3.2.1** (Superior Taxi Trip Market). *An origin (destination) of a taxi trip is called a “Superior Taxi Trip Market” if the probability distribution function \( f(\theta_i) \) is monotonic decreasing (increasing) in \( \theta_i \), for each destination (origin) and for each driver \( i = 1, \ldots, I \).*

A real world example is presented in Figure 2. These are a bivariate histogram and a scatter plot of geographic coordinates of NYC yellow cab trip origins in Manhattan below 59th street. The highest peak, in panel (a) is at West 42nd Street and Broadway, Times Square, and the highest of bars representing frequency of trip origins gradually decrease as for sites further away from Times Square in any direction. Time Square is thus a STM, as origin or destination relative to any other locations in Manhattan below 59th street.\(^{10}\)

### 3.3 Certainty Equivalent Distance

Consider an OD trip with \( t^*(\theta) \), the fare rule given by the taxi regulatory agency. Recall that the fare rule \( t^*(\theta) \) is strictly increasing and applies for all OD trips in the metropolitan area. Let \( \hat{\theta}_i \) and \( \theta_i \) be the trip distances of driver \( i \)'s desired route and the optimal route respectively. Assume the origin is a STM, and let \( f(\theta_i) \) and \( g(\theta_i) \) be the spatial distributions for the trip from the origin and the return trip from the destination to the origin respectively. By definition 3.4.1, \( f(\theta_i) \) and \( g(\theta_i) \) are monotonic increasing and monotonic decreasing in \( \theta_i \) respectively.

As shown in panel (a) of Figure 3, the greater is the length from the STM origin of a trip with \( f(\theta_i) \), smaller is the chance driver \( i \) finds subsequent passengers at the destination of a trip. In panel (b), on the other hand, the greater the length of a trip with \( g(\theta_i) \), the greater is the chance driver \( i \) finds subsequent passengers at the destination of a trip.

The spatial distributions \( f(\theta_i) \) and \( g(\theta_i) \) introduce driver \( i \)'s route choice behavior under uncertainty to the taxi fare mechanism. In other words, driver \( i \) chooses a trip route that maximizes the expected fare, rather than the given fare for a present trip. Let \( F(\theta) \) and \( G(\theta) \) be cumulative distribution functions of \( f(\theta_i) \) and \( g(\theta_i) \) respectively. Define expected fares \( \int t(\theta)dF(\theta) \) and \( \int t(\theta)dG(\theta) \), for the trips with \( f(\theta_i) \) and \( g(\theta_i) \) respectively. Assume now that driver \( i \)'s payoff for having the OD trip with distance \( \theta_i \) is the expected fare \( \int t(\theta)dF(\theta) \),

\(^{10}\)The data in Figure 1 are a random sample of 300,000 observations, drawn from New York City Taxicab Limousine Commission’s GPS tracking data during 2008 to 2010. Note that I use the standard random sampling without any stratification for the entire region of NYC yellow cabs, and draw the histogram and the scatter plot only for midtown and downtown Manhattan from the 300,000 random sample.
instead of the given fare $t^*(\theta_i)$. In the same way, the expected payoff is $\int t(\theta)dG(\theta)$ for another trip with $g(\theta_i)$ and $\theta_i$. This implies that driver $i$ chooses a trip route with the distance $\hat{\theta}_i$ to maximize her utility, but the realization of the utility is not a given fare $t^*(\hat{\theta}_i)$ but the expected fare that consists of the given fare $t^*(\hat{\theta}_i)$ and the risk of having a vacant taxi for a period.

The desired route distance $\hat{\theta}_i$, which maximizes the driver $i$’s utility (2.1), is hidden information to all but the driver $i$. Instead of using $\hat{\theta}_i$ as it is, here I redefine it as the certainty equivalent distance that yields the same utility level as the optimal distance $\theta_i$. For simplicity, let $u_i(\theta_i)$ be driver $i$’s utility (2.1) with taxi demand $y(\theta_i)$ and the given fare $t^*(\theta_i)$. Let $B(X)$ be a space of the utility function, which is bounded. The following lemma asserts that there exists a unique $\hat{\theta}_i$ for a given $t^*(\theta_i)$ and the space distributions $f(\cdot), g(\cdot)$,

Lemma 3.3.1 (Existence of the trip distance $\hat{\theta}_i$). If the fare system $t(\theta)$ is defined on the space $B(X)$ and increasing in trip distance $\theta$, then:

(a) there exists a unique $\hat{\theta}_i$ such that

$$u(\hat{\theta}_i) = \int t(\theta_i)dF(\theta_i),$$

for all driver $i$ with the probability distribution function $f(\theta_i)$.

(b) there exists a unique $\hat{\theta}_j$ such that

$$u(\hat{\theta}_j) = \int t(\theta_j)dG(\theta_j),$$

for all driver $j$ with the probability distribution function $g(\theta_j)$.

Lemma 3.2.1 provides a necessary condition for the following theorem that ties the comparison between $\hat{\theta}_i$ and $\theta_i$ to the BIC (3.1). In other word, we can examine whether the fare rule $t^*(\theta)$ satisfies the BIC (3.1) and the truthful implementability condition $\{\theta_i = \hat{\theta}_i\}$ for the taxi trip demand $y(\theta)$, which is the ex-ante efficient allocation,

Theorem 3.3.1 (Truthful Implementation of Taxi Trip Demand). The taxi trip demand function $y(\theta)$, and the associated taxi trip market allocation $x^*(\theta) = (y(\theta), t^*(\theta))$ is truthfully implementable, if the fare system $t^*(\theta)$ yields

(a) for each driver $i$ with $f(\cdot)$,

$$\hat{\theta}_i \leq \theta_i \text{ such that } u(\hat{\theta}_i) = \int t(\theta_i)dF.$$
(b) for each driver $j$ with $g(\cdot)$,

$$\hat{\theta}_j \leq \theta_j \text{ such that } u(\hat{\theta}_j) = \int t(\theta) dG.$$  

Figure 4 is a simple graphical representation of theorem 3.3.1. As shown in panel (a), the expected fare given by the optimal route distance $\theta_i$ is greater than driver $i$’s utility with $\theta_i$, so the distance $\hat{\theta}_i$, which gives the same utility level of the expected fare with $\theta_i$, is less than $\theta_i$. In this case, the fare rule $t^*(\theta)$ gives enough incentive for driver $i$ to choose the optimal route $\theta_i$. The opposite case is shown in panel (b), where the expected fare given by the optimal route distance $\theta_i$ is less than driver $i$’s utility with $\theta_i$, so the distance $\hat{\theta}_i$, which gives the same utility level of the expected fare with $\theta_i$, is greater than $\theta_i$. This is the case where the fare rule $t^*(\theta)$ fails to lead driver $i$ to choose the optimal route $\theta_i$.

[Figure 4 about here.]

### 3.4 Analysis of Taxi Fare Mechanism

My analysis of the taxi fare mechanism focuses on driver $i$’s route choice with respect to a fare rule, given by the taxi regulatory agency. For an OD trip, driver $i$ chooses a trip route with distance $\hat{\theta}_i$ which yields him taxi fare $t^*(\hat{\theta}_i)$ from the passengers who determine the taxi trip demand $y(\hat{\theta}_i)$. The fare rule $t^*(\theta)$ is designed to lead driver $i$ to choose the optimal route with distance $\theta_i$ that simultaneously minimizes $t^*(\theta)$ and maximizes $y(\theta)$. The truthful implementation of $y(\theta)$ through the allocation $x^*(\theta) = (y(\theta), t^*(\theta))$ can thus be achieved by \(\{\theta_i = \hat{\theta}_i\}\).

The primary assertion of this analysis is that the fare rule $t^*(\theta)$, and monotonic increasing in $\theta_i$, cannot satisfy BIC for all OD trips, given the trips have different spatial distributions of finding subsequent passengers. To illustrate, I consider a superior taxi trip market. Suppose that the fare rule $t^*(\theta)$ yields \(\{\hat{\theta}_i = \theta_i\}\) that satisfies the theorem 3.3.1 for a trip with $f(\theta_i)$, the trip which originates from the STM. It is immediate that the unique $\hat{\theta}_i$, the trip with $g(\theta_i)$, the return trip to the STM, cannot satisfy theorem 3.3.1 because $\hat{\theta}_i$ is greater than $\hat{\theta}_i$ for all $\theta_i$. This counterexample shows the fare rule $t^*(\theta)$ does not work for OD trips with the different spatial distributions. This assertion is summarized in the following Proposition,

**Proposition 3.4.1.** If the cumulative probability distribution $F(\cdot)$ the first order stochastic dominates the cumulative distribution function $G(\cdot)$, then a unique trip distance $\hat{\theta}_i$ for the trip with $f(\theta_i)$ is less than a unique trip distance $\theta_i$ for the trip with $g(\theta_i)$, for a $t^*(\theta)$ and $\theta_i \in [\theta, \bar{\theta}]$. 

This proposition shows that the desired route distances of driver \( i \) can differ by the spatial distribution of finding subsequent passengers at the destination. For round trip, the distance \( \hat{\theta}_i \), driver \( i \)'s desired route distance for the trip originates at the STM, is greater than the distance \( \tilde{\theta}_i \), the driver \( i \)'s desired route distance for the trip returning to the STM. Although both trips have the same fare rule \( t^*(\theta) \) and the optimal route distance \( \theta_i \), the consequence of the truthful implementation of the taxi trip demand by the fare rule \( t^*(\theta) \) may be different.

There are two possible scenarios results from Proposition 3.4.1 and theorem 3.3.1 can be considered. First, if both \( \hat{\theta}_i \) and \( \tilde{\theta}_i \) are lower than the optimal route distance \( \theta_i \), then the fare rule \( t^*(\theta) \) truthfully implements the taxi demand at \( y(\theta_i) \). Second, if \( \hat{\theta}_i > \theta_i > \tilde{\theta}_i \), then the fare rule \( t^*(\theta) \) fails to implement the demand truthfully at \( y(\theta_i) \) for the trip to the STM. In this case, driver \( i \) has an incentive to choose a trip of length longer than optimal, resulting in inefficiency of the taxi fare mechanism.

In conclusion, a monotonic increasing fare rule \( t^*(\theta) \) may fail to yield the truthful implementation of the taxi trip demand \( y(\theta_i) \) for trips with the same optimal route distance but different probabilities of finding subsequent passengers at the destinations. The consequence can be an inefficient allocation of the taxi fare mechanism. In other words, drivers are more likely to extend trip distance longer than optimal for the trip to a STM, and this possibly shrinks the taxi trip market by reducing demand for taxi trip, if the passengers are aware of the drivers’ behavior. This inefficient taxi fare mechanism is caused by i) the monotonic increasing fare rule, and ii) the participation constraint that the drivers cannot deny trip requests from passengers.

4 The Econometric Model

In this section, I discuss issues in the empirical analysis to test the theoretical prediction that is summarized by Proposition 3.4.1. The goal of this section is to justify the validity of statistical identification of a proposed econometric model and the statistical consistency of its GMM estimations using the yellow cab trip data. First, I propose a conditional expectation of the distance difference between drivers’ desired route and the optimal route for a given fare rule as an observable model for the taxi fare mechanism. Then I introduce the NYC yellow cab trip data, the corresponding econometric model to the observable model, and variable creations from the data to use for the econometric model estimations. Details for computations of the GMM estimations are discussed in the appendix 3.
4.1 The Observable Model for Taxi Market Mechanism

In this section, I propose an empirically observable model for the taxi fare mechanism. Using statistical inferences, the observable model can be used to examine the truthful implementability condition and the Pareto efficiency of a given fare rule. Let $\delta$ be a hypothetical parameter that is the difference between $\theta_i$ and $\hat{\theta}_i$ for any OD trips with a given fare rule $t^*(\theta)$. We can now consider a population model for the difference. That is, for all driver $i = 1, \ldots, I$,

$$\theta_i - \hat{\theta}_i = \delta + u_i, \quad (4.1)$$

where $u_i$ is driver specific unobservable error term. The parameter $\delta$ offers a way to examine empirically the truthful implementability and the Pareto efficiency of the fare rule $t^*(\theta)$ and the associated taxi fare mechanism. Positive $\delta$ implies that the $\theta_i$ is statistically greater than $\hat{\theta}_i$, which satisfies the theorem 3.3.1. In this case, the taxi fare mechanism with the given fare rule $t^*(\theta)$ are therefore statistically Pareto ex-ante efficient. Negative $\delta$, on the other hand, implies that the taxi fare mechanism with $t^*(\theta)$ fails to satisfies the theorem 3.3.1, which means that it fails to implement the demand for taxi trip and therefore the mechanism is statistically Pareto ex-ante inefficient.

Statistical identification of the $\delta$ is the most crucial part at this population model stage, and a key component of this identification is how to treat the error term $u_i$. A sample mean of the difference $I^{-1} \sum_{i=1}^{I} (\theta_i - \hat{\theta}_i)$ becomes a statistically consistent estimator of $\delta$, if we simply assume that the error term $u_i$ is an independent and identically distributed random variable with mean zero and a constant variance. But the zero mean assumption for $u_i$ has neither economic implications nor statistical justifications. Furthermore, the error term $u_i$ may contain unobservable variation by optimal route distance $\theta_i$, which can vary over time due to traffic conditions. The optimal route distance $\theta_i$ for an OD trip is typically shortest route as measured on the roadways. But the shortest route may not be optimal for either passengers or drivers due to expected traffic congestion such as rush hours, or unexpected traffic jam due to accidents ahead or constructions. Thus, randomly given traffic conditions are an important factor in route choice.

In order to control for the error term $u_i$ properly, here I consider a conditional expectation of the difference $\theta_i - \hat{\theta}_i$ given the trip fare rule, and the spatiotemporal information sets. That is, for each driver $i = 1, \ldots, I$,

$$E \left[ \theta_i - \hat{\theta}_i \mid t(\hat{\theta}_i), \Omega_s, \Omega_t \right] = \delta, \quad (4.2)$$
where \( t(\hat{\theta}_i) \) is the trip fare resulting from the driver’s desired route distance \( \hat{\theta}_i \). \( \Omega_s \) and \( \Omega_t \) are spatiotemporal information sets that represent time and spatial variation respectively. Note that the information set \( \Omega_t \) can be thought as a set of all possible independent variables that vary over time but are spatially invariant, and the information set \( \Omega_s \) as the set of the independent variables that vary over space but are time invariant. The identification condition for \( \delta \) is that the unobservable error term \( u_i \) is exogenous: uncorrelated with the trip fare and the spatiotemporal information sets \( E[u_i|t(\hat{\theta}_i), \Omega_s, \Omega_t] = 0 \).

### 4.2 The Data

I use the NYC yellow cab trip data from 2008 to 2010 to estimate the hypothetical parameter \( \delta \) using the model (4.2). In March 2004, the TLC mandated installation of new electronic fare meters which allowed credit card payments for fares in all NYC yellow cabs under the program title “Taxicab Passengers Enhancement Project”.

11 Embedded in the meter is a trip record collection system that gathers the geographic coordinates, date and time of the trip origin and destination, the total trip time and distance, and the associated taxi fare, and reports the data to the TLC’s server. The data used in this study are total of 378,532,118 trips recorded from January 2008 to November 2010.

I use the trip distance recorded on the fare meter as the drivers’ desired route distance \( \hat{\theta}_i \) for each trip. For the optimal route distances \( \theta_i \), I make use of the Traffic Analysis Zone (TAZ) polygon, a street block level geographic polygon that is commonly used for urban transportation studies. I use the calculated as a proxy variable the shortest polygon-to-polygon route distances along the New York City metropolitan area’s road network as a proxy variable.

12 For more detail, see the U.S. Department of Transportation’s the Federal Highway Administration website.

13 \( 3.78^2 \times 10^{16} \) shortest distances need to be calculated to use point-to-point distance, and a numerical optimization is required for each calculation. Furthermore, many of trip originate or terminate in the same street blocks (although from different building number). I treat all trip origins, or destination points in the same geographic polygon as the same trip origin, or destination, and use the calculated shortest distance as a representative distance for all the origin and destination points. This expects to reduce the computation burden without causing significant measurement error and improves efficiency of data management. See the appendix for details of variable creation.

---

11 For more detail, see the Taxicab Passenger Enhancements Project (TPEP) Archive at the TLC’s website on TPEP.

12 For more detail, see the U.S. Department of Transportation’s the Federal Highway Administration website.
4.3 The Econometric Model

In this section, I derive an econometric model of route distance difference (4.2), and uses it to identify and estimate the parameter $\delta$. Recall that the conditional expectation (4.2).

$$E \left[ \theta_i - \hat{\theta}_i \mid t(\hat{\theta}_i), \Omega_s, \Omega_t \right] = \delta,$$

for each driver $i$, $i = 1, \ldots, I$. To estimate this empirically, I consider a linear regression model for the distance difference $\theta_i - \hat{\theta}_i$ with control variables for the trip fare $t^*(\hat{\theta}_i)$, and the spatiotemporal information sets $\Omega_t$ and $\Omega_s$. That is

$$y_j = \delta + x_j \beta + z_j \gamma + u_j, \quad j = 1, \ldots, N, \quad (4.3)$$

where $y_j = \theta_j - \hat{\theta}_j$ is the difference between the optimal distance and the recorded distance in the taxi meter, $x_j$ is a vector of control variables for fare, and $z_j$ is a vector of control variables for time and date of a trip $j$, so that $z_j \in \Omega_t$. Note that the subscript $j$ indicates a taxi trip as an observation, not taxi driver $i$. I obtain $y_j$ by subtracting the recorded distance in the taxi data from the calculated OD distance that corresponds to the trip. $x_j$ consists of all itemized fares and its related variables such as payment method, indicators for special fare zone, number of passengers, vacancy time from the end of previous trip. $z_j$ consists of indicators for the hours, day of the week, and month of trip origins.

The independent variables $x_j$ and $z_j$ each of have standard measurement units such as dollars for taxi fare, trip time in minutes, vacant cruising time before each trip in minute, and hour, days, or months at trip origin and destination. But the spatial variation of taxi trip has no standard measurement unit. Therefore I use instrumental variables (IV) of the spatial information set $\Omega_s$, in order to control for endogeneity of $x_j$, and to avoid estimating its own effects on $y_j$.  

I use indicator variables for the origin and destination ZIP Codes as a set of IV $w_j$ under the notion that the origin and destination of a taxi trip are determined solely by passengers, and are therefore exogenous to the drivers’ route choice. In addition, the location of the trip origin and destination is one of the main determinants of the trip time and the vacant

---

14The itemized fares consist of the meter fare, toll, tip, MTA surcharge, and etc. The sum of these itemized fare becomes the total fare amount charge.

15A valid IV is required to be uncorrelated with the error term and correlated with endogenous independent variables, and therefore, the IV’s measurement unit does not matter as long as it satisfies these condition. The spatial variation of taxi trip can be locations of its origin and destinations, or trip route direction from a area to another area, and there is no standard measurement unit that guarantees not to cause measurement error in a model estimation. Thus, the robustness check of which an estimate of the spatial variation is consistent with the choice of the measurement unit has to be performed, if the spatial variation is considered as the independent variables. And the robustness check is expected to be quite computationally burdensome.
cruising time, which are the endogenous independent variables due to the unobservable traffic condition. The IVs $w_j$ are therefore expected to be the relatively strong instruments for the endogenous independent variables, freeing the estimation from the weak instruments problem of using large dataset, pointed out by Bound et al (1996). The other reason to use zipcode polygons, rather than the traffic analysis zone (TAZ), is that the number of observations within a TAZ polygon is quite small and becomes too small to compute a matrix inverse for $\sum_j w_jw_j'$.\textsuperscript{16}

5 Empirical Analysis: Taxi Fare Mechanism

In this section, I present the results of the GMM estimation of the econometric model (4.3) using the NYC yellow cab trip data. The primary focus is to test the theoretical prediction, drawn by Proposition 3.4.1 that the monotonically increasing fare rule has different implications for the taxi fare mechanism’s Pareto efficiency for trips with the heterogeneous trip characteristics. First, I present descriptive evidences that Manhattan can be identified as a STM in New York City. I explain identification strategy of $\delta$ for a particular trip characteristic based on taxi trip occurrences and fare rules.\textsuperscript{17} Then I interpret the $\delta$ estimates for complete set of subsamples.

5.1 Identification Strategy

I use the taxi trip origins and destinations to stratify the data by trip characteristic and fare rule. The trip characteristic is whether it originates from and/or terminates in a superior taxi trip market (STM), and thus it is determined by drivers’ anticipation of finding subsequent passengers at the destination. All licenced taxis in New York City, TLC are required to comply with the trip fare rules of the TLC. The applicable rules, taken from the 2010 TLC rule book, are:

i) $\text{\$2.50 initial fare.}$

ii) The charge for each additional unit is $\text{\$0.40 by} \quad$

$$1) \text{one-fifth of a mile, if the travel speed is 12 mph or more}$$

\textsuperscript{16}Note that the estimation of the parameters $\delta$ and $\beta$ in (4.3) requires the computation of the inverse of $\sum_j w_jw_j'$ for the IV estimation. The numerical computation of the inverse matrix might fail if some of indicator vectors in $w_j$ have too few observation of ones, out of the total number of observations because an algorithm for the matrix inversion may mistake the matrix as a singular.

\textsuperscript{17}The specific trip characteristics refers to a spatial distributions of potential passengers of a trip that has a STM as it’s origin, or destination. Note that drivers’ anticipation of finding next passengers at the destination of a given trip determines the trip characteristic.
2) 60 seconds, if the travel speed is less than 12 mph.

There are additional fare rules for specific trip locations such as

iii) $45 flat rates for JFK airport to Manhattan,

iv) Negotiated fare rates for trips outside of the city,

etc. These fare rules are represented by $t^*(\theta)$ for the New York City taxi fare mechanism.

Thus, we can estimate $\delta$ for the fare rule signified by i) and ii) by using the subsample of taxi trips for which the rule is applicable. The $\delta$ for a fixed fare rule can also be estimated by using the subsample of trips between JFK airport and Manhattan, which is governed by the rule iii). Note that the fare rule $t^*(\theta)$ in Proposition 3.4.1 is assumed to be monotonic increasing in trip distance $\theta_j$. So fare rules i) and ii), which apply to trips inside the city area, can be used to test Proposition 3.4.1 if we identify a STM in the city area.

I use Manhattan and the New York City area as STMs. Using Table 1 and the TLC’s fare rule, we can identify that Manhattan can be treated as a STM among the five New York City boroughs here the metered fare rule is applied. Table 1 presents trip frequencies and proportion for each trip location and time frame for the taxi trip data. As shown in ninth and tenth rows, Manhattan has about 85% of taxi trips’ origins and destinations, whereas Brooklyn and Queens have about 10% and 4% of them respectively. Trips within the city comprise about 99% of all trips, whereas trips outside of the city are less than 1% for every reported time frame such as rush hours and weekdays.

The other necessary condition for a STM is passenger density monotonic decreasing with increasing distance from the STM. Figure 5 presents the spatial distributions of taxi trip origins and destinations by 5-digit ZIP Code polygons. The spatial distributions for both origin, in panel (a), and destination, in panel (b), are most densely distributed in Manhattan, and they are gradually dispersed through the other boroughs and outside of the city.

In sum, we can estimate $\delta$ for a metered fare rule with these subsamples of trips: i) Within Manhattan, ii) From Manhattan, and iii) To Manhattan. Then, we can compare the $\delta$ estimates from the sample i) and ii) to examine how the metered fare rule determines Pareto efficiency of the taxi fare mechanism differently with the different trip characteristics. This comparison is based on the notion that Manhattan is a STM among the five New York City boroughs.

---

18 There are a few additional special fare zones and trips such as trips from, or to Nassau and Westchester counties, and fare shares for group rides in Manhattan during rush hours. Please see the TLC’s rule book §§ 1-70 to § 1-73.
5.2 Primary Results

The GMM estimates of the econometric model (4.3) with seven different samples from the taxi trip data are reported in Table 2. The first row of the table is $\delta$ estimates from each sample. The first column is the GMM estimate with the entire trip sample, the next three columns, from the second to the fourth, are the GMM estimates with New York City related trip samples, and the rest of three columns, from the fifth to seventh, are the GMM estimates with Manhattan related trips. Note that I control for time variation of trips using dummy variables of hours, days of the week, and months of the trips’ origin to estimate the models in Table 2. I also control for spatial variation using dummy variables of origin, and destinations’ ZIP Codes but these dummy variables are used as instruments to perform GMM estimations.

I interpret only the $\delta$ estimates’ signs and statistical significances, in order to examine whether the taxi fare mechanism is Pareto efficient. Recall that positive $\delta$ implies that a given fare rule satisfies the theorem 3.3.1, meaning the fare rule truthfully implements the demand for taxi trips and yields a Pareto efficient taxi fare mechanism. So, the positive and statistically significant $\delta$ estimates in the first row of Table 2 can be interpreted as which the given fare rule yields a Pareto efficient taxi fare mechanism.

The first column in Table 2 report the GMM estimate of (4.3) for the entire NYC yellow cab trip sample. The overidentification test statistic, reported at the second bottom row, is zero so that the set of IVs controls for the endogeneity caused by unobserved spatial variations well. The $\delta$ estimate is -167.14, which is statistically significant (the standard error is 21.30)\(^{19}\). This negative $\delta$ estimate implies that, in overall, the TLC’s fare rules fail to implement the demand for taxi trips, and therefore the taxi fare mechanism determined by the given fare rules and the demand is inefficient.

The coefficients of Nassau and Westchester county trip dummy variables, in the first column, are -658.25 and 264.75 respectively, and both are statistically significant. Note that the two counties’ trip have a special fare rule that it follows the metered fare rules inside the New York City, and doubled metered fare is charged for the portion of trips outside of the city area. These two coefficients of the special fare trips can be interpreted jointly with the

\(^{19}\)The direct interpretation of the $\delta$ estimate -167.14 is that the yellow cab drivers are likely to choose trip routes that make about 167 additional miles from its shortest routes’ distances, but this is an invalid interpretation. The unconditional mean and median of the distance difference for the entire trip sample are -0.3161 and -0.0900 respectively, and even its minimum is -49.76. The $\delta$ estimate is again a statistical measure of the conditional expectation (4.2).
\( \delta \) estimate as: i) the \( \delta \) estimate for the Nassau county trip is 97.61 \((-167.14 + 264.75)\), which is positive, so the fare rule for Nassau county trip yields a Pareto efficient taxi fare mechanism, ii) The \( \delta \) estimate for the Westchester county trip remains negative in which the fare rule incurs an inefficient taxi fare mechanism. The other independent variables that might change the \( \delta \) estimates’ signs are surcharge (61.77), credit card dummy (35.56), meter fare (1.25) and trip time (13.12), but the maximum values for the surcharge and credit card dummy are one, and the median values of meter fare and trip time are 7.7 and 9.88 respectively. So, none of them are likely to change the \( \delta \)’s sign\(^20\).

From the second to the fourth column in Table 2 are the GMM estimates for the three New York City area related trip samples. The \( \delta \) estimate for the trips within the city area under the metered fare rule is -5.15 but statistically insignificant. The \( \delta \) estimates for the trip from the city area, and for the trips to the city area, with the negotiated fare rule are 19.33 and 0.62 respectively but the \( \delta \) estimate for the trips to the city area is statistically insignificant. The overidentification test statistics for all three estimates are about zero, so its IVs well control for the endogeneity. These \( \delta \) estimates imply that the metered fare rule incurs an inefficient taxi fare mechanism. But this inefficiency of the taxi fare mechanism for trips within the city area is statistically inconclusive.\(^21\)

The last three columns are estimates of \((4.3)\) for the Manhattan related trip samples. The \( \delta \) estimates for all three trips are statistically significant and positive except for the trips To Manhattan sample. And the overidentification test statistics are all about zero. These \( \delta \) estimates imply that the metered fare rule yields a Pareto efficient taxi fare mechanism for trips within Manhattan and trips departing Manhattan\(^22\). Note that the toll charge coefficient for the trip within Manhattan should be irrelevant, since there is no toll road inside Manhattan but there are about 3\% of trips have a non-zero toll charge in that sample\(^23\).

Overall, the results indicate that the TLC’s fare rule is observed that fails to implement

\(^{20}\)The surcharge rule is i) $1.00 for trips between 4:00 p.m. and 8:00 p.m., weekday excluding legal holidays, and ii) $0.50 for trips between 8:00 p.m. and 6:00 a.m. on all days.

\(^{21}\)The \( \delta \) estimate’s sign for trips within the city area can be changed by the meter fare (15.83), surcharge (65.43), which are positive and statistically significant coefficients in the model. In the same way, the \( \delta \) estimate for trips from the city area can be changed by meter fare (-4.23) only, and the surcharge (-18.45) and credit card dummy (-12.84) are significant but its magnitudes cannot exceed the \( \delta \) estimate even with its maximum values. The \( \delta \) estimate’s sign for trips to the city area can be changed by the number of passengers (-1.03), the trips from Westchester (-2.92) and Nassau county (-4.49).

\(^{22}\)The signs of the \( \delta \) estimates for the trip within Manhattan and the trip From Manhattan can be changed by each of the meter fare coefficients -11.41 and -6.87 respectively and both are statistically significant. For the trip From Manhattan, the surcharge coefficient -41.03 is another statistically significant factor that can change its \( \delta \) estimate.

\(^{23}\)8,997,590 trips have non-zero toll charge amount in the trip within Manhattan sample with its median charge $4.57. These are quite suspicious taxi trip behavior that took toll roads that go to outside Manhattan, but I do not conduct any further examination in this paper and leave it for future research.
the taxi trip demand truthfully for the entire trip sample and the trip to Manhattan sample. Comparing $\delta$ estimates for the trip from Manhattan and the trip To Manhattan, moreover, we can see that they are consistent with Proposition 3.3.2. As shown in Table 1, Manhattan is the source of about 85% of the trip origins occurring in the New York City area, and its fare rule is the metered fare, monotonically increasing in trip distance. Manhattan is thus a STM among all the five New York City boroughs, and as Proposition 3.3.2 predicted, trips from Manhattan have a shorter desired route distance than that of trips to Manhattan. In other words, the metered fare rule does not provide enough incentive to drivers to choose the optimal route distance for the trip To Manhattan. Drivers are likely to make trip longer than optimal, and thus demand for taxi trips To Manhattan is not maximized.

[Table 3 about here.]

The negative $\delta$ for the trip To Manhattan is a cause and consequence of the taxi market failure that 95% of yellow cab pick up have occurred in Manhattan and the city’s airports. The TLC’s metered fare rule fails to lead drivers to choose optimal routes for passengers who request trips to Manhattan. If this drivers’ route choice behavior becomes known information to the passengers, they no longer demand that taxi trip as much as they had been willing to. The 95% taxi pickup in Manhattan is thus a sort of an equilibrium that has been made sequentially by repeating the process that the drivers’ route choices have decreased the demand for taxi trip To Manhattan, and the decreased demand has worsen the drivers’ anticipation to find subsequent passengers when they have trips to go outside of Manhattan.

5.3 Time and Spatial Variation of $\delta$ Estimates

Estimates of the $\delta$ and the coefficients of dummy variables for mass transit terminal trips are reported in Table 3.\textsuperscript{24} None of the dummy variables’ coefficients have the opposite sign with statistical significance that can change signs of the corresponding $\delta$ estimates. The only exception is the trip from the Port Authority Bus Terminal in the within NYC sample, reported at the second column. The coefficient estimate for the trip is 6.89 and it is statistically significant and greater than its corresponding $\delta$ estimate -5.15. This implies that the metered fare rule yields a Pareto efficient taxi fare mechanism for trips from the Port Authority Bus Terminal to any locations in the city area. Looking at the coefficient of trip from the Port Authority Bus Terminal in the From NYC sample, reported in the third

\textsuperscript{24}These are binary indicator variables for whether a trip’s origin or destination is located in a TAZ polygon containing mass transit terminal. The TAZ polygon is a street block level geographic polygon that is commonly used for urban transportation studies. Every mass transit terminal takes place in a whole TAZ polygon, so every trip within the polygon has the terminal as its origin or destination.
column, it is positive, which is consistent with its $\delta$ estimate 19.33, but not statistically significant.

A notable point in Table 3 is the $\delta$ estimates for special fare zone trips. Trips to Newark Liberty Airport from the city area, and JFK airport trips from or To Manhattan have special fare rules that the metered fare plus $15$ surcharge for the Newark airport trips, and $45$ flat rates for the JFK airport trips. The Newark airport trip coefficient of -0.45 in column three is statistically significant, but it is too small to change the sign of the corresponding $\delta$ estimate of 19.33. This implies that the Newark airport trip fare rule yields a Pareto efficient allocation taxi fare mechanism. Similarly, these coefficients for trips to and from JFK are too small to change the corresponding $\delta$ estimate of 19.33. This implies that the JFK trip fare rule, which is a flat rate, cannot change the given Pareto efficiency for both To Manhattan and From Manhattan trips.

[Table 4 about here.]

Estimates of the $\delta$ and coefficients of origin hours dummy variables are reported in Table 4. Note that the origin hour 5 a.m. is excluded from the GMM estimations to keep the matrix $\sum_N x_j x'_j$ non-singular, which is a necessary condition to compute the GMM estimate (7.1). Overall, no notable pattern, which may change the Pareto efficiency of the taxi fare mechanism for each sample, can be found. But the trip for the city area, reported at the fourth column has an interesting pattern that the coefficients for morning rush hours are negative and statistically significant, except 7 a.m., whereas the coefficients for evening rush hours are large, positive and statistically significant.

Estimates for the $\delta$ and coefficients of the origin days of the week dummy variables are reported in Table 5. The origin days of the week dummy variables are binary indicator variables of days of the week at which trip origins were occurred. Note that the origin Sunday is excluded from the GMM estimations to keep the matrix $\sum_N x_j x'_j$ non-singular for the same reason that I exclude the 5 a.m. dummy variable. There are no specific and notable patterns over the coefficients in Table 4 could be found, so it can be concluded that Pareto efficiency of the given fare rule is not be affected by variation in days of the week.

[Table 5 about here.]

---

25 These are binary indicator variables of the hours (24 hours range) at which trip trip origins were occurred. Table 4 reports coefficients for only rush hours, and the rest of them are reported in Appendix Table 1.


6 Conclusion

Mechanism design and big data analytics appear to be standing on opposite sides of the river, one side is advanced microeconomic theory and the other is cutting-edge econometrics. Trying together these far distant method is challenging, but the synergy is quite powerful to reveal unexplored economic phenomena such as taxi drivers’ route choice behavior with respect to a given fare rule, and Pareto efficiency of the fare rule.

The goal of this study is to find the causes of the taxi market failure that 95% of yellow taxi pick-ups occurred in very limited areas of New York City. Also, I seek empirical evidences from a Big Data of taxi trip records as to whether the market failure is caused by inefficiently designed fare rules. Mechanism design is a powerful to analyze agents’ hidden behavior with respect to the principal’s rule in a regulated market. By applying the mechanism design, we can reveal taxi drivers’ route choice behavior with respect to the taxi regulatory agency’s fare rule. Moreover, I show that the taxi fare mechanism can become inefficient for an opposite-direction trip to one that is Pareto efficient, if the fare rule (such as a metered fare) is monotonic increasing in trip distance.

Because of the size, exploratory or descriptive data analysis that uses a Big Data as it is are common in big data analytics, but the drivers’ route choice behavior like unobservable objects cannot be examined with the exploratory or descriptive approaches. The Big Data in this paper have 379 million yellow cab trip records, stored in 132Gbytes of text files, and the raw data are extremely burdensome to manipulate and perform statistical computations with. I perform GMM estimations of the econometric model to test the Pareto efficiency of taxi fare rules for a particular trip characteristic, and to do so, 180Gbytes of an independent variable matrix and 432Gbytes of IV matrix have to be created and used. By using high performance computing systems and techniques, I can compute the GMM estimates and the associated statistics, and show that the metered fare rule for within the city area causes inefficient taxi trips from the other boroughs to Manhattan, whereas the fare rule is efficient for trips from Manhattan to the other boroughs. This inefficient fare rule can be analyzed as a cause and consequence of the yellow cab failure.

I wish this study can encourage active researchers in economics to conduct more interdisciplinary studies within economics, and with other disciplines in science.

References


The New York City Taxi & Limousine Commission (2013): Background on the Boro Taxi program,


The New York City Taxi & Limousine Commission (2013): Taxicab Passenger Enhancements Project (TPEP) Archive,


The New York City Taxi & Limousine Commission (2010): Old Rule Book,


7 Appendix

7.1 Mathematical Proofs

Lemma 3.2.1. Suppose to the contrary that there does not exist a unique \( \hat{\theta}_i \) such that

\[
u(\hat{\theta}) = \int t(\theta) dF.
\]

By the incentive compatibility (IC) and the individual rationality (IR) condition, the utility function \( u(x(\cdot, \theta_{-i}), \cdot, \theta_{-i}) \) is bounded, so it is \( u : \Theta \mapsto [\underline{u}, u(\theta)] \). Without loss of generality, the utility function \( u \) is assumed to be continuous on \( \mathbb{R}^+ \), so the codomain of \( u \) is a subset of \( \mathbb{R}^+ \). Let \( B(X) \) be a space of the bounded function \( u \) and \( X \) be the codomain set \( [\underline{u}, u(\theta)] \subset \mathbb{R}^+ \). Since the set \( X \) is a subset of positive real numbers, \( B(X) \) is a complete metric space with the metric \( \rho(x, y) = |x - y| \).

Define an operator \( T : B(X) \mapsto B(X) \) such that

\[
Tu = \int t(\theta_i) dF.
\]

The operator \( T : B(X) \mapsto B(X) \) satisfies the following conditions:

(a) (monotonicity) For all \( \tilde{u}, u \in B(X) \) with \( \tilde{u} \geq u \),

\[
T \tilde{u} = Tu = \int t(\theta_i) dF.
\]

(b) (discounting) For any \( \alpha \geq 0 \) and for some \( \beta \in (0, 1) \),

\[
[T(u + \alpha)](\theta) = \int t(\theta_i) dF \leq (Tu)(\theta) + \beta \cdot \alpha,
\]

where \( (u + \alpha)(\theta) = u(\theta) + \alpha \).

Hence the operator \( T \) satisfies the Blackwell’s sufficient conditions for a contraction, so it is a contraction mapping.

By contraction mapping theorem, there exists a unique fixed point \( u^* \) in \( B(X) \) that is \( u^* = \int t(\theta_i) dF \) so a unique \( \hat{\theta}_i \) such that \( u(\hat{\theta}_i) = \int t(\theta_i) dF \) must exist. Therefore, the contrary cannot hold.

\[\square\]

Theorem 3.2.1. Suppose to the contrary that a taxi market allocation \( x(\theta) = (y(\theta), t(\theta)) \) satisfies the Bayesian IC with \( \hat{\theta}_i > \theta_i \) such that \( u(\hat{\theta}_i) = \int t(\theta_i) dF \), for all \( i \). By the property of \( t(\theta) \) that is increasing in its argument, \( t(\hat{\theta}_i) \geq t(\theta_i) \), for all \( \hat{\theta}_i > \theta_i \). This implies
\[ u(y(\theta_i), t(\hat{\theta}_i), \theta) > u(y(\theta_i), t(\theta_i), \theta), \text{ and hence} \]
\[ E_{-\theta}[u(y(\theta_i), t(\hat{\theta}_i), \theta, \theta_i)] > E_{-\theta}[u(y(\theta_i), t(\theta_i), \theta, \theta_i)]. \]

This is contradiction of the Bayesian IC.

\textbf{Proposition 2.3.1.} Let \( t^*(\theta) \) be a solution of the principal’s utility maximization problem for both trip with the spatial distribution \( f(\cdot) \) and \( g(\cdot) \) as:

\[
\begin{align*}
\max_{t(\theta)} & \quad E_\theta[u_0(y(\theta), t(\theta), \theta)], \\
\text{s.t.} & \quad u(y(\theta_i), t(\theta_i), \theta_i, \theta_{-i}) \geq u(y(\theta_i), t(\hat{\theta}_i), \theta_i, \theta_{-i}), \\
& \quad u(y(\theta_i), t(\theta_i), \theta_i, \theta_{-i}) \geq u.
\end{align*}
\]

Assume that \( t^*(\theta) \) is a unique maximum for both trip with \( f(\cdot) \) and \( g(\cdot) \). Then it is Pareto (ex-ante) superior than any other \( t(\theta) \), so it is an ex-ante efficient allocation.

By the definition of well-defined market, \( F(\cdot) \), the cumulative distribution function of the probability distribution function \( f(\cdot) \), first-order stochastically dominates \( G(\cdot) \), the cumulative distribution function of \( g(\cdot) \). By the definition of stochastic dominance, the distributions \( F(\cdot), G(\cdot) \) yield the following inequality:

\[ \int u_0(y(\theta), t(\theta), \theta)dF(\theta) > \int u_0(y(\theta), t(\theta), \theta)dG(\theta). \]

Since the operators \( \int dF \) and \( \int dG \) are contraction mapping, which is \( T : B(X) \mapsto B(X) \) in the proof of lemma 2.3.1, the left and right hand sides have unique fixed point individually. Each constrained maximization problems thus has a unique fixed point as its solutions. Since \( y(\theta) \) is decreasing in \( \theta \), and since the principal’s utility is decreasing in \( t(\theta) \) and increasing in \( y(\theta) \), a unique \( \hat{\theta}_i \) for \( t^*(\theta_i) \) with \( f(\theta_i) \) is less than a unique \( \tilde{\theta}_i \) for \( t^*(\theta_i) \) with \( g(\theta_i) \). \qed

7.2 Geographic Information System (GIS) for Variable Creation

The geographic polygons and the road network that I use for calculating \( \theta_j \) come from the New York Metropolitan Transportation Council (NYMTC)’s Best Practice Model (BPM) for analysis of the New York City metropolitan area’s transportation system. BPM uses Traffic Analysis Zone (TAZ) as its geographic polygon layer, and a detailed road network layer that includes interstate highways, roads and expressways, and bridges and tunnels around New York City.

Figure 7 shows the TAZ polygons and the associated road network. As shown in panel (a),
TAZ polygons mostly have street block detail so that we can expect that treating all origin, or destination points within a TAZ polygon will not cause significant measurement error for the distance calculation because those points are mostly located on the same street block. The road network, shown in panel (b), covers almost all existing roads and highways around the city area. In addition, it has an implicit hierarchy that highways are first, expressway next, and ordinary street roads are last.

This road network is used to calculate the shortest distance for each OD trip. The NYC taxicab data requires 3,580 TAZ polygons for both origin and destinations to compute the optimal distances that corresponds to each observations. The number of distance pairs is thus 12,855,810. The manageable number means that 12,855,810. This is a more manageable number of distances to calculate than the square of 378 million different O-D pair distances in the data.

The shortest route distance calculation for each OD trip is done as follows: i) compute geographic coordinates (latitude and longitude) of a topological centroid for each TAZ polygon, ii) find the nearest point on the road network from the centroid, iii) calculate the euclidean distance from the centroid to the nearest point, which is called a centroid connector. Finally, calculate the shortest route distance along the road network. The shortest route distance plus the centroid connector distances at both origin and destination becomes $\theta_j$ for an OD trip.

Figure 8 shows an example of centroids and the associated centroid connectors around JFK airport. The points in panel (a) represent computed topological centroids of each TAZ polygon. Note that there is more than one centroid connectors in a TAZ polygon, and these connect to the nearest roads and streets that pass through the polygon. For each given trip direction, the shortest one is chosen out of all the centroid connectors in a polygon. Note also that the BPM provides a traffic condition hierarchy depending on trip time, but the distance $\theta_j$ that I calculate is the physically shortest route distance without considering the traffic condition.

7.3 Large Scale GMM Estimation

I perform GMM estimations for the regression model (4.3) with the entire dataset, and subsets for specific area trips to isolate the STM. The reason to perform GMM to estimate
(4.3) is that GMM is the most (statistically) efficient among all (statistically) consistent linear model estimators.

The most notable issue for estimation of the parameters \( \delta, \beta, \) and \( \gamma \) in (4.3) is the data size. The entire trip sample case, for example, has 357.9 million observations and 61 independent variables, so the independent variable matrix with a vector of ones becomes 90Gbytes in a 32-bit computer and 180Gbytes in a 64-bit computer. Furthermore, the entire sample has 151 instrumental variables, so the IV matrix with a vector of ones turns out to be 216Gbytes in a 32-bit computer and 432Gbytes in 64-bit computer.

A sufficient amount of memory from the random access memory (RAM) to a central processing unit (CPU) has to be allocated to perform a numerical computation with real numbers, called floating point numbers in computer science, and each floating point number takes 4 (8) bytes on a 32 (64)-bit computing system. The maximum memory allocation per processor, which is for a standard PC or a serial computing machine, is currently 32Gbytes or 64Gbytes, and therefore, no serial computing systems are capable of performing computation on the data matrices of \( w_j, x_j, \) and the vector of the dependent variable \( y_j \) at the same time.

Even though Big Data analytics has become a well-known subject in the computer science field, calculating large scale data matrices is almost impossible with the standard method of reading the data into a personal computer and performing computations to obtain an analysis output.

For this reason, I use a parallel matrix multiplication algorithm, developed by Kress and Shim (2015) (KS hereafter), to compute GMM estimates of the parameters \( \delta, \beta, \) and \( \gamma \) in (4.3) using the 357.9 million observation taxi trip dataset. The KS algorithm is to compute matrix-transpose-matrix multiplication in a parallel way for both dense and spare form matrices. To illustrate how the KS algorithm works for the GMM estimation, here I present the GMM estimator for the model (4.3). Let \( \theta = (\delta, \beta, \gamma) \) be a vector of the parameters in (4.3). The GMM estimator of \( \theta \) is then given by

\[
\hat{\theta} = \left[ \left( \sum_{i=1}^{N} x_i w_j^i \right) \Sigma \left( \sum_{i=1}^{N} w_j x_i^i \right) \right]^{-1} \left( \sum_{i=1}^{N} x_j w_j^i \right) \Sigma \left( \sum_{i=1}^{N} x_j y_i \right), \tag{7.1}
\]

where \( \Sigma \) is a weighting matrix that controls for heteroskedasticity of the error term\(^{26}\) The estimation requires computation of a transposed matrix-matrix multiplication \( \sum_{i=1}^{N} w_j x_j^i, \)

\(^{26}\)Following the way to obtain the optimal weighting matrix for linear GMM estimation, as Hansen (1982), and Wooldridge (2010), I first compute residuals from two-stage least square (2SLS) estimator \( \hat{\theta} \) as

\[
\hat{\theta} = \left[ \left( \sum_{i=1}^{N} x_i w_j^i \right) \left( \sum_{i=1}^{N} w_j w_j^i \right)^{-1} \left( \sum_{i=1}^{N} w_j x_i^i \right) \right]^{-1} \left( \sum_{i=1}^{N} x_j w_j^i \right) \left( \sum_{i=1}^{N} w_j w_j^i \right)^{-1} \left( \sum_{i=1}^{N} x_j y_i \right). 
\]
a vector-matrix multiplication $\sum_{i=1}^{N} w' j y_j$, and a matrix-diagonal-matrix multiplication $\sum_{i=1}^{N} u_j^2 \cdot w_j w'_j$. For the entire dataset, the multiplication $\sum_{i=1}^{N} w_j x'_j$ requires transposing the 357,880,388 by 152 matrix and multiplying it by the 357,880,388 by 62 matrix. Even though the outcome matrix is dramatically smaller (152 by 62), the matrix multiplication is too large to be done in a personal computer with a single CPU and standard statistical analysis packages.

The KS algorithm parallelizes a matrix-transpose-matrix multiplication and compute it without physically transposing the left side matrix physically, so it reduces memory consumption and the total computing time significantly. In addition, the KS algorithm can be used for sparse matrices multiplication. The independent variable matrix $[x]_j$ and the IV matrix $[w]_j$ have about 357 million times 62, and 357 million times 152 elements respectively, but their nonzero elements are 3.1 billion and 2.1 billion respectively, which is about 14% and 5% of the entire elements. When those matrices are converted to the compressed sparse row (CSR) form that the KS algorithm allows the use of, only 14% and 5% of elements need to be utilized for $\sum_{i=1}^{N} w_j x'_j$, by employing the KS algorithm.

For this reason, I perform the matrix-transpose-vector, matrix-transpose-matrix multiplications for (4.3) using the KS algorithm implemented in the message-passing interface (MPI) written in Fortran. The computation is done on the Cray XE-6, a high performance clustered system in the CUNY High Performance Computing Center.

MPI is a portable, standard interface for writing parallel programs using a distributed memory programming model. The Fortran program for this KS algorithm is well implemented in parallel with MPI, and this program is compiled and executed in the Cray XE-6 that is a high-performance clustered computing system. The Cray XE-6 in the CUNY High Performance Computing Center has 2048 CPUs with 2Gbytes memory per core, and it is designed for this types of parallel computing program implementation. See Kress and Shim (2015) for more detail.

Then the (2SLS) residuals can be computed as $\hat{u}_j = y_j - x_j \hat{\theta}$. With the residual, the optimal weighting matrix $\hat{\Sigma}$ is given by:

$$\hat{\Sigma} = \left( N^{-1} \sum_{i=1}^{N} u_j^2 \cdot w_j w'_j \right)^{-1}.$$  

Note that a matrix representation of $\hat{\Sigma}$ is $W' \Lambda W$, where $\Lambda$ is a diagonal matrix with $\Lambda_{ii} = u_j^2$, for all $i = 1, \ldots, N$. Computation of $\hat{\Sigma}$ is thus done by applying the parallel sparse matrix multiplication subroutine with a diagonal matrix that is in between given two matrices.
### List of Figures

1. The New York City Road Network and a Trip Route .......................... 35
2. Spatial Distribution of NYC yellowcab Pickup .............................. 36
3. Spatial Distribution $f(\theta_i)$ and $g(\theta_i)$ ............................ 37
4. Drivers’ Expected Utilities and Expected Fares ............................ 38
5. Taxi Trip Frequency Maps by ZIP Code Polygons .......................... 39
6. TAZ Polygons and Road Network Maps ...................................... 40
7. Centroid and Centroid Connector: An Example ............................ 41
Figure 1: The New York City Road Network and a Trip Route
Figure 2: Spatial Distribution of NYC yellowcab Pickup
Figure 3: Spatial Distribution $f(\theta_i)$ and $g(\theta_i)$
Figure 4: Drivers’ Expected Utilities and Expected Fares
Figure 5: Taxi Trip Frequency Maps by ZIP Code Polygons
Figure 6: TAZ Polygons and Road Network Maps
Figure 7: Centroid and Centroid Connector: An Example
List of Tables

1 Frequency of Taxi Trips in New York City Boroughs (Weekdays) . . . . . . . 43
2 GMM Estimates of Taxi Trip Attributes . . . . . . . . . . . . . . . . . . . . 44
3 GMM Estimates of Mass Transit Terminal Indicators . . . . . . . . . . . . . 45
4 GMM Estimates of Peak Time Indicators . . . . . . . . . . . . . . . . . . . . 46
5 GMM Estimates of Days of the Week Indicators . . . . . . . . . . . . . . . 47
Table 1: Frequency of Taxi Trips in New York City Boroughs (Weekdays)

<table>
<thead>
<tr>
<th></th>
<th>Peak Time (Morning)</th>
<th></th>
<th></th>
<th></th>
<th>Weekday</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Origin</td>
<td>Destination</td>
<td>Origin</td>
<td>Destination</td>
<td></td>
</tr>
<tr>
<td>New York</td>
<td>39,591,545</td>
<td>36,153,286</td>
<td>54,562,324</td>
<td>52,897,571</td>
<td>222,605,544</td>
</tr>
<tr>
<td>(Manhattan)</td>
<td>(0.870)</td>
<td>(0.838)</td>
<td>(0.868)</td>
<td>(0.850)</td>
<td>(0.860)</td>
</tr>
<tr>
<td>Kings</td>
<td>3,835,304</td>
<td>4,751,837</td>
<td>5,567,111</td>
<td>6,569,716</td>
<td>24,125,653</td>
</tr>
<tr>
<td>(Brooklyn)</td>
<td>(0.084)</td>
<td>(0.110)</td>
<td>(0.089)</td>
<td>(0.106)</td>
<td>(0.093)</td>
</tr>
<tr>
<td>Queens</td>
<td>1,883,714</td>
<td>1,819,962</td>
<td>2,540,357</td>
<td>1,987,656</td>
<td>11,221,342</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.042)</td>
<td>(0.040)</td>
<td>(0.032)</td>
<td>(0.043)</td>
</tr>
<tr>
<td>Bronx</td>
<td>210,506</td>
<td>418,835</td>
<td>191,660</td>
<td>750,287</td>
<td>948,900</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.010)</td>
<td>(0.003)</td>
<td>(0.012)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Richmond</td>
<td>2,214</td>
<td>4,509</td>
<td>2,609</td>
<td>10,264</td>
<td>11,529</td>
</tr>
<tr>
<td>(Staten Island)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>[0.999]</td>
<td>[0.997]</td>
<td>[0.998]</td>
<td>[0.997]</td>
<td>[0.998]</td>
</tr>
<tr>
<td>Others</td>
<td>60,251</td>
<td>130,364</td>
<td>99,780</td>
<td>168,522</td>
<td>393,471</td>
</tr>
<tr>
<td></td>
<td>[0.001]</td>
<td>[0.003]</td>
<td>[0.002]</td>
<td>[0.003]</td>
<td>[0.002]</td>
</tr>
</tbody>
</table>

Notes: The proportions of trips in each New York City boroughs are reported in parentheses, and the proportions in square brackets are of the entire trips that are observed at each time durations. The peak time in morning is between 6 am and 9 am, and the peak time in evening is between 5 pm and 8 pm.
Table 2: GMM Estimates of Taxi Trip Attributes

<table>
<thead>
<tr>
<th></th>
<th>Entire Sample</th>
<th>New York City Related Trip</th>
<th>Manhattan Related Trip</th>
<th>Manhattan/Outside NYC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Within NYC</td>
<td>From To Manhattan</td>
<td>From To Manhattan</td>
<td>From To Manhattan outside NYC</td>
</tr>
<tr>
<td>( \delta )</td>
<td>-167.1358</td>
<td>-5.1487</td>
<td>19.3258</td>
<td>0.6188</td>
</tr>
<tr>
<td>Number of Passengers</td>
<td>-0.2452</td>
<td>0.169</td>
<td>0.080</td>
<td>0.254</td>
</tr>
<tr>
<td></td>
<td>(0.233)</td>
<td>(0.080)</td>
<td>(0.080)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>Itemized Fare</td>
<td>-11.4187</td>
<td>-8.7222</td>
<td>9.2182</td>
<td>-1.5810</td>
</tr>
<tr>
<td>(Meter)</td>
<td>-2.667</td>
<td>(1.818)</td>
<td>(1.460)</td>
<td>(0.917)</td>
</tr>
<tr>
<td>Itemized Fare</td>
<td>-13.2142</td>
<td>-12.8449</td>
<td>12.493</td>
<td>-5.2061</td>
</tr>
<tr>
<td>(Toll)</td>
<td>-2.367</td>
<td>(1.842)</td>
<td>(1.398)</td>
<td>(1.173)</td>
</tr>
<tr>
<td>Itemized Fare</td>
<td>-13.2142</td>
<td>-12.8449</td>
<td>12.493</td>
<td>-5.2061</td>
</tr>
<tr>
<td>(Surcharge)</td>
<td>6.028</td>
<td>4.697</td>
<td>5.469</td>
<td>4.246</td>
</tr>
<tr>
<td>(Credit Card)</td>
<td>2.574</td>
<td>2.075</td>
<td>1.765</td>
<td>1.434</td>
</tr>
<tr>
<td>Fixed Fare Zone</td>
<td>4.208</td>
<td>3.728</td>
<td>3.434</td>
<td>2.496</td>
</tr>
<tr>
<td>(From JFK Airport)</td>
<td>0.2656</td>
<td>0.1787</td>
<td>0.0753</td>
<td>0.037</td>
</tr>
<tr>
<td>Fixed Fare Zone</td>
<td>-0.1892</td>
<td>-0.0468</td>
<td>0.5087</td>
<td>-0.0057</td>
</tr>
<tr>
<td>(To JFK Airport)</td>
<td>0.084</td>
<td>(0.066)</td>
<td>(0.192)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>Special Fare Zone</td>
<td>0.2479</td>
<td>-0.4511</td>
<td>0.1343</td>
<td>0.133</td>
</tr>
<tr>
<td>(To Newark Airport)</td>
<td>(0.220)</td>
<td>(0.111)</td>
<td>(0.133)</td>
<td>(0.133)</td>
</tr>
<tr>
<td>Special Fare Zone</td>
<td>-658.2528</td>
<td>-0.1604</td>
<td>-2.9192</td>
<td>0.6815</td>
</tr>
<tr>
<td>(To Westchester)</td>
<td>(87.537)</td>
<td>(0.143)</td>
<td>(0.451)</td>
<td>(0.207)</td>
</tr>
<tr>
<td>Special Fare Zone</td>
<td>264.7458</td>
<td>-0.8230</td>
<td>-1.4910</td>
<td>0.1213</td>
</tr>
<tr>
<td>(To Nassau)</td>
<td>(34.489)</td>
<td>(0.124)</td>
<td>(0.173)</td>
<td>(0.205)</td>
</tr>
<tr>
<td>Hour Fixed Effect</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Weekday Fixed Effect</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Month Fixed Effect</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>OverID Statistic</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td># of Indep Variables</td>
<td>61</td>
<td>50</td>
<td>56</td>
<td>55</td>
</tr>
<tr>
<td># of IVs</td>
<td>151</td>
<td>150</td>
<td>182</td>
<td>183</td>
</tr>
<tr>
<td># of Observations</td>
<td>357,880,388</td>
<td>356,990,974</td>
<td>956,444</td>
<td>287,455</td>
</tr>
</tbody>
</table>

Note: The asymptotic standard error estimates of the optimal GMM estimators are reported in parentheses. The GMM estimations are performed with 5-digit ZIP Code origin, and destination indicator dummy variables as IVs. The ZIP Code dummy variables are chosen, if 0.1% trip origin and destinations are possessed within each ZIP Code polygon, to avoid the matrix singularity problem in GMM estimate computations. The estimated GMM weighting matrix \( \Omega \) is a diagonal matrix with squared residuals \( \sum_{i=1}^{N} \hat{u}_{j}^{2} \cdot w_{j} \), where the residuals \( \hat{u}_{j} \) are obtained from the two stage least square estimation. OverID statistic refers to the test statistics for overidentification restriction (\( N^{-1/2} \sum_{i=1}^{N} \hat{u}_{j} \cdot w_{j} \), (\( N^{-1/2} \sum_{i=1}^{N} \hat{u}_{j} \cdot w_{j} \))^T \( \Omega \) (\( N^{-1/2} \sum_{i=1}^{N} \hat{u}_{j} \cdot w_{j} \))^T \( \chi_{L-K}^{2} \), where \( L \) is the number of IVs, \( K \) is the number of independent variables, and \( \hat{u}_{j} \) is the residuals from the GMM estimates.
<table>
<thead>
<tr>
<th></th>
<th>Entire Sample</th>
<th>New York City Related Trip</th>
<th>Manhattan Related Trip</th>
<th>Manhattan/Outside NYC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Within NYC</td>
<td>From NYC</td>
<td>To Manhattan</td>
</tr>
<tr>
<td>$\delta$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-167.1358</td>
<td>-5.1487</td>
<td>19.3258</td>
<td>0.6188</td>
</tr>
<tr>
<td>Fixed Fare Zone</td>
<td>-4.5392</td>
<td>-0.5283</td>
<td>-0.3244</td>
<td>0.6073</td>
</tr>
<tr>
<td>(From JFK Airport)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed Fare Zone</td>
<td>-0.1892</td>
<td>-0.0468</td>
<td>0.5087</td>
<td>-0.0057</td>
</tr>
<tr>
<td>(To JFK Airport)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>From LaGuardia</td>
<td>-0.0981</td>
<td>-0.1286</td>
<td>0.8953</td>
<td>-0.0057</td>
</tr>
<tr>
<td>To LaGuardia</td>
<td>-0.0101</td>
<td>-0.0654</td>
<td>0.7929</td>
<td>0.0058</td>
</tr>
<tr>
<td>Special Fare Zone</td>
<td>0.4279</td>
<td>-0.4511</td>
<td></td>
<td>0.1343</td>
</tr>
<tr>
<td>(To Newark Airport)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>From Grand Central</td>
<td>0.5716</td>
<td>0.8522</td>
<td>-1.1890</td>
<td>0.6012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.110)</td>
<td>(0.846)</td>
<td>(0.107)</td>
</tr>
<tr>
<td>To Grand Central</td>
<td>-0.2310</td>
<td>-0.0633</td>
<td>-0.1331</td>
<td>-0.1062</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.108)</td>
<td>(0.911)</td>
<td>(0.096)</td>
</tr>
<tr>
<td>From Port Authority</td>
<td>3.7885</td>
<td>6.8908</td>
<td>0.3308</td>
<td>-1.3759</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.444)</td>
<td>(1.446)</td>
<td>(0.428)</td>
</tr>
<tr>
<td>To Port Authority</td>
<td>-2.5269</td>
<td>-1.1294</td>
<td>-0.6525</td>
<td>-0.6152</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.486)</td>
<td>(2.221)</td>
<td>(0.478)</td>
</tr>
<tr>
<td>From Penn Station</td>
<td>0.0296</td>
<td>0.5553</td>
<td>-5.4169</td>
<td>-0.2049</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.038)</td>
<td>(0.226)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>To Penn Station</td>
<td>0.0106</td>
<td>0.042</td>
<td>-0.1947</td>
<td>0.0127</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.037)</td>
<td>(0.296)</td>
<td>(0.035)</td>
</tr>
</tbody>
</table>

Note: The asymptotic standard error estimates of the optimal GMM estimators are reported in parentheses.
Table 4: GMM Estimates of Peak Time Indicators

<table>
<thead>
<tr>
<th></th>
<th>Entire Sample</th>
<th>New York City Related Trip</th>
<th>Manhattan Related Trip</th>
<th>Manhattan/Outside NYC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Within NYC</td>
<td>From To</td>
<td>Within From To To</td>
<td>From Manhattan Outside NYC</td>
</tr>
<tr>
<td>( \delta )</td>
<td>-167.1358</td>
<td>-5.1487 19.3258 0.6188</td>
<td>73.6579 31.8456 -19.3355</td>
<td>29.8041 -4.1160</td>
</tr>
<tr>
<td>6 am</td>
<td>-0.2498</td>
<td>-0.0685 -0.9905 -1.0522</td>
<td>-0.0904 0.9191 -0.3028</td>
<td>0.4931 -0.9475</td>
</tr>
<tr>
<td></td>
<td>(0.112)</td>
<td>(0.085) (0.225) (0.376)</td>
<td>(0.089) (0.119) (0.109)</td>
<td>(0.271) (0.700)</td>
</tr>
<tr>
<td>7 am</td>
<td>0.0846</td>
<td>0.3547 -1.9744 0.8005</td>
<td>-0.1144 0.9460 0.5979</td>
<td>1.0167 -0.2048</td>
</tr>
<tr>
<td></td>
<td>(0.098)</td>
<td>(0.073) (0.236) (0.361)</td>
<td>(0.079) (0.111) (0.099)</td>
<td>(0.292) (0.673)</td>
</tr>
<tr>
<td>8 am</td>
<td>-0.1264</td>
<td>0.0208 0.5583 -6.7789</td>
<td>-0.2433 1.7056 0.6824</td>
<td>1.3629 -2.7039</td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
<td>(0.073) (0.232) (0.355)</td>
<td>(0.078) (0.119) (0.099)</td>
<td>(0.293) (0.670)</td>
</tr>
<tr>
<td>5 pm</td>
<td>0.0150</td>
<td>-0.2242 0.3827 57.1002</td>
<td>-0.2548 3.0043 0.4421</td>
<td>1.6375 18.8968</td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
<td>(0.076) (0.228) (0.346)</td>
<td>(0.087) (0.110) (0.096)</td>
<td>(0.299) (0.696)</td>
</tr>
<tr>
<td>6 pm</td>
<td>-0.2301</td>
<td>-0.1909 -0.0976 56.3670</td>
<td>-0.0027 1.2873 1.4127</td>
<td>3.2757 61.8889</td>
</tr>
<tr>
<td></td>
<td>(0.097)</td>
<td>(0.071) (0.234) (0.351)</td>
<td>(0.082) (0.103) (0.094)</td>
<td>(0.313) (0.651)</td>
</tr>
<tr>
<td>7 pm</td>
<td>-0.3889</td>
<td>-0.2165 6.5597 55.7061</td>
<td>-0.1931 2.3056 4.5423</td>
<td>-3.8859 61.3666</td>
</tr>
<tr>
<td></td>
<td>(0.107)</td>
<td>(0.079) (0.249) (0.343)</td>
<td>(0.086) (0.112) (0.094)</td>
<td>(0.311) (0.658)</td>
</tr>
</tbody>
</table>

Note: The asymptotic standard error estimates of the optimal GMM estimators are reported in parentheses.
Table 5: GMM Estimates of Days of the Week Indicators

<table>
<thead>
<tr>
<th></th>
<th>Entire Sample</th>
<th>New York City Related Trip</th>
<th>Manhattan Related Trip</th>
<th>Manhattan/Outside NYC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Within NYC</td>
<td>From</td>
<td>To</td>
</tr>
<tr>
<td>( \delta )</td>
<td>-167.1358</td>
<td>-5.1487</td>
<td>19.3258</td>
<td>0.6188</td>
</tr>
<tr>
<td>Monday</td>
<td>0.0068</td>
<td>-0.0134</td>
<td>-0.0155</td>
<td>-0.1112</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.031)</td>
<td>(0.115)</td>
<td>(0.142)</td>
</tr>
<tr>
<td>Tuesday</td>
<td>-0.0255</td>
<td>0.0273</td>
<td>0.0346</td>
<td>-0.3338</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.032)</td>
<td>(0.119)</td>
<td>(0.143)</td>
</tr>
<tr>
<td>Wednesday</td>
<td>-0.2865</td>
<td>-0.0386</td>
<td>0.1508</td>
<td>-0.5353</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.033)</td>
<td>(0.119)</td>
<td>(0.143)</td>
</tr>
<tr>
<td>Thursday</td>
<td>0.1960</td>
<td>-0.4035</td>
<td>0.2430</td>
<td>-0.9592</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.045)</td>
<td>(0.118)</td>
<td>(0.156)</td>
</tr>
<tr>
<td>Friday</td>
<td>-0.2948</td>
<td>-0.0994</td>
<td>0.1903</td>
<td>-0.7600</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.040)</td>
<td>(0.115)</td>
<td>(0.150)</td>
</tr>
<tr>
<td>Saturday</td>
<td>0.2499</td>
<td>-0.0932</td>
<td>0.0978</td>
<td>-0.6140</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.029)</td>
<td>(0.118)</td>
<td>(0.147)</td>
</tr>
</tbody>
</table>

Note: The asymptotic standard error estimates of the optimal GMM estimators are reported in parentheses.