Regulatory Distortion: Evidence from Uber's Entry Decisions in the US[1]

Haldun Anil, Braze, and Sara Fisher Ellison, MIT and CESifo

September 2022

[1]We thank Anna Ellison, Caroline Ellison, Kate Ellison, Jia Yi Goh, Tanush Garg, Zoe Greenberg, Eileen Liu, Nahom Marie, Krista Moody, Sicong Shen, James Simon, Allan Wu, and Maren Young for their outstanding research assistance. We also thank Glenn Ellison, Tobias Salz, Christine Zulehner, and numerous seminar participants for valuable comments. Finally Ellison thanks the Paris School of Economics for their hospitality and collegiality. E-mail: haldun@alum.mit.edu, sellison@mit.edu.

Abstract

There is a large and long-standing literature on the distortionary effects of regulations on the functioning of markets. A newer strand of this literature focuses on licensing regulations, such as state-specific licensing of teachers and hairdressers. We seek to add to this literature with the specific case of ride-hailing services, such as Uber. We assemble a new and comprehensive data set of 250 US cities and their regulations regarding hackney services. We specify a stylized profit model for Uber, which is a function of these regulations, and estimate the parameters of the profit function using observed Uber entry decisions into these cities. Our data set and empirical strategy allow us to estimate the differential effects of particular types of regulations, separating out regulations governing safety, governing operations, and erecting entry barriers. We find that safety regulations do not have a distortionary effect on the functioning of the market for hackney services and evading them does not increase Uber's profits. We find evidence that Uber's profits are increased, however, by their ability to evade regulatory entry barriers and regulations governing operations. In other words, those regulations do have a significant distortionary effect on the market. To the extent that safety-related regulations are welfare-enhancing and those erecting entry barriers are welfare-decreasing, our results suggest a welfare-enhancing effect of Uber's entry.

1 Introduction

Since Uber launched its ride-hailing service in San Francisco in March 2009, the transformation of for-hire transportation has been nothing short of remarkable. Google Trends indicates that the search terms "uber" and "lyft" are, together, four times as popular as the search term "taxi" in the US. See Figure 1 for the evolution over time of the relative popularity.

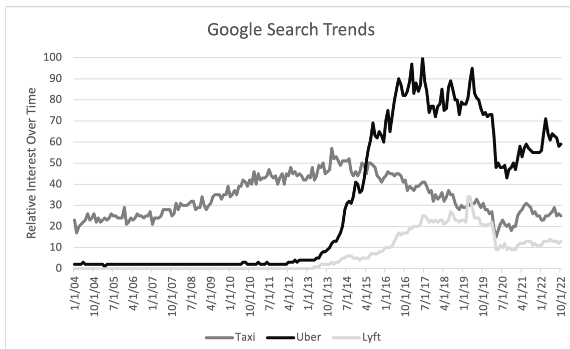

Figure 1: Google Search Trends For Ride-Hailing Services

As another indication, see Figure 2. In the span of a decade and a half, ride-hailing services have literally decimated the storied New York City taxi industry, capturing more than 90% market share in that city.[1] Uber has entered 785 metropolitan areas worldwide in 70 countries, as of 2022, and enjoys a market capitalization of around 60 billion dollars. Although competitors, such as Lyft, have enjoyed substantial success as well, Uber has managed to retain a 68% share of the US ride-hailing market.

Consider also the effect on well-established markets in a few major US cities for buying and selling medallions, or the right to operate a taxicab.[2] Figure 3 shows time series of taxicab medallion prices in six cities.[3] We have noted on the figures the dates at which both Uber and Lyft entered each market. What is clear from these graphs is that the markets

[1] https://toddwschneider.com/dashboards/nyc-taxi-ridehailing-uber-lyft-data/

[2] Most US cities that have active taxicab fleets either do not have medallion systems or do not have active markets in the sale of these medallions. We were able to find fairly complete data on five medallion systems, New York, Chicago, Boston, Miami, and Philadelphia. New York, Chicago, and Boston seem to have the most active markets in medallions and, therefore, more complete data. Note that many other municipalities use other regulatory levers to control or prevent entry.

[3] The data used to create the medallion graphs were obtained from both regulatory websites and private websites.

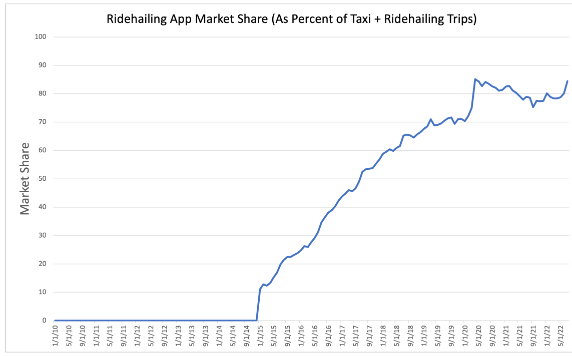

Figure 2: Ridehailing App Market Share

for medallions have collapsed over the past couple of years. Transactions prices leveled off soon after Uber entry and then fell precipitously. Volume also dropped dramatically, often before price fell, perhaps reflecting uncertainty in the market causing a hesitance to transact. This empirical setting, ride-hailing's remarkable ascendancy, and the fall of legacy hackney services, [4] serve as the backdrop for our study of a broader phenomenon, regulatory distortion. There is a vast literature spanning decades on the sources and effects of government regulation of industry. Often, the literature, both theoretical and empirical, focuses on one particular facet of regulation, perhaps price caps, or emissions standards, or licensing requirements. But, of course, in many industries, regulation is not a monolith. It can take many forms, even within the same regulatory framework for a particular industry. The regulatory environments for hackney services in the US are examples of complex frameworks, governing geographical coverage, fares, driver training, car maintenance, emissions, safety, advertising, and communications, to name just a few. We want to exploit the existence of these different frameworks across US cities, along with the actions of one firm, Uber, to shed light on the possible distortionary effects of different types of regulation.

The extant empirical literature on regulation falls mostly into one of four categories. Historical studies present a narrative of the function (or dysfunction) of a particular market over time and the evolution of the regulatory regime as a response. An example of such a study is Hodges (2007), which offers a detailed history of taxicab regulation in New York

[4] Since we are Bostonians, we adopt the official terminology used here, "hackney services." We will also use the more familiar "taxicab" interchangeably with "hackney."

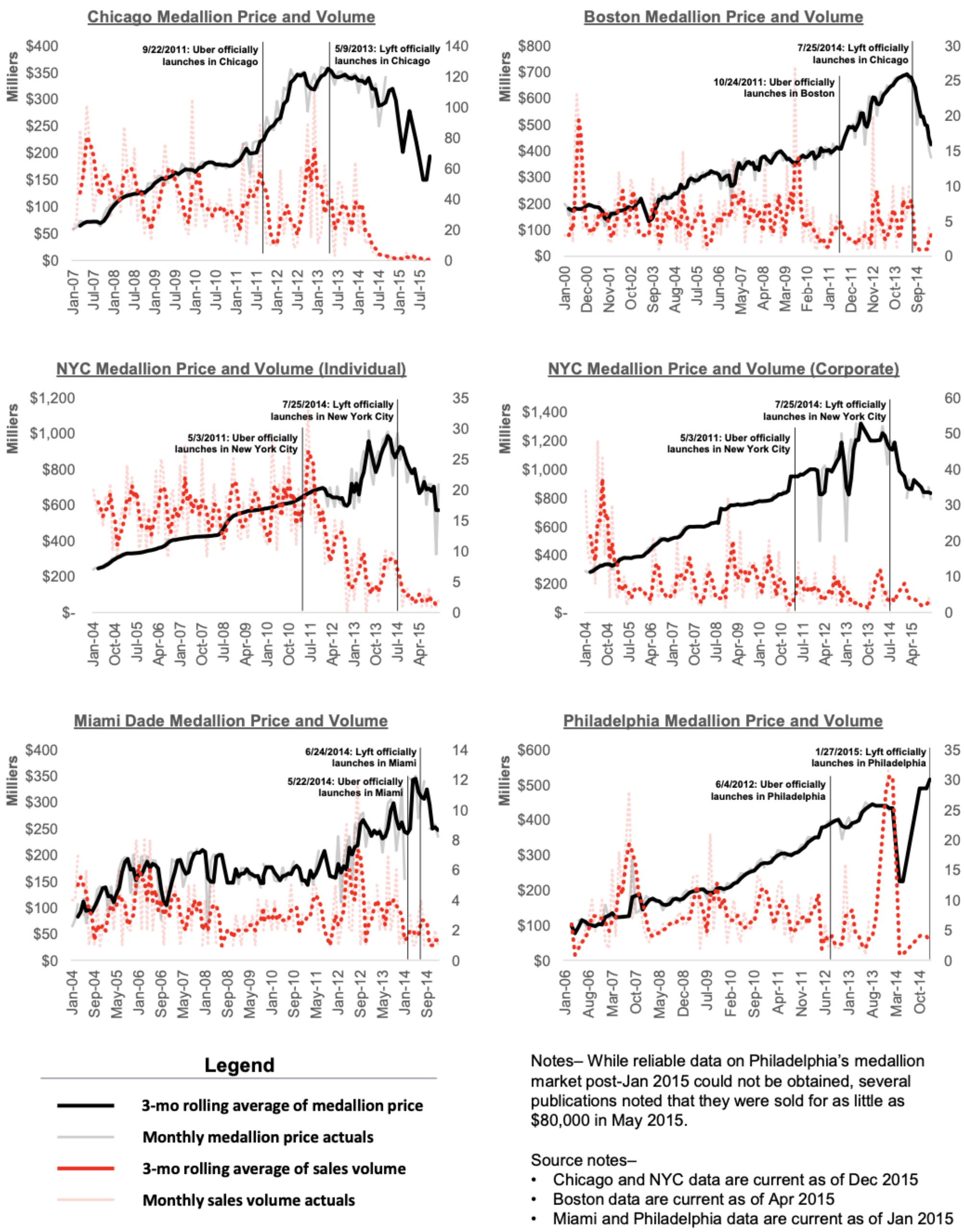

Figure 3: Time series of taxicab medallion prices

City. Event studies, a second category, examine the functioning of a particular market before and after a regulatory event, such as deregulation. Leading examples in this category are Rose (1985), which examines the effects of deregulation in the trucking industry, and Grabowski and Vernon (1992), which studies the effect of the Hatch-Waxman Act on the behavior of generic pharmaceutical manufacturers. A third category is the cross-sectional study, which looks at some outcome such as pricing, competition, or entry, across similar markets that differ in regulatory rules or enforcement. Kyle (2007) performs such a study on entry of pharmaceuticals into different geographic markets in Europe with differing regulatory regimes. Finally, structural studies exploit theory and equilibrium conditions to estimate the effects of regulation, such as in Ryan's (2012) study of the cost of environmental regulation in the cement industry.

Our paper relies primarily on cross-sectional variation to identify effects of interest, like the third category, but does leverage models of entry to provide functional form guidance, as in the fourth category. Our study, however, has two features absent from the existing literature. First, the regulatory regimes we observe are complicated, multi-dimensional, and varied. So we have an unique opportunity to comment on different *types* of regulations in a specific setting. Second, we are using the actions of a firm not subject to the regulations (or at least behaving as if it were not subject to the regulations) to infer effects on the firms that are subject to them.

In the sense outlined above, we think that this paper fits squarely into the classic literature on the effects of regulation. We do, however, want to argue that are special features of regulation in the digital age worth closer study. In particular, new, disruptive entrants as well as entirely new industries, are fueling a sort of regulatory arbitrage. In fact, there are a number of new markets that may exist primarily for the purpose of avoiding taxation and regulation. For many years, online retailers in the US gained significant advantage over brick and mortar retailers by having *de facto* tax-free status.[5] See Ellison and Ellison (2009) for a discussion and estimation of the advantage. Platforms such as Upwork and TaskRabbit allow employers to skirt labor laws in many jurisdictions. Airbnb rentals function much like hotel rooms in many locations but are typically not subject

[5] Recent Supreme Court rulings have significantly eroded that special status.

to the same regulatory structure. Farronato and Fradkin (2022) consider this issue in their study of the welfare effects of peer entry in the accommodation industry. Finally, cryptocurrencies are used by some to evade the scrutiny that comes with traditional banking sector regulations and are traded in many places free of the consumer protections that exist in equities markets. Understanding both regulatory distortion and the motives for regulatory avoidance in the digital economy could become increasingly important.

We have compiled a detailed and comprehensive data set of the regulations imposed on hackney services by 250 cities and towns in the US circa 2015, some of which experienced Uber entry by July 2015 and some of which did not. Furthermore, conditional on Uber entry, we know the date of entry. We can, therefore, examine the pattern of entry over time as a function of the regulatory environment to assess the relative importance of particulars in that environment. Specifically, we formulate a simple model of Uber profit as a function of city demographic, structural, and regulatory characteristics and are able to estimate parameters of this model as well as characteristics of the distribution of entry costs. We will flesh out our empirical strategy further in later sections.

First, though, it is important to note that regulations are not all created equal. They can have vastly different intents and effects on profits, market structure, and consumer welfare. A regulation requiring a potential entrant of hackney services to "prove pubic necessity" is clearly intended to deter entry and protect incumbents. A regulation requiring background checks of taxi drivers is likely intended to alleviate an asymmetric information problem and ensure the safety of passengers riding alone in cars with drivers. In our analysis, we will allow for these different types of regulations to have different effects on market distortion and Uber entry decisions.

The main economic argument in favor of the regulatory structure is the ability to license and monitor owners and drivers to ensure higher quality (availability, better driving skills, knowledge of local area, honesty, cleanliness, safety, environmentalism, lack of discrimination, etc.). Anyone who has ever ridden in a taxi can attest to the fact that hackney commissions have had uneven records regarding ensuring quality along those dimensions. This observation suggests an alternative view of the regulatory structure: hackney commissions might have originally come into existence to ensure quality but they have since been captured by the industry and now work largely to protect monopoly power and capture rents for the taxicab owners and drivers. It is difficult to justify regulations on geographic licensing, for instance, which require a Boston taxi taking a customer to Cambridge to return empty, on quality or efficiency grounds. Nothing we do relies on a specific view of these regulations and their intentions and effects. It is, perhaps helpful, though, to bear these notions in mind as we interpret results.

In addition to shedding light on the role of the regulatory structure in the ascendency of ride-hailing services, an auxiliary but important motivation of our paper is simply to build on an existing literature documenting how a set of technologies is transforming a long-standing industry. Economists have studied the markets for hackney services before, and the advent of ride-hailing services has caused an uptick in research interest. An important paper by Camerer et al (1997) estimated the labor supply elasticity of New York City taxicab drivers and found a large, negative, and significant elasticity. Their explanation, grounded in behavioral theory, is that taxicab drivers shoot for a certain target income each day and, when reached, supply their labor less willingly. A follow-on literature, for instance Farber (2005) and Crawford and Meng (2011), sought alternative explanations or nuance to the original result of Camerer et al. Recent papers by Bucholtz (2017) and Frechette, Lizzeri, and Salz (2016) exploit a rich data set made available by New York City of all taxicab rides from 2009-2013. They both estimate structural models of ride supply and demand to answer various questions about market functioning and perform policy counterfactuals. There is a growing literature looking at ride-hailing services specifically. We mention a few notable papers below. Greenwood and Wattal (2015) use a difference-in-difference approach to exploit the natural experiment of the entry of two different Uber services in California towns. They find a significant drop in the rate of alcohol-related vehicle homicides after the introduction of Uber. Hall and Krueger (2015) provide a descriptive study using proprietary Uber data on drivers, which establishes many interesting facts about the driver population and behavior. Chen et al (2017) study the value of flexibility to Uber drivers, again using proprietary Uber data, and find that drivers earn more than twice the surplus they would relative to less flexible work arrangements. Angrist and Caldwell (2021) also study aspects of the labor market for Uber drivers. Finally, a number of papers have studied the congestion effects of ride-hailing, including Barrios et al (2019), Tardano (2021), and Schaller (2021). Finally, as suggested above, our study is relevant to the literature and current debate on occupational licensing. Many of the regulations applying to hackney services are akin to licensing requirements. Drivers sometimes need special municipality-specific hackney licenses, and firms offering services often need to register and apply for licenses to operate in a specific municipality. Furthermore, these licenses can be limited in number. The labor literature on licensing has pointed out the inefficiencies that licensing requirements might cause by erecting unnecessary barriers to entry and hampering worker mobility. See, for instance, Kleiner and Krueger (2008) and Gittleman and Kleiner (2013). Onerous licensing requirements in some industries, such as hair styling, appear to be clear-cut cases of regulatory capture. In other industries, such as healthcare, there is broad concensus about the need for licensing, but these requirements can still be captured by market participants and used as a device to ensure entry at levels suboptimal for consumers.

So far, the labor literature has documented the prevalence and growth of licensing requirements in detail, and has also looked at the effect on wages, but little work has attempted to quantify the other possible distortionary effects of these regulations. Our study can comment on this issue by seeing how the regulations impact the entry decisions of a potential competitor who would not be subject to them. In other words, regulations which had a large distortionary effect on the market by, say, ensuring a suboptimal number of hackney licenses (from a welfare perspective) would create a large incentive for a potential entrant to exploit the distortion. Observing Uber's entry decisions as a function of licensing and other regulations will help us gauge the magnitude of the distortionary effects.

We find evidence that regulations creating barriers to entry did result in market distortions, leading to significantly earlier entry by Uber. In contrast, we find no evidence that safety-related regulations caused market distortions that were exploited by Uber. The evidence on what we would call operations-related regulations—mandated 24/7 operation, for instance—was mixed.

In the following section, we recount a brief history of hackney operation and regulation.

Section 3 discusses the data set for the analysis. Section 4 presents the empirical model and results. We conclude in Section 5.

2 A Short History of Hackney Services and Regulation

It is difficult to know the exact historical origins of for-hire transportation.[6][7] We do know, though, that a direct antecedent of modern taxicab fleets, for-hire hackney carriage services, began operating in both London and Paris in the early 17th century. These services, obviously horse-drawn at the time, were often hired by innkeepers for their guests. The first regulation of these services appeared in 1635, when British Parliament passed the Hackney Carriage Act. It was followed by the Ordinance for Regulation of Hackney-Coachmen in London and the Places Adjacent in 1654. The first hackney-carriage licenses were issued in 1662. Other cities followed London's lead both in terms of services available and regulation of those services.

A mechanical "taximeter" to measure the fare was developed in the 19th century, giving rise to the term "taxicab," a portmanteau of "taximeter" and "cabriolet." Something akin to a modern taxicab became common around the turn of the 20th century: gasoline-powered, meter-equipped vehicles for hire were operating in major cities in Europe and North America. Demand for services grew, as did the fleets, and by the 1910's, New York City had half a dozen large fleets, mostly owned by automobile manufacturers, as well as thousands of independent owner/drivers.

In New York City in 1934, to protest wages and corruption in the industry, 2000 taxi drivers took over Times Square, crippling the city. Mayor Fiorello La Guardia responded by signing the Haas Act of 1937, introducing the medallion system that exists to this day in New York City. The system was intended to put a strict and binding limit on supply, in the case of New York City, capping the number of taxicabs at 16,900, therefore helping to ensure adequate (typically excess) demand for each taxicab. The number of medallions in New York has actually decreased since 1937 and now stands at 13,437.[8] It is interesting

[6] Gondolas, for instance, were first mentioned in a letter from a Venetian Republic official in 1094 (See "The Rough Guide to Venice and the Veneto.")

[7] Information from this section was taken from various documents and websites we cite in the references. We do not provide the source of every fact separately.

[8] See NYC Taxi & Limosine Commission Factbook.

to note that in the case of the Haas Act, at least, the regulation was demanded by drivers and aimed at controlling entry into the market.

Not all cities have medallion systems, but several major US cities, such as New York, Chicago, Philadelphia, Boston, and Atlanta do. In those cities, medallions can be bought and sold,[9] and, in some cases, active markets and extensive data on prices and quantities exist. Many cities do not have medallion systems, but use other regulatory levers to control or prevent entry of taxicabs. We discuss the nature and scope of regulation in more detail in the data section.

Other regulations governing operations have been enacted in various places. For instance, matching between taxicabs and customers seeking a ride is aided in some cities by extensive use of cab stands, also known as hack stands or cab ranks. These are designated areas where taxicabs can queue up to wait for customers. Some cities, such as Boston and Cambridge, use them extensively. In fact, since a large fraction of taxicabs are waiting in cab stands, hailing a taxicab on the street can often take much longer than walking to the nearest cab stand. New York City, in contrast, does not make such extensive use of cab stands. Some cities mandate that taxicab companies must operate around the clock. Others mandate that cars must be dispatched by radio, although such a regulation seems counterproductive given recent advances in dispatching and wireless communication.

Taxicab fares are routinely capped, and sometimes mandated, by regulation in U.S. cities (as well as internationally). These regulations vary locally or at the state level and are the result of local costs and wages, lobbying by firms, customer input, and other factors. Note that even if the regulation on taxicabs comes in the form of a cap as opposed to a mandate, meters typically charge the cap without the possibility of lowering the fare during periods of low demand or high supply. Such a practice results in the cap behaving more or less like a mandate. In other words, fare stickiness should be thought of as the result of both regulatory restrictions and technological limitations of the meters. The combination of these two factors makes adaptive, or equilibrium, pricing for taxicabs close to impossible. So period of high demand cannot be addressed by temporarily raising fares, thus both inducing a supply response and causing low-valuation passengers to seek alternative transportation.

[9] San Francisco has a medallion system but prohibited the private sale of medallions in 1978.

It should also be noted that taxicabs in the US accept (and expect) tips, so reported fares understate actual fares.[10]

As we discuss in the following section, we assembled a comprehensive data set of hackney regulations in 250 US cities. In doing so, we came across a wide variety of different regulations, not all of which made it into our final data set. In addition to ones mentioned above we found explicit caps on the number of taxicabs, geographic restrictions on pickups, limits on advertising in and on vehicles, dress codes for drivers, required proof of public need for entry, and many others.

A crucial point to emphasize (and a key to our empirical strategy) is that Uber was not subject to the regulatory structures that applied to taxicabs in US cities, or, at the least, behaved as if they were not. This situation resulted from the working definitions that most cities used to regulate hackney industries: a taxicab was a car that could pick passengers up on the street when hailed or park at cabstands and wait for passengers there. Cars that passengers would have to book ahead of time were typically subject to different and much less stringent regulatory structures, if any. Ubers, of course, did neither of the things that would categorize them as taxicabs. Furthermore, Uber argued that its drivers were independent contractors rather than employees, so regulatory compliance, if any occurred, was not Uber's responsibility.

The advent of ride-hailing services has prompted a rethinking of taxi deregulation in many US cities as well as discussions of how cities can bring Uber under their current regulatory regime. Some attempts at deregulation of hackney services predate the entry of ride-hailing services, of course. A notable example is the Irish deregulation in 2000. Barrett (2003) discusses the particulars of the reform and estimates its effects: a three-fold increase in the number of taxis, much reduced passenger wait times, and no discernible reduction in either driver or vehicle standards.

[10]In Boston, UberX had a base fare of \$2 when it first entered. Fares increased \$0.20 a minute and \$1.24 a mile. Compare this with taxis at the time, which had a base fare of \$2.60, increasing at a rate of \$2.80 a mile. Unlike taxis, though, Uber fares can and do vary with supply and demand conditions, so base fares are often in force but not always.

3 Data

The data for our analysis of Uber entry and regulatory environments are drawn from many sources. We gathered data on entry dates from Uber's website (including archived pages from Internet Archive). We gathered city-level demographics and other covariates from various sources, including census reports. Finally, we put together a detailed and comprehensive data set on regulations governing taxi drivers and hackney industries. These data were compiled from city and state websites, the US Department of Labor's website, and phone calls to local regulatory bodies. All of these data sets are at the city level, although the definition of a city varies across data sources. The regulations, for instance, are often at the level of municipality: Boston's regulations differ from Cambridge's, which is across the river. They also each have separate taxi fleets. Uber, however, entered the Boston metropolitan area at once, not Boston and Cambridge separately. In other words, the Uber entry data are at the metropolitan area level whereas regulations could vary within those metropolitan areas. We describe below how we handle these mismatches.

First, however, we describe how we create our measures of regulatory burden. Recall that we want to see how important regulatory explanations are to explaining equilibrium entry behavior of Uber. In order to carry out such an analysis, we must have a measure or measures of regulatory burden. Our construction of measures was guided by a detailed study on the regulatory hurdles imposed on potential entrants to the taxicab market in large cities in Ohio (prepared by the Buckeye Institute for Public Policy Solutions). The detail in that study allowed us to define potential areas of regulation: ranging from location of dispatch operations to appropriate clothing for the drivers, from fuel efficiency and emissions standards to limits on advertisements on vehicles. After identifying these potential areas, we then called or visited websites of local governing bodies and did one of two things: For variables that naturally were binary in nature, for instance, whether taxis were only allowed to pick up at cab stands, we coded a 1 for a municipality that imposed that regulation and 0 otherwise. For variables that were linear in nature, for instance, minimum age of driver, we transformed them to take on values between 0 and 1 where 1 was the most restrictive.

From this larger set, we retained regulations where our data collection effort had reasonable coverage. For instance, we were only able to obtain information on a couple of cities which regulated attire for taxi drivers. We might have assumed that no other cities had such regulations and assigned them 0, but, absent definitive information, we decided to omit the areas on which our information was particularly sparse. We are not overly concerned about these data issues, however. Since our goal is to determine whether a city's regulatory environment had an effect on the timing of Uber's entry into that city, we might expect that Uber's decisions would be based on the *observable* regulatory environment, which should be fairly accurately characterized by our data collection effort.

We then used these variables to create composite indexes for intensity of regulation in three broad categories. The indexes were created to ease interpretation but also out of necessity. Even after only retaining potential areas of regulation where our information was not too sparse, and despite hundreds of hours of effort combing through webpages and phoning agencies, we still had missing values for many cities for at least some of the areas of regulation. Creating the indexes allowed us to retain observations by creating non-missing values of the index when some of the constituent variables were missing.

Our three indexes are $Safety$, *Operations*, and *Barriers*. $Safety$ measures the intensity of safety-related regulations, such as requirements on drivers and vehicles. *Operations* is the index which includes regulations involving hours of operation, size of fleet, method of dispatch, and so forth. *Barriers* includes regulations that we interpreted as having the primary purpose of erecting entry barriers for new operators without other discernible effects. These base indexes were computed using only affirmative information, positive or negative, that we were able to find about each regulation. We were concerned, however, that the indexes were not based on the same regulations for each city due to missing values, so we computed an alternative version for each of these three indexes, denoted with a 0. In this alternative version, we assumed that a specific regulation did not exist in a city if we could find no mention or evidence of it. Not only do we think this assumption is quite reasonable—it should be much easier to find positive evidence of the existence of a regulation than negative evidence of its absence—but it also allows us to retain additional observations. Because it is an assumption, however, we maintain both versions of the regulatory variables and present our results with both.

Table 1 lists the variables which went into each index. The three cities scoring highest in our regulation measures were Los Angeles CA, Tupelo MS, and Modesto CA, each with averages across categories of 0.85 or more.[11] For comparison, Boston MA had an average of 0.57, Newport News VA had 0.73, Oklahoma City OK had 0.35, and Scranton PA had 0.21.

Table 1: Regulatory Indexes

Safety
maximum age of cars
minimum age for drivers
maximum workday
physician's certificate required
fingerprint required
background check required
Operations
maximum fare
dictated fare
24/7 service
cab stand-only pick-up
mandated radio dispatch
restricted advertising on car
max number of taxis per operator
Barriers
fee for driver
fee to start company
fee to license each car
minimum fleet size
insignia approval
proof of public need

Summary statistics for those three indexes are included in Table 1. Also included in the table are summary statistics on the Uber entry data as well as the main city-level demographic variables we gathered. *UberEntry* was converted to an integer where 1 is March 1, 2009, the date that Uber was founded. They entered San Francisco and San Jose on March 4, 2011, which marked the beginning of their operations. The last Uber entry in

[11] The cities scoring lowest, less than 0.05 on average, were Bowling Green KY, Klamath Falls OR, Laurel MS, Mount Vernon IL, Rhinelander WI, Temple TX, Valley City ND, and West Point MS.

our data set is Moline/Rock Island IL, which was entered 1,614 days after San Francisco and San Jose. We have 89 cities which had not experienced entry at the time of our data collection. *Subway*, *Bus*, *LtRail*, *Trolley*, and *Ferry* are indicator variables for whether the city had that particular type of transit system. Some cities have multiple types. We also gathered data on population, population density, average wage, and the percent of households without a vehicle.

As mentioned earlier, an Uber entry event might cover multiple municipalities. One example is when Uber entered Moline and Rock Island, Illinois simultaneously. The two cities are adjacent (and across the Mississippi River from Davenport and Bettandorf, Iowa, incidentally), so treating them as a single market for entry made sense. When considering whether and when to enter, Uber would have taken into account the demographics and regulatory environments in both cities, so we aggregated them using a population weighted average of the demographic and regulatory variables of the two cities along with the sum of their populations.

We should emphasize that the regulatory data are a snapshot from 2015. Gathering historical data on regulations—to periods before 2015—was next to impossible. Anecdotal evidence suggested a lot of statis pre-2015. Some cities had not updated them for decades. The 2015 collection date allowed us to use entry data from 2009 to 2015, giving us a reasonable number of entry events. Furthermore, cities had not started to amend their regulations to either deter ride-hailing entry or accommodate it, so 2015 regulations could be taken as exogenous in a way that, say, 2022 regulations could not be.

4 Empirical Strategy and Results

We start in the spirit of Bresnahan and Reiss' (1991) classic paper on entry of firms into geographically isolated markets. We specify a simple model of Uber's profits and how they relate to its entry decisions. Let

$$\Pi_{it} = \Pi(X_i\beta, t, C_i) = X_i\beta + \log t - C_i,$$

where t is calendar time and i indexes cities. We can think of Π_{it} as being the expected per period profits Uber receives from entry into city i at time t, so Uber will enter if and only if $\Pi_{it} > 0$. The fact that our specification requires that Π_{it} be monotonic in t prevents a situation where Uber would enter and then subsequently exit a city.[12] In other words, Uber should enter around the time when $\Pi_{it} = 0$ (exactly so, if t were continuous). Here, C_i is a city-specific per period operating cost that Uber must incur once it enters city i, and we assume that it is drawn from some distribution. The Xs are, again, city-specific and could include both demographics and structural characteristics of the cities as well as measures of the regulatory environment. Note that we do not include a coefficient on $\log t$ as a parameter in the model since it will not be identified.

It is useful, perhaps, to pause at this point to discuss two issues: first, why regulations to which Uber is not subject would enter its profit function, and second, whether basing an empirical strategy on the actions of one firm is asking too much. On the first point, if regulations are sufficiently distortionary, they create market opportunities for firms that are not subject to them. For instance, if the imposition of advertising bans on taxis was a binding constraint on taxi behavior, Uber's profit potential would be greater in the cities that imposed advertising bans if Uber allowed its drivers to advertise on their cars. It is also worth noting that there could be a number of reasons why a regulation might not enter into a potential entrant's profit function (one not subject to the regulation). For instance, the regulation might have been binding at some point but became obviated by technology. It might not be enforced. Or perhaps an entrant not subject to the regulation might decide to self-impose it. (Such would be the case if Uber prohibited its drivers from displaying advertising on its cars despite not falling under that regulation.)

On the second point, one might ask whether our strategy assumes too much knowledge of Uber's objectives on the part of the researchers, or whether it is placing too much weight on Uber's decisions being guided by sound business reasoning. We have some evidence that Uber did, in fact, make its entry decisions based on measured demand and interest in a particular city.[13]

[12] This situation has occurred, but it is rare in the United States. Anecdotal evidence suggests that Uber has exited cities where regulatory conditions towards ride-hailing companies changes dramatically after its entry. Although rare, these instances could provide interesting case studies on regulatory effects but are beyond the scope of this paper.

[13] Thanks to Tobias Salz for providing this quotation. It appeared at this url: https://thenextweb.com/news/hitting-ground-takes-launch-uber-hailo-citymapper-new-city

> How does [Uber] go about choosing which cities to target? Jambu Palaniappan, Head of Expansion for EMEA and India, says that sign-ups in cities that don't currently have Uber is one factor. For example, prior to the company's launch in Manchester, UK in May this year, he says that we had thousands of customers that had downloaded our app, signed up, and even tried to request rides, so that's one of the big indicators.

In other words, Uber did not enter cities randomly but rather weighed factors such as reasonable measures of consumer demand when deciding where to enter. And, of course, one likely cause of high demand would be distortions in the taxicab market caused by regulation.

Returning to our profit function, we want to determine the entry condition for Uber. In a discrete time setting, we can think of Uber entering at time t_i under two conditions, that its profits at time $t_i - 1$ would have been negative and its profits at time t_i would be non-negative, *i.e.,* if

$$X_i\beta + \log(t_i - 1) - C_i < 0$$

and

$$X_i\beta + \log(t_i) - C_i \geq 0.$$

This is equivalent to

$$e^{C_i} \leq t_i e^{X_i\beta} \qquad \text{and} \qquad e^{C_i} > (t_i - 1)e^{X_i\beta}.$$

If we make a distributional assumption on the cost term, $e^{C_i} \sim \mathcal{E}(\lambda)$, then the probability of Uber entering city i at time t_i is approximately $\lambda e^{X_i\beta}$ because it is the probability of an exponential random variable "failing" in an interval of width $e^{X_i\beta}$ conditional on having survived through the start of the interval.

One may notice this formulation as familiar from the literature on survival analysis. It is a standard and general model for the waiting time of a process (or survival time) with time-invariant explanatory variables. In fact, if we know the t_is, we can recover estimates of β and characteristics of the distribution of C_i from standard survival analysis.[14] Note

[14] See, for example, Cox and Oakes (1984) or Kalbfleish and Prentice (1980).

that the econometrician does not observe the C_is—these behave as the error term in the model we will estimate. We use the particular distributional assumption mentioned above, although one may choose a different assumption or even proceed semi-parametrically. Our distributional assumption, that $e^{C_i} \sim \mathcal{E}(\lambda)$, corresponds to a Cox proportional hazards model with a time-invariant baseline hazard.

The interpretation of a coefficient in a survival model can be explained with an example. Let there be one explanatory variable, Z, a dummy variable for groups 0 and 1. If $e^{Z\gamma} = 2$ when $Z = 1$, the expected waiting time for group 1 is twice as long as for group 0. So a γ of $\log 2 = 0.6931$ would suggest that members of group 1 survive twice as long in expectation as members of group 0. Our interpretation of the coefficients is more straightforward—it comes directly from our specification of the profit function—but we can also relate them to marginal effects on entry decisions in a similar way as above. Do note, however, that these coefficients will only be identified on a relative basis.

Recall that we want to use data on regulatory environment and patterns of Uber entry to comment on how distortionary the regulations we observed were, or, put another way, on the relative gain to Uber from entry into more regulated markets. The idea is simple: in cities with strict and binding regulations governing the behavior of taxicab operators, a firm offering similar services but not subject to the regulations could exploit a significant advantage over the existing taxicab operators and would, therefore, have an incentive to enter that market earlier than expected. If we found that markets with stricter regulatory environments did not experience earlier than expected entry, we could conclude that evading the regulations did not confer a market or cost advantage or that the regulations were not binding or that the entering firm chose not to exploit the potential advantage. To do this, we estimate profit as a function of city-specific regulations, controlling for other factors.

We estimated a Cox Proportional Hazards model on 250 cities, 161 of which had experienced Uber entry by the end of 2015. The analysis we performed accommodated this truncation so we could use data from all 250 cities, even the ones Uber had not entered, to estimate the parameters of the model. The results are presented in Table 5. We also estimated the models with the regulatory variables computed each of two ways, as desribed in the data section. In the "missing excluded" specifications, we excluded a specific regulation

Table 2: Summary Statistics: Profit Function Estimation

Variable	Mean	St.Dev	Min	Max	Obs.
$UberEntry$	2048.09	358.09	733	2358	250
$Barriers$	0.39	0.26	0	1	213
$Safety$	0.57	0.21	0	1	233
$Operations$	0.62	0.27	0	1	193
$Barriers0$	0.33	0.28	0	1	250
$Safety0$	0.53	0.25	0	1	250
$Operations0$	0.48	0.35	0	1	250
$Subway$	0.04	0.21	0	1	250
Bus	0.70	0.46	0	1	250
$LtRail$	0.11	0.31	0	1	250
$Trolley$	0.07	0.25	0	1	250
$Ferry$	0.06	0.23	0	1	250
$VehPerHouse$	1.53	0.23	0.60	2.00	209
$PercentNoVeh$	12.47	7.16	3.40	55.70	209
$EmploymentNum$	649.94	1398.68	30	11160	173
$AnnualMeanWage$	23071.64	2865.71	17530.00	33920.00	250
$Population$	256944.90	596906.40	6585	8175133	250
$PopDensity$	3176.18	2848.06	171.20	27012.40	250

Note: Only 161 of the 250 cities had experienced entry as of the time of our data collection. For those, $UberEntry$ is coded as the number of days after March 1, 2009 they entered. For the other cities, it is coded as the maximum day in our data set, 2358, for the purposes of this table.

from the calculation of the regulatory index if we were not able to obtain any information on that specific regulation for that city. In the "missing imputed to zero" specifications, we made the assumption that the regulation did not exist if we were unable to obtain information about whether a city imposed a specific regulation.

Of course, in such an analysis, one would want to contol for other important determinants of entry into a market, such as population density, average wage, public transportation infrastructure, and number of vehicles per household. Not surprisingly, these and other demographic and structural variables prove to be quite important and explain most of the variation in entry dates that we observe. Interestingly, we see that Uber entered cities with public transportation earlier, even though Uber would be competing with those alternative modes of transporation. The existence of public transportation is likely highly correlated with unmeasured city characteristics, like traffic congestion and high parking rates, which could account for that result. Alternatively, these results could be consistent with a story in which Uber is a complement to public transportation, not a substitute. Hall, Pallson, and Price (2017) consider this possibility.

Of the three types of regulatory variables, one type, $Safety$, was not significant in any of the specifications we tried. In other words, Uber's entry decisions seem unrelated to the regulations involving either safety of passengers or drivers. We find no evidence that Uber being able to evade these types of regulations had an impact on Uber's entry decisions.

Operations had mixed results. In some specifications, we find that Uber entered markets with high levels of operations regulations somewhat early. Recall that operations regulations included capping the number of vehicles per operator, mandating 24/7 operation, and restricting advertising on the car. One could imagine that not being subject to some operations regulations, like a cap on the number of taxis per operator, could afford Uber some advantage over the legacy services. The index was significant despite the fact that many of the regulations in the index would not be particularly relevant for Uber.

The third category of regulations, $Barriers$, also was a significant determinant in Uber entry decisions in some specifications, with Uber entering cities with high barriers to entry sooner than expected. The hazard ratio indicates that, at any point in time, Uber was three times more likely to enter a city with the highest observed barriers to entry versus the lowest observed barriers, conditional on other city characteristics. Another way to think about the magnitude of these effects is to note that high barriers to entry and high levels of operations regulations both matter about as much in Uber's profits as multiplying population by 2.6. Recall that we classified regulations as "barriers to entry" when they existed primarily to deter or prevent entry. We are not surprised to find that cities that took unusual actions to deter or prevent entry of competitors to their existing taxicab operators would be attractive cities for a ride-hailing service such as Uber to enter.

These results are robust to the inclusion of demographics on vehicles, which cuts down our sample size significantly due to missing data.

Our construction of the regulatory indexes was sensible, we think, but still arbitrary. One might prefer a less constrained specification, where each regulation can enter separately.

With the limited sample size and the large number of possible explanatory variables, however, we are quite constrained and would need some type of model selection procedure. Since LASSO and other similar procedures have not been extended to the maximum likelihood techniques we need to estimate our profit function, we simply perform a LASSO regression with the dependent variable being the number of days after Uber's launch that it entered a particular city. For the cities that Uber had not yet entered at the time of our data collection, we substitute the latest entry date we observe. Explanatory variables are the demographic, structural, and (individual) regulatory variables, along with some nonlinear functions and interactions. The LASSO procedure ranks the explanatory variables by their contribution to the predictive power of the model and reports which variables exceeded the default inclusion threshold. Accordingly, we see these results as more of a descriptive complement to our core results rather than a reliable model selection criterion for our analysis. See Table 3 for the results. Note first that the regulatory variables are almost entirely behind both the demographic variables and those describing transportation infrastructure in terms of predictive power. It is also interesting to note that the second most important among the regulatory variables (behind $24/7Service$) is $DictatedFare$, suggestive of the possible profit opportunities inherent in adaptive pricing, like Uber's surge pricing. Interestingly, $MaxFare$ is much lower on the list, below Stata's default threshold for inclusion, which again is consistent with a firm like Uber wanting to sometimes price below the mandated fare and sometimes price above it.

It is well-known that regulations that pull markets very far away from their counterfactual freely-operating state present opportunities for entry by similar firms that are not subject to the regulations. These results offer some evidence on the distortionary effect of the regulatory environment, and which specific regulations Uber benefited from evading.

We are far from making definitive statements about social welfare and consumer surplus on the basis of results from this analysis.[15] We do think it is interesting to note, though, that the largest distortions seem to be generated by the regulations that we think are least

[15] An interesting study by Farronato and Fradkin (2016) looks at entry of Airbnb into various geographic markets and, using detailed data on hotel prices and occupancy rates, can produce estimates of the welfare effects of Airbnb entry. Their focus is not on regulatory avoidance, but their paper provides a blueprint for thinking about welfare issues in similar markets.

Table 3: Order of Inclusion, LASSO procedure

Included Variables	Variables Not Included
$log(Population)$	$RadioDispatch$
$PopDensity$	$Fingerprint$
$CommuterRail$	$PhysicianCertif$
$AnnualMeanWage$	$MaxWorkDay$
$Subway$	Bus
$LightRail$	$LicenseFee$
$24/7Service$	$MinimumFleet$
$Ferry$	$PercentNoVehicle$
$DictatedFare$	$CabStand$
$Trolley$	$PublicNeed$
$Insignia$	$MaxFare$
	$StartFee$
	$MinAgeDriver$
	$DriverFee$
	$MaxAgeCar$
	$Background$
	$RestrictAdvert$
	$VehPerHouse$
	$MaximumFleet$

Note: The variables on the right are variables below the default cutoff below which the LASSO procedure would not include variables in a model meant to predict Uber entry.

likely to protect consumers and result in social welfare gains. This observation suggests that municipalities might want to rethink the regulatory apparatus surrounding hackney services. Of course it may be that Uber and other ride-hailing services are obviating legacy hackney services quickly enough that the question of the optimal form of their regulation is moot, but this analysis could offer some broader lessons on regulatory capture.

It is reassuring that the safety-related regulations seem to either be non-distortionary—requirements that consumers would demand in a free market in any case—or voluntarily adopted by Uber for other reasons.

Uber was the pioneer and first mover in smartphone-based ride-hailing apps, but it was not the only firm entering US cities with similar services during the time period we study. We do not think that studying the entry behavior of Uber's closest US competitor, Lyft, would provide information nearly as clean as that which we use in our analysis. Strategic

Table 4: Results: Profit Function

	Dep Variable: *UberEntry*							
	Missing Excluded				**Missing Imputed to Zero**			
Exp. Variables	Est Coeff	Std Err	Est Coeff	Std Err	Est Coeff	Std Err	Est Coeff	Std Err
Barriers	1.21 **	0.40	0.89 **	0.38	0.94 **	0.36	0.77 **	0.34
Safety	−0.26	0.54	−0.13	0.51	−0.49	0.44	−0.30	0.40
Operations	−0.04	0.37	0.17	0.37	0.44*	0.26	0.39	0.25
Subway	2.09 **	0.75	1.83 **	0.64	1.65 **	0.60	1.45 **	0.54
Bus	0.46	0.33	0.55*	0.32	0.61 **	0.29	0.80 **	0.28
LightRail	0.40	0.37	0.21	0.36	0.48	0.33	0.23	0.32
CommuterRail	−0.12	0.41	0.27	0.39	0.02	0.37	0.25	0.36
Ferry	0.62	0.38	0.44	0.35	0.63*	0.35	0.50	0.33
Trolley	0.36	0.45	0.16	0.45	0.23	0.42	0.18	0.40
AnnualMeanWage	0.00*	0.00	0.00*	0.00	0.00 **	0.00	0.00*	0.00
log(Population)	0.62 **	0.13	0.68 **	0.13	0.66 **	0.12	0.71 **	0.12
PopDensity	0.00 **	0.00	0.00	0.00	0.00 **	0.00	0.00	0.00
PercentNoVeh	−0.08 **	0.02			−0.06 **	0.02		
Observations	162		179		209		250	
Number of cities entered	121		126		151		161	
χ^2 test statistic	122.79		124.30		166.43		208.86	

Note: Significance at the 10% level is denoted with *, 5% level with **.

concerns would have played a much more important role for Lyft if it was primarily entering cities that Uber had already entered, for example. Furthermore, we were not able to obtain similarly comprehensive entry data for Lyft. We do, though, have incomplete data on Lyft's entry into a subset of cities, which we can use to offer a rough characterization of its entry behavior.

Table 5: First 10 Cities Uber Entered

City	Uber Entry	Lyft Entry	Difference
San Francisco	734	1232	498
New York City	794	1972	1178
Seattle	895	1502	607
Atlanta	907	1642	735
Chicago	936	1530	594
Boston	968	1552	584
Washington D.C.	1020	1617	597
Los Angeles	1103	1431	328
Philadelphia	1191	2158	967
San Diego	1195	1585	390
Average Difference			648

Note: Uber and Lyft entry are represented as the number of days after March 1, 2009 that entry occurred.

First, Lyft tended to enter cities substantially after Uber. See Table 5 for the dates when Lyft entered the first ten US cities that Uber entered. The dates of entry are fairly highly correlated, at 0.34, with Lyft entering, on average, 648 days later. Furthermore, if we consider all cities for which we have entry data for both Uber and Lyft through 2015, the correlation increases to 0.68, Lyft entered, on average, 173 days later, and Lyft only entered 18 out of 52 cities before Uber.

This rough examination suggests that any information about regulatory distortion inherent in Lyft's behavior, conditional on Uber's behavior, would be minimal. An analysis of Uber's and Lyft's entry behavior which accommodates Lyft's strategic considerations could be an interesting topic for further study, though.

5 Conclusion

In the introduction, we noted how ride-hailing services had gained such popularity in such a short time. We presented small pieces of evidence of the effects that they are having on legacy hackney services, such as the changes over time of medallion prices. The core of our paper looked at the regulatory environment in 250 cities and towns in the US and the pattern of Uber entry (and non-entry) into them. Keeping in mind that Uber and other ride-hailing services were not subject to the regulations that existed for the legacy hackney services, we argued that unexpected timing decisions for Uber entry as a function of the regulatory environment could give us information on how distortionary the regulations were. We found evidence that regulations that we categorized as "barriers to entry"—regulations that we felt had little justification other then to make the entry of other providers of hackney services into the market costly or impossible—were associated with unexpectedly early entry of Uber into that market. We also found the regulations governing "operations" were associated with early Uber entry. We found little or no evidence that regulations relating to safety were associated with early or late Uber entry.

It is clear that avoidance of many of these regulations could provide a substantial advantage to ride-hailing services. What is less clear is the overall welfare consequences of this avoidance. The main argument in favor of the regulatory structure is the ability to license and monitor owners and drivers to ensure higher quality (availability, better driving skills, knowledge of local area, honesty, cleanliness, safety, environmentalism, lack of discrimination, etc.). As noted before, the bodies charged with regulation of hackney industries have had uneven records regarding ensuring quality. This observation suggests a regulatory capture story, where commissions now work largely to protect monopoly power and capture rents for the taxicab owners and drivers. While we remain largely agnostic on the question of whether the regulations in the hackney industry are overall welfare-improving, we do presume that that some of them have little scope for improving quality. And it is precisely those least likely to ensure quality and safety for which we have found evidence of their market distortion and their importance for entry decisions. Combining the results from the survival analysis on Uber entry with the response of the values of taxicab medallions to the entry of Uber and Lyft, the evidence suggests that ride-hailing services are exploiting the barriers to entry erected to protect legacy hackney services and are eroding their value as a result.

6 References

Angrist, Joshua D., Sydnee Caldwell, and Jonathan V. Hall (2021), "Uber versus Taxi: A Driver's Eye View," *American Economic Journal: Applied Economics,* 13(3), pp 272-308.

Barrett, Sean (2003), "Regulatory Capture, Property Rights, and Taxi Deregulation: A Case Study," *Economic Affairs,* 23(4), pp 34-40.

Barrios, John, Yael V. Hochberg, and Hanyi Livia Yi (2018) "The Cost of Convenience: Ridesharing and Traffic Fatalities," Working paper, Chicago Booth Research Paper No. 27.

Bresnahan and Reiss (1991), "Entry and Competition in Concentration Markets," *JPE,* 99(5), pp. 977-1009.

Brown, Peter Jensen (2022), "A History of the Taxicab," Early Sports and Pop-Culture History Blog.

Buchholz, Nicholas (2021) "Spatial Equilibrium, Search Frictions, and Efficient Regulation in the Taxi Industry," *The Review of Economic Studies,* 89(2), pp. 556–591.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler (1997), "Labor Supply of New York City Cab Drivers: One Day at a Time," *Quarter Journal of Economics,* 112, pp 407-441.

Chen, M. Keith, Judith Chevalier, Peter Rossi, and Emily Oehlsen (2019), "The Value of Flexible Work: Evidence from Uber Drivers," *Journal of Political Economy,*, 127(6), pp. 2735-2793.

Cox, David, and David Oakes (1984), *Analysis of Survival Data,* Chapman and Hall.

Crawford, Vincent, and Juanjuan Meng (2011), "New York City Cab Drivers Labor Supply Revisited: Reference-Dependent Preference with Rational Expectations Targets for Hours and Income," *American Economic Review,* 101(5), pp. 1912-1932.

Ellison and Ellison (2009), "Tax Sensitivity and Home State Preferences in Internet Purchasing," *AEJ Policy,* 1(2), pp. 53-71.

Farber, Henry (2005), "Is Tomorrow Another Day? The Labor Supply of New York City Cab Drivers," *Journal of Political Economy*, 113, pp 46-82.

Farronato and Fradkin (2022), "The Welfare Effects of Peer Entry: The Case of Airbnb and the Accommodation Industry," *AER,* 112(6), pp. 1782-1817.

Frechette, Guillaume, Alessandro Lizzeri, and Tobias Salz (2019), "Frictions in a Competitive, Regulated Market: Evidence from Taxis," *AER,* 109(8), pp. 2954-2992.

Gilby, Walter (1903), *Early Carriages and Roads,* London: Vinton.

Gittleman, Maury, and Morris Kleiner (2016), "Wage Effects of Unionization and Occupational Licensing Coverage in the United States," *ILR Review,* 69(1), pp. 142–172.

Greenwood, Brad, and Sunil Wattal (2015), "Show Me the Way to Go Home: An Empirical Investigation of Ride Sharing and Alcohol Related Motor Vehicle Homocide," Fox School of Business Research Paper No. 15-054.

Hodges, Graham (2007), *Taxi! A Social History of the New York City Cabdriver,* New York University Press.

Kalbfleisch, John, and Ross Prentice (1980), *The Statistical Analysis of Failure Time Data,* Wiley.

Kleiner, Morris, and Alan Krueger (2010), "The Prevalence and Effects of Occupational Licensing," *British Journal of Industrial Relations,* 48, pp 676-687.

Kyle, Margaret (2007), "Pharmaceutical Price Controls and Entry Strategies," *Review of Economics and Statistics,* 89(1), pp 88-99.

Lynn, Michael, Michael Sturman, Christine Ganley, Elizabeth Adams, and Matthew Douglas (2008), "Consumer Racial Discrimination in Tipping: A Replication and Extension," *Journal of Applied Social Psychology,*, 38(4), pp. 1045-1060.

Moore, Adrian, and Ted Balaker (2006) "Do Economists Reach a Conclusion on Taxi Regulation?," *Econ Journal Watch*, 3, pp 109-132.

Rogers, Brishen (2015) "The Social Costs of Uber," *University of Chicago Law Review Online,* 82(1), pp. 85-102.

Schaller, Bruce (2021), "Can sharing a ride make for less traffic? Evidence from Uber and Lyft and implications for cities," *Transport Policy,* 102, pp. 1-10.

Toner, Jeremy (1992) “Regulation in the Taxi Industry,” Working paper, Institute of Transport Studies, University of Leeds.

Tarduno, Matthew (2021) “The congestion costs of Uber and Lyft,” *Journal of Urban Economics,* 122(C).