The Value of Time in the United States: Estimates from Nationwide Natural Field Experiments

Ariel Goldszmidt, John A. List, Robert D. Metcalfe,
Ian Muir, V. Kerry Smith, and Jenny Wang*

Dec 6, 2020

Abstract

The value of time determines relative prices of goods and services, investments, productivity, economic growth, and measurements of income inequality. Economists in the 1960s began to focus on the value of non-work time, pioneering a deep literature exploring the optimal allocation and value of time. By leveraging key features of these classic time allocation theories, we use a novel approach to estimate the value of time (VOT) via two large-scale natural field experiments with the ridesharing company Lyft. We use random variation in both wait times and prices to estimate a consumer's VOT with a data set of more than 14 million observations across consumers in U.S. cities. We find that the VOT is roughly \$19 per hour (or 75% (100%) of the after-tax mean (median) wage rate) and varies predictably with choice circumstances correlated with the opportunity cost of wait time. Our VOT estimate is larger than what is currently used by the U.S. Government, suggesting that society is under-valuing time improvements and subsequently under-investing public resources in time-saving infrastructure projects and technologies.

Keywords: Value of time; natural field experiment; ridesharing; transportation; benefit-cost analysis.

*Goldszmidt, Muir, & Wang: Lyft; List: University of Chicago, Chief Economist, Lyft, & NBER; Metcalfe: Boston University & NBER; Smith: Arizona State University & NBER.
Acknowledgments: We thank Vittorio Bassi, Antonio Bento, Alec Brandon, Jonathan D. Hall, Nathan Hendren, Justin Holz, Matthew Kahn, Shanjun Li, Paulina Oliva, Joseph Shapiro, Ken Small, and Cliff Winston for excellent comments.

"Remember that time is money. He that can earn ten shillings a day by his labour, and goes abroad, or sits idle one half of that day, tho' he spends but sixpence during his diversion or idleness, ought not to reckon that the only expence; he has really spent or rather thrown away five shillings besides." Benjamin Franklin, 1748

1. Introduction

Perhaps taking the lead from Franklin's advice to a young tradesman, the concept of opportunity cost was leveraged by early economists as far removed as Mill (1848) who laid the foundations for the notion; Bastiat (1848) who cleverly elucidated the broken window fallacy; Austrian economists like von Wieser (1876) who applied it to the phenomenon of cost; and all the way to Green (1894) who made it a central feature of his economic decision makers. Today it would be difficult to find an economist who would not place opportunity cost on a short list of key economic concepts that every citizen should understand. The central actor in Franklin's opportunity cost advice, of course, is the value of time (VOT).

Since these early writers, economists for roughly the next century assumed that the VOT for each person could be regarded as a constant, with their wage rate as a reasonable approximation. For example, college students' investments in human capital should include not only the real resource outlays for schooling but also the value of their forgone earnings from spending time in school rather than working. Similarly, the full cost of on-the-job training includes not only the cost of the training itself but also the value of forgone productivity associated with the employees' time spent in that training. Crucially, this wage assumption makes labor market activity the primary basis for how we should judge the contribution of time to economic welfare.

While focusing on labor market activity yields rich insights that elucidate key economic tradeoffs, there was a movement in the 1960s to also explore the allocation, efficiency, and welfare considerations of non-working time. One of the most influential of these contributions was Becker (1965), who proposed a model in which the individual combines time with a market good to produce a flow of services that he labeled "basic commodities." In this manner, he described households as being small factories at their core. His framework unleashed the full economic toolkit to allow analysis of a wide array of issues within the household. Stimulated by Becker, Mincer, and their students, this new "home economics" began to apply economic concepts to a broad set of issues that included selection of a partner, spacing of children, division of labor among household members, divorce, and decision authority within the family (see Greenwood et al. (2017) for a recent overview).

The growth and maturity of this literature has served economics well, yet one important aspect of the Becker model has received less attention: the key features of his approach can also be leveraged to estimate the VOT. Indeed, as we describe more patiently below, there are two major assumptions in the classic early time allocation studies (Becker, 1965; Johnson, 1966; DeSerpa, 1971) that have allowed the analyst to place a value on time: (1) the degree to which a consumer has the ability to flexibly allocate time to different tasks and (2) the degree of complementarity between time and purchased commodities as part of the consumption process.

The extant body of research on time use falls into two primary categories. The first, associated with microeconomic applications (especially in transportation and infrastructure), has largely relied on observing a person's decisions in the face of time and money trade-offs - either actual trip choices made to reduce travel time delays associated with congestion or hypothetical decisions made in the context of stated choice surveys or small-scale experiments (see, e.g., Deacon and Sonstelie (1985); Smith and Mansfield (1998); DellaVigna et al. (2012)). The second has instead focused on tracking the trends in time use for different demographic groups (Aguiar and Hurst, 2007b; Ramey and Francis, 2009). In these latter studies, the VOT is a latent variable implied by the observed differences in time allocations across these groups. Thus when a VOT is estimated, there are specific assumptions made about how money is traded for time savings. For instance, in Aguiar and Hurst (2007b), cohorts of people with low opportunity cost of time, such as retirees, are found to spend more time searching for lower prices than consumers with more limited time availability. By assuming search time can lead to monetary savings, it becomes possible to infer the VOT (see also Ghez et al. (1975); Juster and Stafford (1991); Robinson and Godbey (1999)).

We take this literature in a new direction by combining Becker's work with the early pioneering work in non-market valuation that explored weak complementarity. Formally, weak complementarity relies on the assumption that a person would not value a quality change in something she does not use (Mäler, 1971). This restriction implies that there is a price equivalent (in welfare terms) for a quality change in that resource (Smith and Banzhaf, 2007). Importantly, the VOT literature has yet to appreciate this theoretical connection, even though the original models sought to establish how the allocation of time contributes to the total value created through each use of time. In a setting where an agent has the ability to make choices that reflect consideration of both the time required and the price of a service, this relationship provides an estimate for the opportunity cost of time. Deeper inspection of Becker's proposed empirical framework for household production highlights another feature of time use that can be leveraged: his "technological coefficients" for the time required for each non-work activity provide the ability to consider how the context of a time allocation affects its value. In particular, one interesting aspect of Becker's model is the proposition that choice characteristics matter, or likewise the fundamental assertion that the properties of the situation and the population might matter a great deal in both the allocation and value of time.

With these necessary theoretical conditions in hand, we sought an appropriate testing ground to empirically study the VOT. Our search concluded with the realization that the assumption in the classic studies applies to any situation where one must wait for a service or good. An activity that nearly half of all American adults have used satisfies exactly that assumption: rideshare. Waiting sessions on rideshare naturally provide the basis for estimating the trade-offs between waiting time and price that underlie economic measures for the VOT. To our best knowledge, our paper is the first to recognize the applicability of weak complementarity to situations where one must wait for a good or service. Furthermore, the richness of the rideshare environment permits a deeper exploration into measuring important context specificity (Becker's "technological coefficients") across relevant situations and individuals.

To request a ride via a rideshare service, a prospective passenger opens the app on their phone, can input an intended destination to receive a price quote and estimated wait time, and then decides whether to request (purchase) the ride. Such "take it or leave it" decisions exploit the role of weak complementarity in extensive margin choices. More specifically, when an individual opens the Lyft app we are able to observe how various combinations of wait times and prices affect purchase decisions. In 2015/16 and 2017, Lyft ran two natural field experiments that randomly varied these prices and wait times that were shown to customers in the app as they made their purchase decision. The assumption of weak complementarity ensures that the observed trade-offs then allow us to estimate price changes that are equivalent to a change in wait times. Thus, with weak complementarity in place, the observed trade-off reveals the marginal VOT.

The field experiments spanned 13 cities in the United States that had constituted most of Lyft's largest markets at the time. In total, they cover 3.7 million customers and 14.8 million customer sessions (defined as an interaction with the app in which the customer receives a price and time quote). The exogenous prices and wait times in our field experiments are independent of the passengers' outside options as well as potential unobserved shifters to aggregate demand (i.e., passengers) and supply (i.e., drivers). The exogenous variation combined with passengers' decisions and weak complementarity allow us to estimate the marginal VOT.[1]

An important feature of our work is that Lyft retains the ability to control not only whether people have to wait longer but also how much longer they have to wait. In the experiments, customers are randomly assigned to wait at the market wait time or at least an additional 60, 150, or 240 seconds over the market wait time (additional wait time irrespective of market conditions). With this variation, we can causally estimate how the waiting time elasticity and the VOT vary over the length of the wait time - with variation in length that is unrelated to other market factors (e.g. when and where customers open their app). Our approach allows us to recover an estimate of the VOT over wait time gradients, providing insight into the shape of the VOT function.

[1] Our research is complemented by two contemporaneous studies that use very different identification strategies and data sets (Castillo, 2019; Buchholz et al., 2020). Both studies observe market wait times and prices but use econometric structure as opposed to experimental variation to solve the identification problem for Houston and Prague consumers. Our experimental structure allows for exogenous variation in waiting time and prices, and provides greater resolution in the nature and spatial variation in VOT estimates, and how market conditions experienced by rideshare users (e.g., weather conditions, location (airport, downtown, public transit distance), business trips etc.) affect VOT. In addition, weak complementarity assures that we avoid sensitivity of the results to the specification decisions that generally are associated with a structural approach.

Several insights can be drawn from these two large-scale natural field experiments. We divide them into three main areas. First, we find that consumers are responsive to both wait times and prices. Across both field experiments, we find the time elasticity of demand to be approximately -0.043 (standard error of 0.003) and the price elasticity of demand to be -0.59 (standard error of 0.02). Price elasticities are significantly larger than time elasticities within every city. The estimated price equivalent implied by these two elasticities yields an average VOT across the U.S. of $19.38 per hour (standard error of 1.39; all prices are in 2015 dollars). To explore how our results map across relevant populations, we re-weight our estimates for the elasticities matching the rideshare population with the broader US population. We find no meaningful difference in VOT estimates based on the weighted sample, tentatively supporting their external validity with respect to the broader population of all travelers. We also find that our VOT estimate is stable across the duration of our eight-week experiments.

When evaluating projects, the US Government currently values people's time between 33% and 50% of the wage rate for project appraisal, suggesting a value of at most $14.20 per hour.[2] This one figure is meant to cover all types of travel (e.g., leisure, personal, and commuting to work), except where driving is part of the job (e.g., freight travel where the wage rate is used to impute the VOT) (USDOT, 2015). Our VOT estimate is approximately 35% higher than the currently used rule of thumb by U.S. federal guidelines. More specifically, our estimates imply that the VOTs for different metro areas are approximately 75% of the after-tax mean wage rate and about 100% of the median after-tax wage rate. In every metro region, we find a VOT estimate that is statistically larger than $\frac{1}{2}$ of the after-tax mean wage rate.

A second set of results that emerges from our field experiments relates to waiting time elasticities and the shape of VOT. We find that both are larger over longer periods of wait time, independent of market conditions and the reasons for consumption. This finding implies that the VOT is convex over time. Such convexity is important when considering the appropriateness of transferring or generalizing VOT estimates across space, time, or situations, as we find that the VOT depends critically on the baseline level of time use. Taken together, our first two results are directly relevant for the analysis of both private and public investments. They suggest that society is under-valuing projects that involve time saving infrastructure or technologies and furthermore, the degree of this under-investment increases in the amount of time saved.

Our third set of results leverages non-experimental variation to explore how properties of the situation affect the VOT estimate. Using a simple framework that directs our exploration of how various choice characteristics affect the VOT, we find substantial heterogeneity across contexts.

[2] The recreational demand model literature and environmental regulation use $\frac{1}{3}$ (following Cesario (1976)), and the transportation and infrastructure literature and regulation use $\frac{1}{2}$ (following Small et al. (2005); Small (2013)) of the wage rate. There does not seem to be consistency within the government on the values used for different policies. At best, one might argue that the U.S. Department of Transport is valuing time primarily as it relates to congestion–infrastructure to address congestion related delays whereas recreation is considered leisure travel, therefore the opportunity cost may be argued to be lower since travel may be part of the trip experience.

For example, across regions we find a Spearman correlation of 0.73 ($p = 0.025$) between the mean wage and the VOT estimate. Furthermore, we find that the VOT critically relates to the availability of substitutes for consumers, in that individuals considering a Lyft trip near alternative modes of transportation are more time sensitive than those who do not have readily available substitutes. In addition, signatures of the trip matter a great deal, as the VOT is strongly related to purpose of trip: during the morning and afternoon peak commuting times, the VOT is 50% higher than during off-peak times.[3] Relatedly, weekdays have a 10% higher VOT than weekends, and we find that the VOT is 20% higher in the central business districts of cities than in the suburbs. Finally, we explore several other types of heterogeneity based on our economic framework to find time elasticities consistent with predictions from economic theory and the broader literature (Small, 2013).

We view our research as contributing to several areas of import. For policymakers, our estimates are a significant improvement over those in the existing literature with respect to both the identification and the design of the wait and travel time changes. Our experiment varied the total travel time of the journey, so we are assuming the increase in wait time for the ride is valued in the same way as an increase in the in-car time.[4] They also provide a granular view of how choices vary with the context of time use, consistent with Becker's early insights. Our research also has direct policy implications, as we recommend that policies: (i) account for the VOT heterogeneity with respect to cities, locations within cities, day of week, and time of day when estimating the benefit profile of public projects; and (ii) when this is not possible, adjust the rule-of-thumb VOT estimates up to 75% of the after-tax mean wage rate otherwise.

In this spirit, our estimated VOT varies predictably with economic aspects of the marketplace. This finding has important implications for how we value time in various economic sub-fields. For instance, in transportation, the VOT is usually the pivotal factor in benefit-cost decisions, as it has been estimated that excess urban road congestion led U.S. consumers to spend 5.5 billion hours sitting in traffic. Indeed, some studies (Schrank et al., 2012; Couture et al., 2018) estimate the annual deadweight loss due to congestion in the United States alone at $30 billion. Understanding how best to apply our estimated elasticities to construct Pigouvian taxes designed to reduce the deadweight loss of congestion is an important next step in such research (Arnott et al., 1993; Duranton and Turner, 2011; Finkelstein, 2009; Small, 2013). Furthermore, the VOT in

[3]We acknowledge that the context of the wait in our field experiments is heterogeneous. Some consumers will be waiting for the car on a street corner, while others will be getting ready to leave their building, others working, etc. Given that some consumers can do other things while they wait for the car (e.g., peruse their emails, texts, etc), the wait might not be necessarily boring or painful, so our estimates might be viewed as a lower bound of the VOT that may be measured in less comfortable or productive situations (i.e., caught in gridlock traffic).

[4]Once you assume a linear time constraint (i.e., total hours = sum of allocations) in the model, we assume each part being allocated is a perfect substitute for another. Because wait time exhibits weak complementarity with the Lyft ride, we now have the VOT derived at the margin by the price equivalent change in the price of the ride. Linearity of the time constraint allows us to use this margin and apply it to other types of time at the margin, so it becomes a general estimate of the VOT. Empirically, to test this, we need experimental random variation in price and time for people who are randomly allocated to either wait time or in-car time (where the base price and total time are the same in the two scenarios). Such a study does not exist.

car transit is an important parameter when estimating the demand for public transit, evaluating any proposal for federal funding of infrastructure, and evaluating the impact of other more climate-friendly (green) transit options, such as a carbon tax (Parry and Small, 2009; Chen and Whalley, 2012; Anderson, 2014; Basso and Silva, 2014). More generally, given that the VOT usually constitutes the largest share of total benefits in infrastructure projects, our estimates open up the possibility of more efficient allocation of resources in the economy.[5]

In terms of linking to the broader literature on a host of policy decisions, many view time as the ultimate scarce resource. These are many policy- and market-relevant activities that require a period of waiting time before ascertaining utility from the commodity, such as time waiting for: a table at a restaurant; a delivery of a good; a store or government office to open (e.g., renewing a driving license or ID, or waiting to be seen at a Veterans Administration center); a dentist or a doctor; a voting booth to open; accessibility to buildings (e.g., how much longer does the handicapped-accessible ramp/elevator/parking/office take); or a car or public transit to get to a destination. For the economy as a whole, the VOT depends on how different people respond to the market and non-market signals in allocating their monetary resources and time (Juster and Stafford, 1991). These allocation decisions of time impact where people live (Wheaton, 1977; Van Ommeren and Fosgerau, 2009; Su, 2018; Kreindler and Miyauchi, 2019), how they supply their labor (Aguiar and Hurst, 2007b; Aguiar et al., 2013, 2017; Benhabib et al., 1991; Gelber and Mitchell, 2012; Goldin, 2014; Gronau, 1973; Mas and Pallais, 2017, 2019), how they commute (Small et al., 2005; Bento et al., 2017; Hall, 2020), how they invest in their health (Besley et al., 1999; Miller and Urdinola, 2010; Philipson et al., 2010), and what goods they buy (Nevo and Wong, 2015). VOT estimates have also become increasingly important in international debates about productivity and national accounting (Krueger et al., 2009; Nordhaus, 2009; Aguiar and Hurst, 2016), for the welfare estimation of business cycles (Aguiar et al., 2013), and the VOT is a central feature in governments and companies as a basis for investment decisions for the supply of intangible and service goods within economies.

The remainder of our study is structured as follows. Section 2 provides key theoretical underpinnings of the classic literature and outlines how we leverage these features to estimate the VOT using two field experiments. Sections 3 and 4 report empirical results from our two natural field experiments and detail our framework for exploring heterogeneity. Section 5 addresses the issues in identification and external validity of our estimates, and section 6 provides a discussion linking our work to the current policy landscape. Section 7 concludes. The online appendix includes additional empirical analysis.

[5] Moreover, our VOT estimates relate to understanding the economics of online platforms (Goolsbee and Klenow, 2006; Chen et al., 2014; Allcott et al., 2019) and the amount of bureaucracy in government policymaking (Sunstein, 2018).

2. Theory and Market Context

In this section, we describe the theoretical landscape for valuing time and link it to our modeling approach. What results is a set of necessary experimental conditions that must hold to deliver theoretically-consistent estimates of the VOT. We then detail the market context for our two natural field experiments.

2.1 Theoretical Framework

The current modelling of time use and value in the literature stems from Becker's (1965) insight that time is required for all consumption activities. Most discussions of his contribution equate it with the origin of home production, and focus on Becker's argument that an individual "produces"—what Becker describes as "basic commodities"—which are then consumed.[6] These basic commodities are the services derived when market goods are combined with time, and these services are what contribute to well-being and motivate choices for individuals. While the concept of home production is certainly important, Becker's description of consumption also introduced two other features that are important to our research design. The first feature highlights the role of restrictions on how private goods and time enter preferences. This dimension of the classic framework is best illustrated by first considering a simple form of Becker's model. Assume the household consumes two basic commodities, $Z_i, i = 1, 2$, which are service flows, and in our case, Z_1 is the service flow from rideshare travel, and Z_2 is the service flow from all other goods. The household production functions are Leontief as in equation (1):

$$T_1 = t_1.Z_1 \qquad T_2 = t_2.Z_2 \qquad x_1 = a_1.Z_1 \qquad x_2 = a_2.Z_2 \tag{1}$$

Our specification assumes one private good x_i per basic commodity. In our case, x_1 is the rideshare trip, and x_2 are all other goods. Time is allocated exclusively to each activity and there is no multi-tasking. Both a_i and t_i are technological coefficients in Becker's model. t_i is the time in discrete units for each unit of Z_i, and a_1 is the amount of rideshare service that is needed to produce Z_1.[7]

Assume a time constraint with $\overline{T}$ the total time available and T_w the amount of work time which is priced at w. The individual also faces a budget constraint with p_i, where $i = 1, 2$, the prices for private goods, the wage income, wT_w, and non-wage exogenous income (R). Becker's

[6]There are a number of contributions using the household production logic to model consumption expenditures as well as in describing alternatives to the unitary model of individual behavior. A good access point is the review in Browning et al. (2014), where the work using household production in alternative models of individual choice is summarized.

[7]a_1 is important because it allows us to interpret a local marginal condition that links the Becker model and weak complementarity (see below). Thus a_1 is a function of the waiting time, which allows us to illustrate if consumers are not producing" Z_1 (i.e., services from travel with ride share), they do not care about waiting time.

household production for consumption activities is equivalent to a restriction on preferences that treats time and private goods as perfect complements.

The second feature highlighted by Becker's model arises from the fact that time use is described to be *specific to a particular consumption task*. That is, the framework allows time to be uniquely linked to each activity a person undertakes. As a result, a natural interpretation is that the analysis can consider how the context for using one's time affects its value. For instance, time spent commuting on a weekday morning can be different from commuting in the evening or over weekends.

The importance of restrictions on how time and goods enter preferences, and the assumption an amount of time is uniquely linked to the use of each good (rather than allowing time to jointly produce two or more basic commodities), arises when we consider the two possible values for time. These different values are implied by the indirect utility function for Becker's model. Substituting the time constraint and production functions into the budget constraint, we have the general form given by equation (2) (see appendix A for all equation derivations).

$$\overline{V} = V(wt_1 + p_1.a_1, wt_2 + p_2.a_2, w\overline{T} + R) \tag{2}$$

The first possible value of time is defined when we consider increasing or decreasing the time endowment $\overline{T}$. Such a shift in the time endowment implies the marginal value of time is equal to w. Such an approach is usually operationalized using stated preference surveys or by increasing the amount of time available for people to make choices (e.g., sleeping less during each day which is unlikely without a technology).[8] The second possible value of time arises because we can define what might be termed the "supply" price or time cost of each activity, which depends on both the wage and the technology of home production (i.e. the t_i's). In this case, the supply price is equal to wt_1 for the first consumption activity and wt_2 for the second.

Since the developments in the literature in the 1960s, the contributions on the allocation and value of time have, for the most part, missed these key features of Becker's model. They observed that the model had the same implications for VOT because the marginal value of adding to the time endowment remained the wage rate. This result is conditional on the optimal allocation of

[8] Small et al. (2005) impose a linearity assumption on the indirect utility function assumed to underlie an individual's choice of whether or not to use an express lane for a trip. The express lane has a toll and an anticipated travel time while the conventional lane has no toll but a longer anticipated travel time. Linearity assures the ratio of the coefficients for the toll and the travel time in the choice model reveal a marginal value of time (see their equations (1) and (2)).

time among activities with different costs.[9] Our conclusion is readily illustrated if we consider how the model describes adjustment in response to a change in the marginal value of time (w). As equation (3) illustrates, the supply of labor takes into account the reallocation of time among activities:

$$\frac{\overline{V}_w}{\overline{V}_R} = \overline{T} - t_1 Z_1^* - t_2 Z_2^*, \tag{3}$$

Equation 3 is derived from the partial derivative of equation 2 using Roy's identity as amended for this model, where Z_1^* and Z_2^* are the utility maximizing choices for the two basic commodities (rideshare and all other commodities). Changes in the wage rate change the "prices" of each of the basic commodities as illustrated by the last two terms on the right side of equation (3). Adjustment in the amounts consumed determine the time requirements (with the assumed Leontief technology) and labor supply is the residual component of the time endowment.

To focus attention and simplify matters, Becker assumed the time requirements for each activity were fixed. The joint roles for preference restrictions between goods and time and the ability to take account of the context of when and how time is being used are the important elements in his framework for our research. Of course, demonstrating this point requires an ability to control both the time requirements for some set of activities and their prices. Such control allows the analysis to test whether time allocation is largely a matter of labor–leisure decisions. In Becker's model, an agent's choice of her allocation of non-work time matters according to equation (3). When the assumption of a Leontief technology for household production is relaxed, however, control over the prices and time requirements alone will not ensure recovery of the value of time for each use.[10] Another preference restriction is needed.

Fortunately, for some activities a different form of complementarity, weak complementarity, provides sufficient information to value time. These are activities that require a period of waiting time before ascertaining utility from the commodity, such as waiting for: a table at a restaurant; a delivery of a good; a store to open; a dentist; a doctor; a voting booth to open; or a car or public transit to get to a destination. As Mäler (1971, 1974) showed, weak complementarity means that underlying changes in features of the good or service are only important to actual consumers of the good or service. He recognized that even without the assumption of perfect

[9] Aguiar and Hurst (2007a) develop their approach for estimating a value of time by assuming that optimizing households exploit a shopping technology as another mechanism for substituting time for goods outside the labor market. They maintain that the price paid for private goods is a function of the time allocated to shopping. Greater search time yields lower prices. They also acknowledge that the price paid can depend on shopping needs or the number of items that might be involved in a household's search activities. So the value of time is derived as the shadow price of time allocated between the shopping and household production technologies (see their equations (1) and (2)). The ability to freely substitute time between these two uses assures the marginal value of time is equalized between these activities "outside" the labor market. As a result they use estimates of this price function to estimate the value of time (see their figure 1).

[10] When the household production technology is assumed to be more flexible, the VOT in each use depends on the marginal technical rate of substitution between time and goods at the optimal consumption levels for the basic commodities. To estimate this requires detailed information on the technologies involved as well as all the goods' prices.

complementarity between a nonmarket good and a private good, weak complementarity makes it possible to estimate the demand price for the nonmarket good from the information contained in the demand function for the private good.[11]

Our case provides a direct parallel to his example. If the demand for the rideshare service is zero, then the demand for a shorter wait time to obtain that service is also zero. Mäler was careful to spell out how the assumption of weak complementarity applies only at the individual level (or to an aggregate over homogeneous groups of consumers). He also noted that to use it in measuring the willingness to pay for a discrete change in a nonmarket good (like time), the specification of the Marshallian demand for the private good needs to ensure that the corresponding Hicksian demand function for that good has a finite choke price.

For our purposes, however, we are not attempting to estimate the demand for rideshare trips; we are simply interested in the marginal VOT. Our use of weak complementarity can be illustrated with a simple amendment to the Becker model. We acknowledge at the outset that our example abstracts from important details that could be used in a full structural model of these decisions.

We remain with the assumption that the first basic commodity is the travel services "produced" by using a rideshare company. We distinguish two types of time involved in ride share services. The first, designated T_1, is a summary measure for the travel time associated with rideshare trips. The second type of time, T_1^a, is exogenous from the perspective of the individual. It represents a second summary measure for the waiting time for the rideshare trips while the individual waits for the drivers involved in these trips to arrive. The user knows these times on each trip occasion when selecting her trips. We assume waiting time is an indicator of the quality of the service. Longer waiting time implies a lower quality travel service can be produced. Weak complementarity implies that the individual does not value waiting time if she does not produce travel services. Thus, our modification to the Becker model must embed this assumption. For this example assume the technical coefficient linking Z_1 (our measure of the produced travel services) to x_1 (our measure of the rideshare services required) is a function of waiting time. So we replace

[11] For certain environmental applications this condition can be controversial—a person may want to maintain high levels of air quality (and visibility) at the Grand Canyon but never plan to visit the site, or want to protect the Arctic Wildlife Preserve but not consider a wilderness adventure there. These omitted values are the existence values of such goods. As he noted:

> *It is, however, not necessary that the environmental service and the private good be perfect complements in order to carry out the steps involved in determining demand price. A much weaker condition of complementarity is the following: if the demand for a private good is zero, then the demand for some environmental service will also be zero. If, for example, the private good is swimming in the lake and the environmental service is the quality of that lake, then it is very reasonable to assume that if a person does not use this lake for recreation, he is indifferent to the quality of the water. . . . It would therefore seem that this weak complementarity condition has very broad applications, although it cannot be applied in cases where option values are involved* (Mäler 1974, 183).

this component of the Becker model with:

$$x_1 = a_1(T_1^a)Z_1 \tag{4}$$

We assume $a_1' > 0$, implying increased waiting times reduces the amount of "constant quality" travel services produced by ride share (i.e., Z_1 gets smaller) This specification assures that if $Z_1 = 0$, then changes in T_1^a will not affect the individual's well-being, consistent with weak complementarity between waiting time and the travel service basic commodity.[12]

We can now return to equation (2) and adjust the indirect utility implied by Becker's model to reflect the assumed role of T_1^a in producing constant quality travel services. Equation (5) provides the modified indirect utility:

$$\bar{V} = V(wt_1 + p_1a_1(T_1^a), wt_2 + p_2a_2, w(\bar{T} - T_1^a) + R) \tag{5}$$

When we make this change and use duality to consider the marginal value of reducing waiting time, equation (6) results:

$$\frac{V_{T_1^a}}{V_R} = \frac{V_1}{V_R}p_1a_1' - w \tag{6}$$

The terms on the right-hand side of the equation corresponds to the marginal "cost" to the user of increases in the waiting time. We can see this by noting that Roy's identity implies $-V_1/V_R = Z_1$. Our argument suggests that the price equivalent value of waiting time is $(p_1Z_1a_1')$, and that could be equal to, less than, or greater than the wage rate. This is an empirical question. Shorter (longer) waiting time reduces (increases) the incremental cost of producing travel services with the ride share, Z_1. This intuition leads to Equation (7):

$$\pi = -\frac{V_{T_1^a}}{V_R} = p_1Z_1a_1' + w \tag{7}$$

Of course, this case represents *one* simple example. Our general point is to show that the example "works" because of weak complementarity. From equation (5), the specification of the measure for ride share trips (x_1) implies that there is an equivalent change in p_1 that can be represented by a change in T_1^a.

Our main theoretical contribution to the Becker model is that a weaker preference restriction

[12] Several qualifications apply to this stylized example. If we were attempting to model both the decisions to use rideshare and the number of trips to take, our specification would need to reflect that T_1^a and a measure for the count of trips are related. Depending on how we added these details, connecting the model to measures of rideshare services, such as trips, the time used in traveling, and the time spent waiting, nonlinearity could be introduced into the budget constraint. Such a detailed formulation would rely on the specific functional assumptions made in identifying estimates for the value of time. Since our goal is simply to illustrate how weak complementarity allows us to measure the marginal value of time by implying a welfare equivalent link between price changes and wait time changes, we avoid spelling out these connections.

than was required in Becker's model is possible for valuing time. In the context of rideshare, we have this form of weak complementarity as people wait for the ride. An additional important feature of the model is the specificity of the value of time can be linked to its particular use. This arises in rideshare because the timing and position of rideshare trips can be linked to a wide range of purposes, tasks, and contexts, allowing for an assessment of how the properties of the situation affect VOT measures. Taken together, we now have the theoretical machinery to value time through changes in prices and wait times of the good.

In this spirit, leveraging the Lyft rideshare platform, our model is specific to a trip session, in which a passenger opens up the app and receives a price and waiting time quote for a potential trip. Individual responses to the terms of a ride are indexed by passenger (i) and session (j). We assume that incomes and other prices faced by individuals opening the Lyft application are fixed across the experimental groups due to randomization, and focus our attention on the utility realized with and without requesting a ride. Equation (8) begins the process of formalizing the decision associated with requesting a ride:

$$V_{ij} = v(P_{ij}, T^a_{ij}) + \varepsilon_{ij} \tag{8}$$

Let V_{ij} be the utility associated with selecting a Lyft ride at a price of P_{ij} for individual i and session j. Let T_{ij} be the wait time indicated to individual i in session j, and ε_{ij} be a random error capturing unobserved (to the analyst) features of the circumstances of choice.

We assume that each individual compares the realized utility from a Lyft ride with that of a default condition that we do not observe. We assume the default option is specific to each individual and session and provides utility W_{ij}. A Lyft is requested if and only if $V_{ij} > W_{ij}$. By independently randomizing both P and T^a, we can recover estimates for the parameters used to describe the choice process in equation (8).

The price of a Lyft ride is the product of a base price, B_{ij}, that depends on the characteristics of the request (timing, route, and vehicle type) and a price multiplier (called Prime Time) $(1+PT_{ij})$. This multiplier is set dynamically in response to local demand and supply conditions. B_{ij} and PT_{ij} are indexed by i and j because the records of sessions allow the timing of the session and the individual to be distinguished.

In our case, we observe whether an individual selected a Lyft ride but cannot completely characterize the features of the alternative set when the ride is not chosen. This is where our assumption of weak complementarity provides the "traction" needed to recover a VOT. As noted earlier, when waiting time is a weak complement to the rideshare service associated with each trip, a change in wait time is equivalent, from a welfare perspective, to a change in the price of the trip (Smith and Banzhaf, 2007). The important implication is that we do not need to know anything about a passenger's labor supply decisions to recover an estimate for the opportunity cost for their time; the price of the Lyft ride serves this role. Thus, with appropriate exogenous variation in both trip prices and wait times, consumers' actual choices allow us to identify the key threshold trade-offs between time and money that serve to recover the VOT.[13]

The decision process in our model begins with the assumption that each individual is considering a "local" trip, that is within his or her metropolitan area. We observe everyone who opens the Lyft application during our experimental period. Yet, we do not know their complete set of outside options, and because of this limitation we consider a variety of approaches to organize sessions to account for our hypothesized differences in how these outside alternatives influence W_{ij} for each individual. There are several implications of this constraint on what can be observed. The first of these is the selection effect directly associated with knowing only those who open the Lyft app; we discuss the implications of this external validity issue in Appendix Section H. Moreover, because we cannot characterize W_{ij}, we do not have a reference or baseline condition that we would expect in defining a reference utility level to measure a VOT. This limitation influences our interpretation of the estimates. Without specifying the outside alternative, we estimate choice probabilities relative to a normalizing alternative.[14]

2.2 Empirical Model

Our primary model specification for the choice process is contained in equation (8), which compares V_{ij} with W_{ij} using the log transformation for the price and wait time as determinants of the observed trip request, $R_{ij}(0,1)$. We assume that passenger i in their j^{th} session receives greater utility from requesting a ride, V_{ij}, than the alternative W_{ij}, and thus the request, R_{ij}, is made according to:

$$R_{ij} = \beta_1 \ln P_{ij} + \beta_2 \ln T^a_{ij} + \varepsilon_{ij} = \beta_0 \ln B_{ij} + \beta_1 \ln(1 + PT_{ij}) + \beta_2 \ln T^a_{ij} + \varepsilon_{ij} \tag{9}$$

Here P_{ij} is the offered price, T^a_{ij} the offered wait time for the ride, B_{ij} is the base price, and ε_{ij} is the unobserved error term. The price of a ride is the product of a base price B_{ij} and a price multiplier, $(1+\mathrm{PT}_{ij})$, which is set dynamically in response to local supply and demand conditions. By using the log transformation for the price and wait time in equation (9), the choices reveal the preference parameters needed to recover our measure for the price equivalent of waiting time without knowing the base price for each request.[15] This formulation allows a separation of the base price and the price multiplier as determinants of the request, and identification of the price

[13] We do not need to specify a global model of time use with every commodity and price in it. Weak complementarity with random prices and wait time at the point of purchase is enough.

[14] In general, we do not know if a person actually considered the alternatives that were available, and only know that they were feasible when a decision to select a mode was made. As McFadden (1974) demonstrated, given the specification of the factors influencing an individual's choices and an assumed choice set, the inability to know a specific default alternative or all the possibilities does not prevent one from assessing the relative importance of each determinant using a random sample of the hypothesized alternatives, together with assumptions that characterize the choice process. For many applications, this constraint on the information available is not important to the results.

[15] We considered other functional forms as robustness checks; see Table C.23 in Appendix C. We also discuss this in section 3.4.1.

effect through the experimental variation in the price multiplier. Since we observe the base price only for a self-selected sub-sample of sessions (as described below in the first field experiment), we include this effect as one of the components of the model's error.[16]

Our *price equivalent of a unit change in the wait time* is defined by the marginal rate at which the passenger is indifferent in trading off units of waiting time with units of monetary cost for the trip. For our basic model with time and price in logs, as in (9), this rate is (minus) the marginal rate of substitution of T^a_{ij} and P_{ij}:

$$-\frac{dP_{ij}}{dT_{ij}} = \frac{\partial U_{ij}/\partial T^a_{ij}}{\partial U_{ij}/\partial P_{ij}} = \frac{\beta_1}{\beta_2}\frac{P_{ij}}{T^a_{ij}} \tag{10}$$

The value for this price equivalent depends on the values of P_{ij} and T^a_{ij} as well as the values of β_1 and β_2.

Given that ε_{ij} in equation (9) includes the default alternative for each person opening the Lyft app, as well as many unobserved factors that will simultaneously affect both the utility of rides and that of the default alternatives, it is reasonable to expect that it will not be exogenous to wait time and prices. We address this limitation in two ways. First, we use experimental variation to construct instruments for the price multiplier and waiting time terms and use two-stage least squares to estimate our model. As a result, consistent estimation of β_1, β_2 can be realized with instruments that provide exogenous variation in $1+\mathrm{PT}_{ij}$ and T^a_{ij} that are independent of W_{ij}, B_{ij}, and ε_{ij}. The power of using a field experiment with randomization of both prices and wait time is that we do not need to make any further assumptions about customer behavior. Second, our large sample allows the definition of a variety of sub-samples that identify different circumstances in which the Lyft app is opened. Some of these outside factors can also be expected to affect the (baseline) ride prices and waiting times.

Our model is therefore estimated with two-stage least squares (2SLS), with the first-stage equations as:

$$\ln(\mathrm{ETA}_{ij}) = \gamma_0 + \gamma_1\mathbf{T1}_{ij} + \gamma_2\mathbf{T2}_{ij} + \gamma_3\mathbf{T3}_{ij} + \gamma_4\mathbf{T4}_{ij} + \gamma_5\mathbf{T5}_{ij} + (controls) + \eta_{ij} \tag{11}$$

[16] In early discussions of the random utility model, the framework was used to describe a choice among a discrete set of alternatives. As a rule, when there were more than two possibilities, a logit framework was often adopted. More recently, these estimators have been generalized to allow for unobserved heterogeneity in preferences by specifying some coefficients as random variables and relying on mixed logit estimators. Mixed logit avoids the restrictive assumption of independence of irrelevant alternatives (IIA) with a simple logit approach. When used for measuring willingness to pay the distributions for the coefficient for the price of alternatives need to be restricted to assure consistent welfare measures. When the alternatives are limited to two possibilities (taking a Lyft trip or not), and one is not attempting to account for a set of observed characteristics distinguishing the alternatives, then linear regression methods are often used. Ordinary least squares (OLS) and two-stage least squares (2SLS) provide robust strategies for estimating the parameters needed to recover the marginal VOT and evaluate the sensitivity of the results to the circumstances characterizing the context of the choice. We report in Appendix C the sensitivity of our conclusions to alternative estimators in Table C.24

$$\ln(1+\mathrm{PT}_{ij}) = \delta_0 + \delta_1 \mathbf{T1}_{it} + \delta_2 \mathbf{T2}_{it} + \delta_3 \mathbf{T3}_{ij} + \delta_4 \mathbf{T4}_{ij} + \delta_5 \mathbf{T5}_{ij} + (controls) + \zeta_{ij} \tag{12}$$

and the second-stage equation is:

$$\mathrm{Request}_{ij} = \beta_0 + \beta_1 \ln(1+\mathrm{PT}_{ij}) + \beta_2 \ln(\mathrm{ETA}_{ij}) + (controls) + \varepsilon_{ij}. \tag{13}$$

Here T1 through T5 are dummy indicators of the experimental treatment assignments (described below), and *controls* include a vector of fixed effects for passenger and session types (e.g., a user's number of past rides with Lyft) as well as time and location, controlling for the unobserved variation in W_{ij} and $\ln B_{ij}$.

The estimated β_1 and β_2, as well as assumed values for P_{ij} and ETA_{ij}, allow recovery of the price equivalent of a unit of waiting time. We evaluate our price equivalent at the control average waiting time and control average price on completed rides during our experimental period. The incremental price equivalent is estimated by:

$$\frac{\hat{\beta}_1}{\hat{\beta}_2}\frac{\bar{P}}{\bar{T^a}},$$

where $\bar{P}$ and $\bar{T^a}$ are the average price and waiting time for control units.[17]

In Section 3 below, we fully describe our two field experiments to identify how agents trade off units of waiting time with units of monetary cost. Our first field experiment randomly assigns Lyft users to a control or one of five treatments (T1 through T5 in our 2SLS model) varying (increase, decrease, or market) price and/or wait time (increase or market). Each individual remains with the assignment they are given at the time of their first opening of the Lyft application during our experimental time period (eight weeks). The realized values for both wait time and price depend upon the circumstances of the session; the Prime Time multiplier and waiting time are thus endogenous variables to the request decision. We use the features defining our treatments to define instruments for both variables. We also add additional controls that identify the features of the context in which a person opens the app. We use the 2SLS approach above to estimate our choice equations, but also report alternative estimates in Appendix Section C.

A second, complementary, field experiment—in which only ETA increases are randomized (while prices remain at market values)—explores more fully the circumstances of choice and how they allow us to measure heterogeneity. In particular, extra wait time is varied across location-time blocks (as opposed to across users) which provides (i) a robustness check on the results from our first field experiment and (ii) additional insights into how users' time elasticities may vary

[17]The decision to evaluate the VOT at the control average price and waiting time is somewhat arbitrary, and due to the nonlinearity of the expression for the VOT in price and waiting time, the VOT expression evaluated at the average price and waiting time may differ from the average VOT. We address this concern in Appendix L by constructing a VOT estimate for each observation using observation-specific price and waiting time predictions (and semi-elasticities); this process produces a full distribution of VOTs across sessions rather than a single estimate.

over a different set of changes in wait time. Both field experiments are discussed more patiently below. Overall, our approach of randomizing price and wait time across consumers, together with knowing whether they decided to take the ride (based on both price and time), provides a set of VOT estimates unique to the literature in terms of approach, diversity of situation, and scale.

2.3 Context: Background on Lyft and Data

Lyft is a ridesharing platform that matches consumers (passengers) searching for motor vehicle transportation with independent contractors (drivers) providing the service.[18] Passengers access Lyft through a smartphone app (Figure 1 shows the passenger user interface at the time of the experiment). The app shows the nearest driver's estimated time to arrival (ETA) and the Prime Time (PT) price multiplier active for the potential passenger's current location. ETA is an estimate of the time, in minutes, that it would take the nearest driver to reach the passenger's location from the moment their request is accepted by the driver. Prime Time (PT) is a dynamic mechanism that increases prices from a base level to balance the local amount of Lyft ride requests (demand) and the local available pool of Lyft drivers (supply). A PT multiplier of +25% means that a ride will cost 25% more than the usual base fare (which is a deterministic function of distance and time of travel for each metro area).

In addition to showing an ETA estimate and the current PT level, the app also allows the passenger to input a destination. At the time of the first field experiment, users could see the ETA estimate and PT multiplier and request a ride without entering a destination. If a user did enter a destination, they were shown an estimated range for the trip's cost, not including the effects of the PT multiplier on the price. These estimated ranges were only seen by the passenger in approximately one-third of sessions in our sample for the first experiment. Because the ranges were expressed in terms of the base price, they were not affected by the experimental treatments. Our primary model for describing the factors influencing Lyft requests estimates the price effect using the experimental variation in the PT multiplier and does not require information about the base price.[19]

Each opening of the Lyft app by a passenger starts a session, which is the primary unit of observation for our analysis. A session ends either when the passenger takes a trip or after 30 minutes of inactivity. So, for example, if a user closes the app without taking a ride and reopens it within 30 minutes, both these interactions count as a single session. For each session, Lyft records the following information: passenger's unique ID code, whether the rider is registered as

[18] The following discussion of the Lyft ridesharing platform describes what existed in late 2015 through early 2016, the time period during which the first experiment was conducted. Lyft also now provides options for bikes, scooters, transit, and rental cars on its platform.

[19] Because price ranges are given only when a potential rider enters a destination, this information could affect the response to the experimental variation in the multiplier; we address this issue by considering the sample of respondents who enter a destination separately from those who did not as part of our robustness analysis in Tables C.26, C.27, and C.28 in Appendix Section C.

a business user, the local start time, the passenger's current location (latitude and longitude), counts of how many requests the passenger makes and rides the passenger completes, as well as the ETA and PT shown to the passenger in the session. ETA and PT may vary over the course of one session due to real-time changes in local supply and demand. As such, our analysis focuses on the last shown ETA and PT in the session, as these are the ones faced by the passenger at their final decision node of whether to request a ride.[20] Thus, all of our discussion of the ETA and PT will refer to the last value presented in a session (unless otherwise specified).

From these data, we can define a number of variables to characterize the circumstances facing the potential passengers as they made their choices. For example, using a deidentified, unique ID for a passenger, we can determine how many Lyft rides that passenger has taken. Using session start times, we can categorize a session as taking place on a weekend evening, during the morning commute, or in any other time category. And, using location data, we can determine if a passenger is at an airport or at a downtown or suburban location for each metro area in our sample. Finally, note that Lyft offers various ride modes, including Classic (the standard mode), Lyft Line, now called Shared, (in which several passengers share a single car with multiple pickups and dropoffs), and Lyft XL (which offers larger vehicles). At the time of the experiment, the majority of rides (about three-quarters) were Classics. We include all ride types in our analysis, and consider exclusion of non-Classic sessions as a robustness check (see Appendix Section K).

3. Design and Results For Field Experiment 1

To identify how agents trade off time and money, we begin with a first natural field experiment (see Harrison and List (2004) for the various field experiment definitions) that randomly assigned consumers to one of several treatments that differ in both the realized wait time and price (i.e. Field Experiment 1). The second natural field experiment, which we denote as Field Experiment 2, is described in Section 4.

3.1 Design of Field Experiment 1

The process that determines both the price multiplier and the waiting time implies that both variables are endogenous. The PT algorithm is designed to raise prices during periods of relative high demand/low supply; similarly, the number of available drivers at the time each potential rider opens the app in relation to the others who do so at the same time in a location will determine the estimated wait time. Our first field experiment involved nine cities in the U.S. (San Francisco, Austin, Atlanta, Miami, Los Angeles, San Diego, Boston, Seattle, and New York City)

[20] ETAs vary over the course of a session in 51.3% of sessions, while PT varies in 11.9% of sessions. The intra-session variation in ETAs is caused by drivers continuously moving during the course of a session, and is generally small. For example, in 76.4% of sessions, the difference between the maximum and minimum ETAs shown is one minute or less.

for eight weeks between December 2015 and January 2016, which involved 720,059 customers and 5,177,358 individual sessions.[21]

At the start of the experiment, 37% of all users in each city were randomly assigned to a control or one of five treatment groups. The experimental treatments ensure that there is exogenous variation in the components of each algorithm determining the PT multiplier and the wait time. The algorithm for wait time can increase the wait time above the normal arrival time, but cannot reduce the wait time. As a result, the wait time variations are limited to a high and normal (ETA). Each is matched with three possible treatments for the price algorithm: low, normal, and high. The sub-sample assigned to the normal ETA and normal price combination is treated as the control group. The remaining 63% of users associated with other sessions are excluded from our main analysis.[22]

The proportion of users randomly assigned to each group is shown in Table 1.[23] Table C.1 compares the features of the control and the treatment groups in terms of the available variables for describing each session. Based on these variables, the randomization achieved balance on the observable covariates.

Throughout the eight weeks of the experiment, each individual remained in the same treatment. Thus, a user assigned to the low price, high ETA treatment group had all of his or her sessions during the eight weeks subjected to the same algorithm, which would lead to potentially lower price and higher waiting times than what would be the case if they were in the control group. Of course, in practice, specific values for the wait time and price multiplier vary depending on how local conditions affected the outcomes produced by each algorithm. We understand the potential selection effects that might occur from this long-term design (although we use a different design in the second field experiment), and we analyze such effects thoroughly in Appendix Section I.

Price variation was achieved by modifying each user's PT multiplier, increasing it (for high price treatment groups) or decreasing it (for low price treatment groups) based on local market imbalances. This modification has the effect of raising or lowering a user's effective price, but not necessarily in every session.[24] Waiting time variation was achieved by removing drivers from the nearest driver queue of each affected passenger: the nearest driver, all drivers whose ETA

[21] Sessions in one city, Nashville, were dropped due to implementation problems with the experimental treatments. Including Nashville data in our full-sample regressions does not significantly change our point estimates. See Figure B.3 for the number of sessions per day over the full duration of the experiment (eight weeks).

[22] These sessions were not used as additional control observations because they may have been subject to other experiments conducted at Lyft concurrently with our field experiment.

[23] Treatment group assignments are determined by applying a hash function to each user's unique Lyft ID code. Since the assignment of each user to a treatment was random, we can assume that users in each treatment group are a representative sample of ride share users who open the Lyft app in the affected cities during the time of the first experiment.

[24] More concretely, because PT takes values in a fixed, discrete set (0%, 25%, 50%, etc.), the change in the algorithm's sensitivity to market conditions may not always result in a different PT level. For example, if the market has much more supply than demand, both the normal and the more sensitive, high price algorithm may find that the optimal PT level in the allowed set is 0%.

was within 30 seconds of that of the nearest driver, and one additional driver were removed from the queue. This removal has the effect of increasing ETA by at least 30 seconds for all passengers subject to the high ETA treatment, but potentially by considerably more than 30 seconds, especially when there are few drivers near a passenger. Because each passenger's price and ETA treatments are independent, we can identify the coefficients for both price and time effects on the demand for Lyft rides.

Table 2 displays average ETA, PT, and completed ride price by treatment group. The randomization was successful: our high ETA treatment increases the ETAs by an average of approximately 1.6 minutes, an increase of about 52%. The high price treatment increases average PT levels from about 10.0% to 16.3%, while the low price treatment decreases the average PT levels to about 3.6%. These PT differences result in completed ride prices that are about 2.5% higher for passengers receiving high price treatments and 2.0% lower for passengers receiving low price treatments.[25]

Table C.5 and Figures 2 and 3 show the distributions of PT and ETAs for each of the six treatment groups.[26] Appendix B provides graphs of the distributions for each variable, indicating that each of the treatments shifts the distributions in the intended directions. Tables C.3 and C.4 in the Appendix report p-values from Kolmogorov-Smirnov tests of the hypotheses that the distributions of average ETA and PT across users differs between treatment groups; the results suggest that the ETA distributions are approximately identical across PT treatments and vice versa, consistent with the independence of these components of the experimental treatments.[27]

3.2 Results of Field Experiment 1

3.2.1 Summary Statistics

Lyft customers in the experiment had, on average, about seven sessions and about four sessions with a completed ride (see Table C.2). Customers in the high ETA and high price treatments had fewer sessions, ride requests, and completed rides than control passengers, while passengers in the normal ETA, low price treatment group had slightly more than the control. These differences are consistent with a treatment effect on passenger behavior, which we explore further below.[28]

Before moving to the formal analysis, we consider the effects of the various treatments on passengers' demand behavior. The lower panel in Table C.2 shows the average number of sessions

[25] The effects of treatment on completed ride price are smaller than the effects on quoted PT because PT itself affects the probability that a session will result in a completed ride.

[26] Two sessions had recorded ETAs of 0 minutes. These were dropped from the data, so that the log transformation could be applied to ETA.

[27] Figures B.1 and B.2 in the Appendix indicate that these treatments are in effect consistently throughout the course of the experiment.

[28] Figures B.7, B.8, B.9, and B.10 highlight the heterogeneity in our sample. While the majority of our sessions come from San Francisco and Los Angeles, we have a large number of observations from the other six cities in the experiment. The distributions of sessions over days of the week and hours of the day are relatively balanced, though weekend and late afternoon/early evening times are the best represented time periods in our sample.

which had ride requests for each treatment group. Figures 4 and B.11 display the demand rate, defined in two ways. The first, Figure 4, uses the total requests for service compared to those opening the Lyft app. The second definition uses completed rides in place of requests. Some requests are not completed because a passenger is not matched to a driver, or the passenger or driver cancels a ride before it is finalized. These situations amount to 0.6% and 10.0% of the total requests during experiment 1, respectively.[29] As expected, the high ETA and price treatments decrease request rates relative to the control, while the low price treatment increases request rates. The small confidence intervals around the means suggest that these differences are statistically significant at the 95% level, and the magnitude of the differences of request rate between treatment groups—which is as large as four percentage points between the normal ETA, low price treatment and the high ETA, high price treatment suggest the outcomes reflect economically consistent responses to the differences in the circumstances of choice. For example, the high ETA, low price treatment had a slightly higher request rate than control, and the high ETA, normal price treatment had a slightly higher request rate than the normal ETA, high price treatment.[30]

3.2.2 Empirical Estimation

Our dependent variable is a discrete indicator for a request for the service (1 for request, 0 otherwise) with $\ln(\mathrm{ETA})$ and $\ln(1 + \mathrm{PT})$ as the independent variables of direct interest. We also have a set of controls that include fixed effects for the location, local hour of week and week of year, user experience with Lyft (decile of pre-experiment lifetime rides), and user type (whether the use has a business profile). Since the data generating process implies ETA and PT will be endogenous, 2SLS is our preferred estimator. As part of a robustness analysis, we estimated a probit model (instrumenting $\ln(\mathrm{ETA})$ and $\ln(1 + \mathrm{PT})$) of our main specification and report these results in Table C.24 in Appendix C. Standard errors are estimated clustering within passengers

[29] Figure B.4 in the Appendix considers how the average number of rides a passenger in each treatment group takes evolves over the course of the experiment, relative to the average number of rides a control passenger takes. The five series are approximately comparable at the outset of the experiment, appear to spread out with the duration of the experiment, and then appear to stabilize about 10 days after a user's first session in the experiment.

Figure B.5 in the Appendix is a similar plot comparing session rates (that is, percentage of passengers opening the app) between the treatment groups over time. Expected effects of the treatments on rates of use of the platform are observed: passengers facing the higher prices and waiting times become less likely to return to the platform over time. Price and waiting time thus have both intensive- and extensive-margin effects on demand, impacting not only the passenger's probability of requesting after opening the app, but also the probability that the passenger opens the app again in the future. We explore how this extensive margin behavior impacts the VOT estimation in section 3.4.

Finally, our observations are distributed evenly over passenger/experience levels. That is, we observe a near equal number of sessions for passengers with 0 rides and over 50 rides before the start of the experiment. The presence of this heterogeneity across regions, time, and passengers allows us to investigate how the value of time varies across circumstances.

[30] In addition to these during-experiment demand effects, we also find some evidence of treatment effects persisting beyond the end of the experiment; see Appendix E.

and are assumed independent across passengers.[31]

Our estimate for the VOT uses the average price and ETA for the control treatments associated with each sample definition. For most of our models, this is defined as:

$$VOT = \frac{\beta_1}{\beta_2} \frac{\overline{Price}}{\overline{ETA}} \tag{14}$$

β_1 and β_2 are semi-elasticities of demand (request rate) with respect to waiting time and price, respectively. To recover estimates for the relevant elasticities, we divide these semi-elasticities by the average request rate $\overline{Request}$. As previously noted, both $\ln(\mathrm{ETA}_{ij})$ and $\ln(1{+}\mathrm{PT}_{ij})$ are endogenous in (13), so we estimate the βs via two-stage least squares (2SLS), with first-stage equations (11) and (12). [32, 33] The F statistics from the first stage regressions, both with and without other controlling covariates, confirm the strength and relevance of the instruments. Standard errors for our estimates for the VOT are derived using the delta method.[34]

Our approach to exploring the sensitivity of our VOT estimates follows from the basic structure

[31] This covariance structure would arise if, for example, the true data generating process had β_1 and β_2 as individual–level random effects. Some studies, including Small et al. (2005), have estimated such random (or mixed) effects models directly. Endogeneity of our explanatory variables and the individual–level randomization of our experimental treatments render this approach difficult in our context. Instead, we simply cluster our standard errors to account for the possibility that β_1 and β_2 vary between passengers. Note also that clustering at the individual user-level is consistent with our experimental treatments, which are randomized at the individual.

[32] First stage results by region can be found in the Appendix Section C in Table C.29 and C.30.

[33] The necessary conditions for 2SLS to give consistent estimates of βs in (13) are that our instruments be orthogonal to the error term ε_{ij} and correlated with the endogenous variables $\ln(\mathrm{ETA}_{ij})$ and $\ln(1{+}\mathrm{PT}_{ij})$. Exogeneity follows from the fact treatment randomization, but may fail if exposure to the treatment has a cumulative effect on passengers, which affects their future behavior outside of the effect on their received ETA and Prime Time within each session. Appendix Figure B.5 suggests that such cumulative effects on demand may be present. As part of our robustness analysis, we estimate the model using only each passenger's first session in the experiment and also only observations in the first week of the experiment. We find little variation in the estimates for the VOT over the number of sessions or weeks of the experiment, suggesting no clear selection effect due to consumers who decide not to use the app. As a result, the appearance of small cumulative effects do not appear to affect the time/price tradeoffs we estimate; see Tables C.19, C.20, and C.21 in the Appendix.

[34] See, e.g., Davidson et al. (2004). Treating price and ETA as fixed, define the nonlinear function $g : \mathbb{R}^3 \to \mathbb{R}$ by:

$$g(\beta_0, \beta_1, \beta_2) = \frac{\beta_1}{\beta_2} \frac{\overline{Price}}{\overline{ETA}} \tag{15}$$

so that $g(\hat{\beta})$ is our estimator of the VOT. The Jacobian of g is

$$Dg(\beta_0, \beta_1, \beta_2) = \left(0, \frac{1}{\beta_2} \frac{\overline{Price}}{\overline{ETA}}, -\frac{\beta_1}{\beta_2^2} \frac{\overline{Price}}{\overline{ETA}}\right) \tag{16}$$

which exists (provided $\beta_2 \neq 0$) and is always nonzero. Assume that Σ is the asymptotic variance–covariance matrix of $\hat{\beta}$, that is, $\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \Sigma)$. By the delta method,

$$\sqrt{n}(g(\widehat{\beta}) - g(\beta)) \xrightarrow{d} N(0, [Dg(\beta)]\Sigma[Dg(\beta)]^T) \tag{17}$$

Then if $\hat{\Sigma}/n$ is any estimator of Var[$\hat{\beta}$] with $\hat{\Sigma} \xrightarrow{p} \Sigma$, a consistent estimator of Var[$g(\hat{\beta})$] is:

$$\frac{1}{n}[Dg(\hat{\beta})]\hat{\Sigma}[Dg(\hat{\beta})]^T \tag{18}$$

of the Becker model. The first of these is the selection of functional form describing how wait time and price influence requests. By transforming the price and time using logs we can estimate the parameter describing how people respond to price knowing only the PT multiplier. Second, to recover a measure of the VOT we need to select a point for evaluating the implied time/price trade-off. As noted earlier, we use the average values for these variables from the control treatments to estimate the semi-elasticities. Finally, to develop insights into how the circumstances of each person's choice affect the VOT, we estimate the VOT using a set of sub-samples motivated by an extension of the Becker theory, which is presented below.

3.2.3 VOT Estimates

Tables 3 and 4 provide the first- and second-stage results for the main model respectively, estimated for the full experiment 1 sample.[35] Table 3 demonstrates that our experiment worked as expected in changing wait times and prices in the correct directions as specified by our experimental groups. In addition, we have strong instruments in the first-stage regressions. In the second-stage results in Table 4, coefficients on $\ln(\mathrm{ETA})$ and $\ln(1+\mathrm{PT})$ are estimated to be -0.026 and -0.330 when controls are included; without controls, the coefficient on $\ln(\mathrm{ETA})$ is smaller in absolute magnitude, and the coefficient on $\ln(1+\mathrm{PT})$ slightly larger in absolute magnitude.

Using the full sample where the overall average request rate is 64.2%, we find estimates implying that wait time and price elasticities of demand are -0.0427 and -0.5942, respectively.[36] Our results provide strong support for the conclusion that Lyft requests are influenced by differences in both wait time and the price multiplier associated with the trip. Taken together, our estimated coefficients on $\ln(\mathrm{ETA})$ and $\ln(1+\mathrm{PT})$ imply a VOT of $\$19.38$ per hour (s.e. = $\$1.39$ per hour[37]) at baseline (i.e. control) waiting time of 3.08 minutes and price of $\$13.83$ (the average actual fare paid by control riders in the sample).

3.3 How do the Properties of the Situation Affect the VOT?

In this section we combine the logic of the Becker model with the temporal and spatial delineation in our sample to consider how situational features affect our estimates of the elasticities and the VOT. Recall from Section 2.1 that in Becker's model with Leontief household production, time and

In practice, we use the standard cluster-robust 2SLS variance-covariance estimator for $\hat{\beta}$, with clustering at the passenger level; see Baum et al. (2003).

This standard error ignores the uncertainty introduced by estimating $\overline{ETA}$ and $\overline{Price}$ from the data. We also estimated the standard error for our main specification using a pairs–cluster bootstrap (Cameron and Miller, 2015) on our full estimation procedure, and the results were similar to those returned by the delta method.

[35] Standard diagnostic tests of endogeneity and overidentifying restrictions are provided in Tables C.17 and C.18 in the Appendix. The results of OLS estimation of equation (13) are in Table C.6.

[36] Our price elasticity estimate is consistent with Cohen et al. (2016), whose mean point estimate for the price elasticity of demand using Uber records for 2015 is -0.57.

[37] A pairs–cluster bootstrap (Cameron and Miller, 2015) with $B = 999$ replications yielded a standard error of $\$1.40$ per hour and a bootstrap-t confidence interval (Efron and Tibshirani, 1994) of $(\$16.71, \$22.19)$.

private goods are perfect complements. When this restriction is relaxed, the wait time and price of a rideshare trip enter the indirect utility function without restriction, and we may consider how each of these influences our measure of the value of time.

Consider the willingness to pay WTP for a price increase ($1 + PT_0$ to $1 + PT_1$) and waiting time decrease (ETA_0 to ETA_1). W is defined by equation (19).

$$V(1 + PT_1, ETA_1, m - WTP) = V(1 + PT_0, ETA_0, m) \tag{19}$$

Denoting the marginal value of time (V_{ETA}/V_m) by π and the demand for rideshare trips by T, we have the following second-order expansion for WTP:

$$\begin{aligned} WTP \approx \pi(ETA_1 - ETA_0) - T(PT_1 - PT_0) + \frac{1}{2}(\pi_{ETA} - \pi\pi_m)(ETA_1 - ETA_0)^2 \\ - \frac{1}{2}(T_{PT} + TT_m)(PT_1 - PT_0)^2 + (\pi_{PT} + T\pi_m)(ETA_1 - ETA_0)(PT_1 - PT_0). \end{aligned} \tag{20}$$

When the changes in price and waiting time exactly offset each other, we have $WTP = 0$. If we also assume that $T = 1$ and that the adjustment of trips to price (T_{PT}) and income (T_m) are neglible, we can solve for the ratio of the price change to the equivalent waiting time change:

$$\frac{PT_1 - PT_0}{ETA_1 - ETA_0} \approx \pi + \frac{1}{2}(\pi_{ETA} - \pi\pi_m)(ETA_1 - ETA_0) + (\pi_{PT} + \pi_m)(PT_1 - PT_0). \tag{21}$$

Equation (21) shows that, to first order, the ratio of the price change to the equivalent waiting time change equals the value of time π. The equation further shows that this ratio is also influenced by the size of the waiting time change ($ETA_1 - ETA_0$), the sensitivity of the VOT to the waiting time (π_{ETA}), and the sensitivity of the VOT to income (π_m).

This conventional model fails to capture another of Becker's insights: the context of choice. The simple Leontief version of his model captured this effect by allowing time to be linked in different ways to different private goods. We observe only one private service: rideshare trips. Nonetheless, we can observe how the price equivalent changes in different situations, indexed by location, time of day, and weather condition. This rich detail in the circumstances of people's choices allows the data to add resolution to what a largely unrestricted model may imply for the VOT. By selecting sub-samples distinguished by the location, timing, and other features of passenger sessions, we can evaluate how each affects potential users' decisions with randomly varied prices and wait times, yielding heterogeneity tests across three bins of the data:

1. **Individual trip features** that are correlated with the opportunity cost of time, which include purpose of trip, time and day of week of trip, and specific location of trip. Such monetary values are driven by what activities are crowded out by greater wait times.

2. **Other trip signatures** such as weather conditions, the baseline ETA, and who bears the

marginal cost of the trip. In this case, when the cost burden is shared by an employer the rider should be more sensitive to temporal changes.

3. **Market variables** that are correlated with the opportunity cost of time, which include local wages and the available substitutes, such as access to alternative public transportation. Such variables affect the nature and shape of the demand curve for time.

Taken together, data generated across both natural field experiments combine to shed important insights on factors within each bin. In each bin, a mixture of experimental and non-experimental variation is utilized, and we take caution in interpreting results based on the naturally-occurring variation and selection below. The features which define each bin may be correlated with other features affecting preferences and the circumstance of choice; in Appendix J we attempt to address this concern in a secondary experiment by reweighting observations within each bin to balance the distributions of other relevant observable covariates.[38]

3.3.1 Individual Trip Features

Table 5 summarizes our findings concerning trip and market features.[39] Consider first the different ways of describing the timing for choices: by day of the week and by an hour of the day. In the latter case, we also distinguish weekends (Saturday and Sunday) from weekdays. Our reasoning above yields predictions on how the various situations should affect VOT estimates. More specifically, in the case of distinguishing days of the week we would expect that the opportunity cost of time is higher during weekdays than on weekends. Our estimates in Table 5 support this conclusion, as weekdays experience ETA elasticities larger than we observe on weekends. And, VOTs average nearly \$20 during the week versus around \$18 on the weekend. While these differences are meaningful, overall they are not consistently different at conventional significance levels.

Yet, when we consider the time of the day, distinguishing weekdays from weekends, we find pronounced and statistically significant differences in the VOT estimates at the $p < 0.01$ level. These differences mainly arise in the morning and evening commuting rush hour periods for weekdays (6 to 10 AM and 4 to 7 PM) where we observe larger ETA elasticities (especially in the mornings) and higher VOTs—averaging above \$24—compared to non-commuting time blocks. For weekends, we observe less variation throughout the day, as we might expect.

These estimates suggest that when the wait time crowds out work time the ETA elasticity and the VOT estimates are higher than when it crowds out non-work time. In this manner, the

[38]We also indirectly examine the impact of reliability by checking the robustness of our results when we include controls for late arrivals in passengers' previous interactions with the Lyft app in Appendix Section D. Similarly, we attempt to address the potential bias that may arise from not accounting for the effect of in-vehicle time on passenger demand in Appendix Section F. Finally, we check the sensitivity of our results to the presence of always- or never-requesters in our sample in Appendix Section G, and to user selection in Appendix Section I.

[39]The corresponding full regression tables are C.7 and C.8 in the Appendix.

pattern of results observed in Table 5 are consonant with the notion that individual trip features predictably affect the relevant primitives in the model.[40]

3.3.2 Other Trip Characteristics

Table 6 considers other features of the trip that potentially affect the marginal sensitivities to price and waiting time.[41] Our data from this first field experiment allows us to distinguish choices along the dimension of wait time unpleasantries as well as who bears the burden of the trip. Field experiment 2 provides the necessary variation to explore how baseline ETA variation impacts VOT estimates and therefore provides a sense of the curvature of the VOT. We discuss those results in Section 4.

The top panel of Table 6 provides our full sample estimates, and as a comparison the middle panel provides relevant estimates when there is rain, snow, and no precipitation.[42] We find that the trips with adverse weather conditions suggest a clear and discernible impact on the VOT estimates, suggesting that precipitation makes longer waiting times considerably less desirable. For instance, in the presence of snow, the ETA elasticity is more than two times higher than when there is no precipitation. This enhanced time sensitivity leads the VOT estimate to be nearly 40% higher than no precipitation trips ($26.56 versus $19.04). The effect of rain, while muted compared to snow, is also significantly larger than when there is no precipitation.

While both rain and snow cause the ETA elasticities to be higher, and in turn increasing the VOT estimates, tempering these increases are the price elasticity increases that we observe. While such increases are indeed small, we expected that as weather conditions worsen consumers would become less sensitive to price changes. These estimates may reflect that baseline PT is higher in such cases (15.2% in rain and 20.0% in snow).

The bottom panel of Table 6 provides insights into how the burden of cost affects the relevant elasticities. We use information on whether the trip was taken by a business or a non-business user, and make the assumption that the cost burden is lower for the former either because their firm pays the expense or if that is not the case at least the expense can be used a tax write-off. In both cases, the consumer bears less than 100% of the cost of trip burden.

We find that business users' ETA elasticities and VOTs are different from those of non-business users in expected ways: non-business users have a price elasticity that is roughly 50% larger (-0.605 versus -0.399), and a value of time about 25% higher. Inspecting the relevant elasticities, we note that this difference is driven by business users showing much less sensitivity to price, presumably reflecting the fact that business users may not themselves bear the full monetary

[40] In Appendix I, we find that these heterogeneity results are robust to correcting for user selection into different trip contexts.

[41] The corresponding full regression tables are C.9 and C.10 in the Appendix.

[42] We obtain historical realized weather data from Dark Sky at the geohash4-hour level, and match this information to sessions based on the passenger's geohash4 and the starting hour of the session.

cost of a trip.[43]

3.3.3 Market Factors

Results on the effects of market factors on the VOT are summarized in Table 7.[44] We distinguish choices by location in three ways: by metro area; by prominent locations within each metro area (airports and downtown locations); and by distance to the nearest public transport station/stop.[45] Point estimates of the VOT vary between metro areas, but we fall short of rejecting the null of homogeneity across regions ($p = 0.154$).[46] While the estimates for Atlanta, Austin, and Seattle are distinctive, with the VOT for the first two metro areas smaller than our overall sample estimate, only the results for Atlanta would be judged as having a marginally significant difference from the estimate for the rest of the sample. We discuss more of the city-level estimates with respect to wages in Section 5.1.

Distinguishing downtown locations, the estimates indicate a greater VOT for downtown locations compared to non-downtown.[47] Trips initiating at airports have numerically larger VOT estimates, but the small sample size for these cases prevent judgement of statistical significance. The grouping of trips based on distance to nearest public transit stop yields VOT estimates that decrease with distance, consistent with available substitute reasoning. However, this category is also likely to serve as a proxy for non-downtown locations, which have lower VOT estimates.[48] In the second experiment below (section 4.2.1), we have random variation in wait time that is completely orthogonal to the base ETA.

4. Design and Results for Field Experiment 2

4.1 Design of Field Experiment 2

Our analysis of the second natural field experiment has three primary goals: (1) replicate the main insights gained from the first field experiment on ETA elasticities and various heterogeneity tests; (2) assess the robustness of the observed heterogeneity to sample reweighting adjustments; and (3) explore how the base level of wait time affects ETA elasticities and how exogenous changes

[43]We are unable to assess the robustness of precipitation and business/non-business results to user selection into contexts because our weather data does not extend far past the beginning of the first experiment and business status does not vary within user.

[44]The corresponding full regression tables are C.11, C.12, C.13, and C.14 in the Appendix.

[45]Public transit stop/station locations are taken from the National Transit Map, and distance is calculated as the great circle distance between a passenger's location and a stop.

[46]When considering the marginal demand effects of price and waiting time separately, we do reject the nulls that time ($p = 0.046$) and price ($p < 0.0001$) semi-elasticies are homogeneous across regions.

[47]Maps of the areas tagged as "downtown" can be found in Figure B.6.

[48]Results in Appendix I suggest that the distance to transit results are not primarily driven by user selection, while the downtown/non-downtown and airport/non-airport may be. We are unable to effectively control for selection into regions due to limited within-user variation in region (only 19% of users have a pre-experiment session in more than one region. Full tables of regression results are available in Appendix C.

in the length of wait time affect the VOT. That is, by precisely varying wait time at the pickup location and hour (as opposed to across users in Field Experiment 1), we develop insights into how the timing of different wait times along with their location affects time elasticities. To do these chores, we randomized location-hour blocks into one of the following four groups: (1) Control (10% of the city sample); (2) ETA plus at least 60 seconds (5%); (3) ETA plus at least 150 seconds (3%); and (4) ETA plus at least 240 seconds (2%). To increase ETAs, we removed the closest drivers from the dispatch queue until the ETA was increased by at least the treatment amount.

The experiment took place for eight weeks from April 2017 to June 2017 in Los Angeles, San Francisco, New York City, Chicago, New Jersey, Boston, Philadelphia, Washington, D.C., Miami, and Atlanta, involving 3.3 million passengers and 9.7 million sessions (5.9 million of which had requests for service). The randomization used in this second experiment differs from that of our first experiment in that treatments were assigned at the pickup location-hour level across the whole eight-week period. Twenty percent of all location-hour blocks in these cities in this eight week period were in the experiment.[49] Price multipliers were not randomized in this experiment; consequently, our focus is on waiting time elasticities rather than VOT estimates. Figures 5 and 6 and Table C.32 show that the experimental conditions and the treatment groups have the expected variation and overall effects on demand.[50]

4.2 Results of Field Experiment 2

To summarize the insights gained from our second field experiment, we split our results into two groups. First, we summarize the overall time elasticity estimates from field experiment 2, and explore how the properties of the situation affect the estimated ETA elasticity. In doing so, we explore how well results from field experiment 2 map into those from field experiment 1. Second, we delve into the shape of the estimated time elasticities. Section 5 focuses on the policy implications of our results from the combination of both field experiments, digging deeper into the relationship between VOT and local wages as well as exploring how our preferred estimates relate to preferred VOT values used by policymakers.

[49] The location unit used for this randomization is a geohash7, which has a width and height of at most 153 meters. Counts of geohash7s per region can be found in Table C.31. In all the analysis of data from this experiment, standard errors are clustered at the geohash7-hour level, to match the clustering of the randomization scheme.

[50] Figures B.12 and B.13 show the timeline of the ETAs and PT respectively over the course of the experiment. It is clear that both the ETAs and PT remain pretty stable over time apart from the holiday season, especially New Years Eve where both prices and ETA increase dramatically.

4.2.1 Time elasticities and how they vary: a replication

The first column of Table 8 provides estimates of the time elasticity from field experiment 2. As shown in the top row of Table 8, the estimated ETA elasticity for the full sample is -0.043.[51] This aggregate estimate is quite close to the ETA elasticity estimate from field experiment 1; indeed, we fail to reject the homogeneity null at the $p < 0.05$ level for these two estimates. When we focus on data drawn from the roughly 200,000 customers who are involved in both field experiments, the estimates are also statistically indistinguishable: we fail to reject the homogeneity null at the 5% level.[52]

The remaining estimates in column 1 of Table 8 explore heterogeneity in ETA elasticities along days and various times of the week. Concerning such trip features that affect the opportunity cost of time, we find qualitatively similar patterns in the ETA elasticities as we found in the first experiment: weekdays have larger time elasticities than weekends, and weekday elasticities are largest during peak commuting times. Similar to experiment 1, these results show the increased importance of time when it relates to opportunity cost during commute hours versus other hours.

Column 2 in Table 8 shows how the ETA elasticities vary with other trip characteristics such as inclement weather and business users. In this case, we again find that time elasticities are higher in rainy conditions: -0.047 versus -0.042 when there is no precipitation.[53] In this case, we find that business users exhibit a larger time elasticity than non-business users (-0.049 vs. -0.043) at the $p < 0.01$ level, despite facing lower ETAs on average (see Table C.40).

In column 3 of Table 8, we show how ETA elasticities vary spatially. Consistent with our findings in experiment 1, we find differences across cities in our sample, with New York and Washington, D.C. having the largest elasticities (-0.086 and -0.054) and Miami and San Francisco having the lowest elasticities (-0.021 and -0.031). Furthermore, non-downtown sessions have a larger time elasticity than downtown sessions (-0.048 vs. -0.039) at the $p < 0.01$ level, though the average ETA is higher in the non-downtown sessions than the downtown sessions (Table C.42).

Finally, we again find that distance to public transit matters: the largest ETA elasticities are observed for passengers more than 800 meters from their nearest public transit stop. This result runs counter to economic intuition: passengers with better outside transportation options should respond more elastically to higher waiting times. Distance to transit, however, is correlated with other relevant characteristics of the trip, such as base ETA (Table C.43). In Appendix J, we

[51] In these regressions, we include $\ln(1 + \mathrm{PT})$ as a control variable for consistency with the model used for the first experiment, though it is endogenous to the decision to request. Our estimates of time semi-elasticities replace ETA with an instrument constructed as predictions from a first stage regression including the treatment indicatros and and fixed effect controls for region, geohash5, local hour of week, local week of year, business user, and decile of user lifetime rides (see Table C.33). These controls are also included in the second stage equation for Lyft requests. Consistency of the estimates for our time semi-elasticity relies on the exogenous treatment effects and the common controls included in both the first and second stage equations. The endogenous price multiplier will not be correlated with our experimental variation in the ETAs each user faces, and as a result will not affect the consistency of our ETA semi-elasticity of demand estimate. The ETA treatments do not affect $\ln(1 + \mathrm{PT})$, as we show in Table C.34.

[52] These results can be found in table C.35.

[53] The second experiment had no sessions for which the reported precipitation type was snow.

adjust for these differences in the distribution of other covariates through sample weighting, and find that the ETA elasticity is largest for passenger near public transit, as expected from the theoretical model (Table J.6).[54]

4.2.2 The shape of time elasticities

An important feature of the second experiment is that we have the ability to observe multiple levels of experimental ETA increases. As a result, we can estimate the elasticities and VOTs implied by different increases in the wait times and can gauge the implied nonlinearity in the relationship, independent of the base wait time.

To operationalize our approach, we consider that our four levels of ETA in the field experiment provide the ability to contrast three levels of wait time, each of which may serve as an instrument to estimate the ETA elasticity of demand at different points of the demand curve. We consider three sub-samples of our data: all Control and Plus 60 sessions, all Control and Plus 150 sessions, and all Control and Plus 240 sessions. In each sub-sample, we estimate our main demand equation, using the indicator of treatment (respectively Plus 60, Plus 150, and Plus 240) as a single instrument for $\ln(\text{ETA})$. The resulting coefficient estimate is a weighted average of the expected derivative of demand with respect to $\ln(\text{ETA})$ over the ETAs between the two treatment levels in the sub-sample (Angrist et al., 2000). We divide this quantity by the average request rate in the sub-sample to obtain an estimate of the ETA elasticity of demand over the ETAs in each sub-sample. As ETAs increase from one sub-sample to the next, comparing these elasticities gives insight into how the ETA elasticity of demand responds to increases in ETA.

Table 9 and 10 provide empirical results of this analysis. Overall, across every base ETA level we find that increasing the wait time reduces demand. This finding represents a good rationality test. In addition, we find that time elasticities are increasing in base ETA: at a one minute base ETA, the elasticity ranges between -0.005 and -0.018, whereas at a ten minute base ETA the elasticity ranges between -0.230 and -0.289.

The interpretation of the elasticity of -0.026 in the first row and third column of 9, for example, is that toward the lower end of the demand curve with a 3 minute base ETA, a 1% waiting time increase translates to a 0.026% decrease in the quantity demanded. As we can see in the second row, a 1% increase in ETA up to the middle portion of the demand curve reduces demand by 0.044%, and the third row shows that a 1% increase in ETA up to the upper end of the demand curve reduces demand by 0.066%. This pattern suggests that time elasticities increase as we move further above the base waiting time.

We find that this result is consistent at all but one base ETA between one and ten minutes; time elasticities are larger for larger increases (see Table 9). In Table 11, we also find that this

[54] In experiment 1, we find that correcting for user selection into different distance bins does *not* significantly affect our ETA elasticity estimates; see Table I.5.

finding holds across all cities in our sample. We also find that ETA elasticities are increasing in ETA: (i) in both downtown and non-downtown sessions (Table C.46); (ii) across all days of the week (Table C.47); and (iii) across all hours of the day (Table C.48 for weekdays and Table C.49 for weekends).

There could be many mechanisms underlying the observed shape of the estimated time elasticities, but we are unaware of any previous causal evidence that shows the marginal value of time varies dramatically over base wait time. Nevertheless, one concern in interpreting our analysis on the shape of the time elasticity is that consumers' responses may reflect a belief that a longer ETA implies greater congestion, and then erroneously expect that their travel time would be longer than what was stated on the app. Under this interpretation consumers might be responding to the changes in the expected ETD in addition to changes in ETA.

We believe that this interpretation is not supported by the structure of the experiment. First, if ETA increases by one minute on the app, then the ETD is also updated by one minute, to reflect the increase in pickup time. Second, it seems unreasonable to expect that consumers believe the quoted ETA but not the quoted ETD. Third, the consumer might believe that as time moves further away from the present, predictions become less accurate, and as a result the variance in ETD increases. We lack the necessary data to seriously investigate such an effect, but leave the question open for future research.[55]

5 Issues in identification and external validity of our estimates

We next discuss potential threats to the internal and external validity of our natural field experiments. We focus on functional form assumptions (section 5.1), app competition (section 5.2), endogenous opening up of the app (section 5.3), experimental contamination (section 5.4), and external validity of our experimental results (section 5.5).

5.1 Functional form and specification

Our main econometric specification has wait time entering the demand equation in logarithmic form. This functional form assumption implies diminishing marginal cost of wait time. We assess the robustness of this assumption in Tables C.23 and C.25 by running our main regressions with ETA entering in levels, and find slightly larger VOTs (but not at conventional significance levels). The first experiment involves only a binary ETA treatment which does not precisely pin down an exact ETA increase (in sparse markets, for example, the treatment may induce larger absolute ETA increases than in dense markets); as a result, the experimental data provide no empirical basis for deciding between different functional forms. The second experiment randomizes sessions

[55]Table F.1 suggests that, in the 2017 experiment, ETA elasticities do not vary with the estimated trip time (pickup to dropoff), which suggests that the elasticites and VOTs we have estimated capture passengers' responsiveness to pre-trip waiting time, and not some combination of pre-trip and in-vehicle time.

between multiple levels of ETA increases, which provides some rough insight into the local shape of the demand curve, but is still insufficient for pinning down the exact functional form.

Another decision we made is to use a simple linear probability model as opposed to a more sophisticated hierarchical model. A potential issue with LPM is that of aggregation. Suppose that there are two types of consumers, one elastic to price and inelastic to wait time and the other inelastic to price and elastic to wait time. The average VOT may be underestimated when taking the ratio of average time and price responsiveness. In Appendix L, we instead use an interacted model to estimate price and time coefficients for different combinations of session features, compute VOTs within each combination, and then average across the population. The resulting mean VOT is within two standard errors of our main estimate.[56] A more complex random-coefficients approach would require imposing further structural assumptions on the model that are difficult to verify.[57] One attractive feature of our approach is that we can recover the VOT with the Becker model and weak complementarity without structural assumptions.

5.2 Competition

Another potential concern is that Lyft had competition in the market during the experiment, so multi-homing behavior could potentially cause the outside option to be non-static. Importantly, our randomization is orthogonal to the outside option at that time in which the customer uses the app. It is unlikely that outside providers would have been able to track and respond to the random variation induced at the session-level in real time.

5.3 Endogenous opening of the app

Our VOT is estimated only when riders are considering using the service. That is a key feature of the weak complementarity assumption in the Becker model for estimating a VOT. However, we do not observe behavior outside of opening and using the app. If users only open Lyft when particularly in a hurry, then our VOT estimates may tend to overstate the overall average VOT.

We explored this potential issue with two distinct approaches. First, as a proxy for urgency of use, we compare VOTs across quartiles of user rides taken in the 28 days before the first experiment in Table C.16. The thought is that those who use the app frequently are using it for reasons beyond being in a hurry. We find a slight decreasing trend of VOT with respect to prior 28 day rides. However, the VOT of the most regular users remains only about 20% lower than our main estimate for the full sample; and this difference is not significant at the 5% level, suggesting that selective opening of the app does not greatly bias our results. One shortcoming with this first

[56]These results may be sensitive to the choice of features used to define groups of sessions.

[57]For example, Buchholz et al. (2020) utilizes a hierarchical model with individual-specific coefficients to estimate the entire distribution of VOT by assuming (i) a particular parametric form to the distribution of individual-specific heterogeneity and (ii) that heterogeneity in time and price responses not captured by additive effects for a few locations and time categories is individual-specific; the full model is then specified by a (relatively) small number of parameters, which can be estimated by Bayesian or maximum likelihood methods.

approach is that usage could be endogenous to other things that are related to VOT, and so one would need to randomize urgency, wait time, and price to provide exact insights.

Our second approach to the endogenous opening of the app is to examine VOT across individual usage rates. The idea being that those customers who use the Lyft app regularly (e.g., for daily commute) likely have a stronger preference for Lyft as a means of transport and so do not simply use it when they are particularly in a rush. Alternatively, those consumers who use it only a few times per month are primarily using it when they are in a rush. Importantly, we do not find significant variation in the VOT by lifetime prior rides (see Table C.15 in the Appendix).

5.4 Contamination

When a user decides whether or not to request a ride, their decision affects total supply in the market, which may in turn influence the request decisions of other users (this is analogous to the displacement effect found in Crépon et al. (2013)). This channel introduces interaction effects in the experiment: the outcome of one session can be dependent on the treatment assignments of other sessions, since those treatment assignments affect the overall available supply pool. While these interaction effects technically violate the SUTVA assumption, we argue that the violation is minor and relatively unimportant, because (i) these interaction effects will on average uniformly affect all sessions in the market, across all treatments and in and out the experiment; (ii) the closest data available to shed light on the size of these interaction effects (described in Appendix J) suggests that their magnitude is small (on the order of a few seconds); and (iii) as we argue in Appendix M, despite this violation of SUTVA, we can still interpret our estimates as weighted average causal derivatives of demand.

5.5 External validity

Many of the above considerations relate to generalizability, or external validity, of our VOT estimates. As List (2020) notes, "all results are externally valid to some setting, and no result will be externally valid to all settings." Given that we view our results as speaking to not only tests of theory, but also to any application that involves time, the populations of people and situations to which our estimates apply merits serious consideration. In terms of sampled population, one key feature of our sample is that nearly 60% of individuals in the US have used rideshare, so consumption of the good itself is quite common. And, importantly, as shown in Appendix H, when we re-weight our user sample to more closely match the characteristics of the US population, our VOT estimates are not significantly changed. This suggests that the VOT of those customers who use ridesharing services is not significantly different from those individuals who do not use ridesharing services.

A next important consideration is whether there is sufficient similarity in relevant situational conditions to generalize our results (List, 2020). A first key feature is whether the experiment places the agent on an artificial or natural margin when making key decisions. In this manner, it would be difficult to find a better setting than the one studied in this paper to estimate the VOT within the theoretical construct of the classic studies. Indeed, our field experiment leverages key features of common purchase decisions in any familiar setting that involves price and quality trade-offs: several alternative choice sets are provided naturally and purchase decisions are made in a familiar form. In this spirit, such choices would be made whether or not our natural field experiment was being conducted.[58] We view this aspect of our approach as quite attractive.

A next important consideration is whether our two field experiments present relevant conditions to explore temporal trade-offs that are sufficiently similar to other settings. The richness of our data provides several variants that speak to very different situations. We have estimated wait time trade-offs at various opportunity cost levels: when wait times crowd out work time, leisure time, airport time, and nearly any other moment. Likewise, we have considered how liberal changes in the substitute set affect VOT estimates and how weather and aspects associated with who pays for lower wait time affects VOT. Where our approach has less coverage revolves around the circumstances of the wait (standing curbside rather than in gridlock traffic) and for time-tradeoffs that involve considering the opportunity cost of hours, rather than minutes, of wait time. We trust that future work in different settings and over different choices can complement our VOT estimates.

6. Implications for public policy

In this section, we use our VOT estimates to consider their implications for public policy. First, we offer some comparisons with the city's wage rate to examine whether local wage rates provide an important correlate with our VOT estimates. Second, we provide examples of how our VOT estimates impact certain public policy analyses that rely on estimates of the opportunity cost of time to assess the associated public actions. Third, our estimated price and wait time elasticities are nicely complemented by contemporaneous work that leverages structural models to estimate

[58] Importantly, our study does not require selection into surveys or the experiment, unlike several others in the empirical VOT literature. Our study also does not rely on strong structural assumptions (see above) for identification.

the VOT.[59]

6.1 The relationship between VOT and wages

Our region-level estimates of the VOT allow a test of how well conventional rules-of-thumb for approximating VOT hold across different metropolitan areas in the United States. Many of these rules-of-thumb for the VOT are formulated as percentages of the wage rate. For example, the USDOT (2015) recommends that the VOT be approximated as 50% of hourly earnings for personal travel (the same as Small (2013); others, particularly in the recreation demand literature stemming from early research by Cesario (1976), use 33% of hourly earnings as an approximation (see Phaneuf and Smith (2005)).

To assess the relevance of the various approximations, we use the estimates from our first field experiment across regions in an array of 16 possible rule-of-thumb approximations. The 16 rule-of-thumb approximations are given for all possible combinations of using the median or mean wage; using pre- or post-tax earnings; and using $1/3$ or less, less than $1/2$, greater than $1/2$, or 100% or more of the resulting hourly earnings. We use one-sided tests to explore whether the VOT is below $1/3$, below $1/2$, above $1/2$, and above 100% of earnings, for a total of 16 tests per region. The tests are all one-sided asymptotic z-tests with level $\alpha = 0.05$, based on the VOT estimates, standard errors, and median and mean wages reported for each region in Table C.11.[60] For a robustness check, we explore all 16 approximations with both a log and a linear specification to estimate the VOT.[61] For simplicity, we assume a constant 25% tax rate across all regions for both median and mean wages.

Table 12 summarizes our empirical results for the log and linear specifications. Each cell

[59] Buchholz et al. (2020) identifies 30 locations within Prague and considers ride choices for different days of the week, time periods during the day, weather conditions, and the location (inside or outside of structure) where choices are made. The location choice for consumers is based on a random utility specification that assumes price and time effects have random coefficients with means varying based on the features used to distinguish trips. The random error in both the price and time coefficients is assumed to reflect individual heterogeneity. A VOT in a location is established in a second estimation step. The model maintains that the choice to move from one location to another identifies the difference in the values of time for the origin of the trip compared to other locations. Their estimates for the VOT's for each location exploit the panel nature of their sample and use the random coefficient component of their model to construct the data for a linear program that is used to estimate the values of time for classes of individuals in each location. They rely on a control function strategy to help take account of price endogeneity. Their VOT estimates display sensitivity to the timing of the time that is similar to our findings, with higher values for peak commuting times on weekdays. While the absolute magnitudes of their price and time elasticities are about 10 times larger than our estimates, the relative size of the relevant elasticities are comparable. Moreover, during work times, the value of time implied by their model is approximately equal to the average wage rate. The second study, due to Castillo (2019), develops a structural model for the demand, supply, and market matching of Uber rideshare users and suppliers in Houston. The demand component of the model is estimated with a linear choice model that relies on the location specific price multipliers together with the rounding of surge adjustments to provide sufficient exogenous variation in prices and avoid bias in estimating price effects on rideshare choices. This study finds a much greater value of time, about two dollars a minute. Yet, the qualitative distinctions isolated for weekend and weekdays conform well to our findings, and the absolute magnitude of the price elasticity is comparable to our estimates.

[60] We take the median and mean wages as constants for the purposes of the hypothesis testing.

[61] To recall, the linear specification differs from the log specification because we replace $\log(\mathrm{ETA})$ in the first and second stages (equations (11) and (13)) with the level of ETA in hours. This specification keeps the price (multiplier) in log, as base prices (which would be necessary for calculating the level of price) are generally unobserved in our data.

records the number of regions (of eight) for which the null hypothesis that the VOT in that region equals the rule-of-thumb approximation of that cell is rejected at the 5% level. Overall, we find evidence in favor of rejecting the $1/3$ or less earnings rule for most regions, and have the least evidence in favor of rejecting the $1/2$ or more pre-tax mean wage and 100% or more post-tax median wage earnings rules. We do not reject the hypothesis that the VOT is above $1/2$ earnings for any region, but do reject the hypothesis that the VOT is above 100% earnings for some regions when using the linear specification. We repeat the same exercise using the city VOTs estimated using a linear functional form for the demand equation, of the same form considered in Section 3.3.2 (see Table C.25 for the city-level VOTs). The test results are broadly similar between specifications. Table 13 shows a similar pattern for the pre-tax median and mean wage values, with slightly more cities with a VOT over 100% of the pre-tax mean wage being rejected.

We provide an ocular depiction of these results in Figure 7, which plots our point estimates of the VOTs and their 95% confidence intervals alongside the post-tax mean and median wages (and the policy-relevant 1/2 and 1/3 of these quantities) for each region. Our region-level estimates also provide insights into the cross-sectional relationship of wages and the estimates for the opportunity cost of time. Using our log-functional-form estimates, we compute a Spearman (rank) correlation of 0.73 ($p = 0.025$) between the mean wage and the value of time estimate across regions. In Table 14, we apply the log transformation to the mean wage and our VOT estimates and run OLS and GLS regressions to estimate the cross-sectional elasticity of the VOT with respect to the wage level.[62] Our results imply wage elasticities of the VOT of 1.7 and 1.3 with the OLS and GLS specifications, respectively.

While these results are particularly striking for the first experiment, results from the second experiment suggest that these comparisons might underestimate the estimated connection because there is convexity in the VOT. By combining the multiple levels of ETA increases in the second experiment with the price variation of the first experiment, we can quantify how the VOT varies by magnitude of the increase in the waiting time. Specifically, for the five regions which overlap in the two experiments, we use the first experiment to estimate the average price and price (semi-) elasticity of demand, and the second experiment to estimate the average ETAs and ETA (semi-) elasticities of demand between each of the four treatment levels. We then combine these results to estimate the VOT between each of the four experimental ETA levels of the second experiment. To ensure that our price and time elasticity estimates are comparable yet independent, we subset both experiments to observations from the five overlapping regions, remove airport sessions from the first data, and remove from the first experiment observations from any users who appear in the second experiment.

Table C.50 provides the results of this analysis. Here, low, medium, and high refer to estimates

[62]The GLS specification uses as the weighting matrix the inverse of the estimated covariance matrix of the vector of log VOT estimates across regions, which is derived from the estimated covariance matrix of the vector of price and time semi-elasticities across regions using the delta method.

derived from contrasting Control and Plus 60 treatments, Control and Plus 150 treatments, and Control and Plus 240 treatments, respectively. We note that for the full sample and each individual region, the estimated VOT is highest when using the largest ETA treatment level as an instrument, supporting the conclusion that the value of a marginal minute is increasing as more minutes are added to the base wait time. Figure 8 provides a visual for these results, and plots them against the mean after-tax wages of the regions. The figure shows that the shape of the VOT elasticity over the changes in wait time impact whether the VOT is less than or equal to the after-tax wage rate in each city.

6.2 The impact of our VOT on public policies

As aforementioned, the VOT is an important input into many policy decisions. For example, transportation projects often have benefits that largely derive from the time savings associated with new or improved infrastructure, and similarly, the design of optimal road pricing depends crucially on a measure of travelers' VOTs. Policymakers often use cost-benefit analyses (CBAs) as a part of the information considered in prioritizing projects.[63] Many such analyses have embedded assumptions underlying the benefit (or the cost) estimates that rely on the types of proxy methods we reviewed earlier

Most cost-benefit analyses use the recommendations put forth by the Department of Transportation (DoT) in 1997, which was later revised in 2014. Both DoT reports tether the VOT to the median (or mean) wage in an area. While the recommendations allude to the complex set of factors that likely affect values of travel time across various contexts (e.g., comfort of transportation, reliability, magnitude of time change, etc.), there did not exist sufficient reliable data to incorporate these considerations more concretely into the estimated VOT. The reports recommend the VOT be set according to the trip purpose and typical wage rate: the median (or mean) wage should be used for business travel (i.e., on the clock) and half the median (or mean) wage should be used for personal travel.

Some of our analysis is consistent with the rule-of-thumb approach in the DoT reports: we find that the mean values of time are indeed correlated with the mean/median wage rates of various geographical regions, with an elasticity near 1. We also find that values of time are closer to the full mean/median wage rates during typical work commuting hours than at other times, consistent with the idea that travel related to work should be valued at a rate closer to the full hourly wage rate.

Our results, however, have the distinct advantage of being estimated with data that are exceptionally rich in source time variation, granularity, and breadth. As such, we are able to more directly address heterogeneity in VOT that has been posited but not empirically verified. In par-

[63] Executive Orders beginning with EO12291 first issued in 1981 require cost-benefit analyses for major new rules. In addition agencies now routinely use CBA as part of the resource allocation and decision-making process.

ticular, we find that the VOT is generally higher than 50% of the median/mean wage across most cities, and there is strong evidence of substantial heterogeneity by context (e.g. time of day, day of week, within-city downtown vs non-downtown areas). Our specific recommendations to policymakers are thus two-fold:

1. When feasible, account for the great deal of VOT heterogeneity with respect to cities, days of week, and times of day.

2. If data are unavailable or unreliable, adjust the mean VOT estimates up to 75% of the after-tax median/mean wage rate.

To demonstrate the potential relevance of our suggestions, we revisited a set of cost-benefit analyses from 1998–2017 that used a VOT for cities that were included in our experiment. We calculated how total costs, total benefits, and cost-benefit ratios change as we use (1) our region-specific empirical point estimates and (2) 75% of median/mean wage adjustment for value of time. We report the differences in Table C.51.

In each case, we find that the estimates used in the reports are lower than our region-specific estimates, so the total benefit (and thus benefit-cost ratio) tends to increase, by values ranging from 2% to 52%. While the qualitative result does not change for the cost-benefit analyses we examined (benefit-cost ratios are above 1 for all cases, both pre- and post-adjustments), it is still apparent from this exercise that a more directly measured, empirically-based VOT can substantially affect the relative benefits and costs of important projects compared to DOT's rule-of-thumb estimate. Indeed, our estimates are large enough to suggest that there was an important lot of projects that were not undertaken that would have provided substantial net benefits. Such "missed" projects are those tilted towards ones that have significant time savings, suggesting that current mis-measurement of temporal values yields under-investment in time-saving projects.

We also compute benefit-cost ratios following the simpler rule using 75% of mean/median wage (adjusted from the 50% of mean/median wage recommended by DOT). Among this set of cost-benefit analyses, this adjustment tends to yield values of time and measures of benefit that are usually close to but sometimes higher than the region-specific empirical point estimates. We view our estimates—based on exogenous price and time variation and rich, granular observations—have the potential to more efficiently allocate of resources and encourage technologies that induce more time savings for individuals.

7. Conclusion

Having gotten this far in our study you have surely invested a fair amount of time. We hope that such time was indeed an investment, and not ill-spent. This is because time is the ultimate scarce resource, and its value has deep implications for a range of economic phenomena and investment decisions. Our starting point is a literature from the 1960s that had deep implications for our understanding of the family, the household, and time allocation more generally. We leverage insights from these classic time allocation theories to provide a theoretically-consistent but updated approach to estimate the VOT. The theory carefully directs two large-scale natural field experiments on the Lyft platform to estimate the causal effects of wait time and price on ride-share demand.

We report several interesting insights. First, we estimate a VOT that is roughly $19 per hour (2015 prices). This estimate is 75-80% of the mean wage rate for the various regions in our experiment, which is quantitatively different from the findings of previous empirical studies on the VOT (Small et al., 2007) and is greater than the existing US policy guidelines on the VOT (USDOT, 2015). Second, we document that, consistent with standard microeconomic models (Becker, 1965; DeSerpa, 1971), the VOT is related to the opportunity cost of time, the available substitute set, and other key features of the trip that impact marginal benefits and marginal costs. Third, taken in aggregate, our research has key implications for policy. Specifically, we recommend that policymakers: (i) account for the great deal of VOT heterogeneity with respect to cities, locations within cities, day of week, and time of day; and (ii) adjust the rule-of-thumb VOT estimates up to 75% of the after-tax mean wage rate otherwise.

We view our VOT estimates as not only adding unique measures to a rich literature, but also providing a key link to the classic time allocation literature. Important areas where this research agenda goes from here can be found in caveats to our research. First, we do not examine the value of reliability in passenger travel or through using ridesharing companies, such as Lyft. We acknowledge that this value could be important and separate from the VOT, but is beyond the scope of this paper. Future studies should consider the value of reliability—the extent to which waiting times vary about their mean—simultaneously with the VOT. On a rideshare platform, one possible approach would be to run a multi-modal waiting time and price experiment, and to model the passenger's choice of not only whether to take a ride, but also of which ride mode (e.g., Classic or Shared) to take. Shared ride modes typically have more variable waiting times (today quoted as a range on the Lyft app), so such a cross-mode comparison might shed light on the value of reliability. Second, we focus on passenger travel in our data and ignore the VOT for rail, air, and freight travel. Combining the various travel modes and exploring their interplay is an ambitious research agenda that promises to lend deep positive and normative insights.

References

Aguiar, M., M. Bils, K. K. Charles, and E. Hurst (2017). Leisure luxuries and the labor supply of young men. Technical report, National Bureau of Economic Research.

Aguiar, M. and E. Hurst (2007a). Life-cycle prices and production. *American Economic Review 97*(5), 1533–1559.

Aguiar, M. and E. Hurst (2007b). Measuring trends in leisure: The allocation of time over five decades. *Quarterly Journal of Economics 122*(3), 969–1006.

Aguiar, M. and E. Hurst (2016). The macroeconomics of time allocation. In *Handbook of Macroeconomics*, Volume 2, pp. 203–253. Elsevier.

Aguiar, M., E. Hurst, and L. Karabarbounis (2013). Time use during the great recession. *American Economic Review 103*(5), 1664–96.

Allcott, H., L. Braghieri, S. Eichmeyer, and M. Gentzkow (2019). The welfare effects of social media. Technical report, National Bureau of Economic Research.

Anderson, M. L. (2014). Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. *American Economic Review 104*(9), 2763–96.

Angrist, J. D., K. Graddy, and G. W. Imbens (2000). The interpretation of instrumental variables estimators in simultaneous equations models with an application to the demand for fish. *The Review of Economic Studies 67*(3), 499–527.

Arnott, R., A. De Palma, and R. Lindsey (1993). A structural model of peak-period congestion: A traffic bottleneck with elastic demand. *American Economic Review*, 161–179.

Basso, L. J. and H. E. Silva (2014). Efficiency and substitutability of transit subsidies and other urban transport policies. *American Economic Journal: Economic Policy 6*(4), 1–33.

Bastiat, F. (1848). *Propriété et loi, Justice et fraternité*. Paris.

Baum, C. F., M. E. Schaffer, and S. Stillman (2003). Instrumental variables and GMM: Estimation and testing. *The Stata Journal 3*(1), 1–31.

Becker, G. S. (1965). A theory of the allocation of time. *Economic Journal*, 493–517.

Benhabib, J., R. Rogerson, and R. Wright (1991). Homework in macroeconomics: Household production and aggregate fluctuations. *Journal of Political Economy 99*(6), 1166–1187.

Bento, A., K. Roth, and A. Waxman (2017). Avoiding traffic congestion externalities? the value of urgency. Technical report, Technical Report, Working Paper.

Besley, T., J. Hall, and I. Preston (1999). The demand for private health insurance: do waiting lists matter? *Journal of Public Economics 72*(2), 155–181.

Browning, M., T. Crossley, and J. K. Winter (2014, April). The measurement of household consumption expenditures. IFS Working Papers W14/07, Institute for Fiscal Studies.

Buchholz, N., L. Doval, J. Kastl, F. Matějka, and T. Salz (2020). The value of time: Evidence from auctioned cab rides. Technical report, National Bureau of Economic Research.

Cameron, C. A. and D. L. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317–372.

Castillo, J. C. (2019). Who benefits from surge pricing? *Available at SSRN 3245533*.

Cesario, F. J. (1976). Value of time in recreation benefit studies. *Land Economics 52*(1), 32–41.

Chen, Y., G. Y. Jeon, and Y.-M. Kim (2014). A day without a search engine: an experimental study of online and offline searches. *Experimental Economics 17*(4), 512–536.

Chen, Y. and A. Whalley (2012). Green infrastructure: The effects of urban rail transit on air quality. *American Economic Journal: Economic Policy 4*(1), 58–97.

Cohen, P., R. Hahn, J. Hall, S. Levitt, and R. Metcalfe (2016). Using big data to estimate consumer surplus: The case of Uber. Technical report, National Bureau of Economic Research.

Couture, V., G. Duranton, and M. A. Turner (2018). Speed. *Review of Economics and Statistics 100*(4), 725–739.

Crépon, B., E. Duflo, M. Gurgand, R. Rathelot, and P. Zamora (2013). Do labor market policies have displacement effects? evidence from a clustered randomized experiment. *Quarterly Journal of Economics 128*(2), 531–580.

Davidson, R., J. G. MacKinnon, et al. (2004). *Econometric theory and methods*, Volume 5. Oxford University Press New York.

Deacon, R. T. and J. Sonstelie (1985). Rationing by waiting and the value of time: Results from a natural experiment. *Journal of Political Economy 93*(4), 627–647.

DellaVigna, S., J. A. List, and U. Malmendier (2012). Testing for Altruism and Social Pressure in Charitable Giving. *Quarterly Journal of Economics 127*(1), 1–56.

DeSerpa, A. C. (1971). A theory of the economics of time. *Economic Journal 81*(324), 828–846.

Duranton, G. and M. A. Turner (2011). The fundamental law of road congestion: Evidence from u.s. cities. *American Economic Review 101*(6), 2616–52.

Efron, B. and R. J. Tibshirani (1994). *An introduction to the bootstrap*. CRC press.

Finkelstein, A. (2009). E-z tax: Tax salience and tax rates. *Quarterly Journal of Economics 124*(3), 969–1010.

Gelber, A. M. and J. W. Mitchell (2012). Taxes and time allocation: Evidence from single women and men. *Review of Economic Studies 79*(3), 863–897.

Ghez, G., G. S. Becker, et al. (1975). The allocation of time and goods over the life cycle. *NBER Books*.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review 104*(4), 1091–1119.

Goolsbee, A. and P. J. Klenow (2006). Valuing consumer products by the time spent using them: An application to the internet. *American Economic Review 96*(2), 108–113.

Green, D. I. (1894). Pain-cost and opportunity-cost. *Quarterly Journal of Economics 8*(2), 218–229.

Greenwood, J., N. Guner, and G. Vandenbroucke (2017). Family economics writ large. *Journal of Economic Literature 55*(4), 1346–1434.

Gronau, R. (1973). The intrafamily allocation of time: The value of the housewives' time. *American Economic Review 63*(4), 634–651.

Hall, J. D. (2020). Can tolling help everyone? estimating the aggregate and distributional consequences of congestion pricing. *Journal of the European Economic Association*.

Harrison, G. W. and J. A. List (2004). Field experiments. *Journal of Economic Literature 42*(4), 1009–1055.

Johnson, M. B. (1966). Travel time and the price of leisure. *Economic Inquiry 4*(2), 135.

Juster, F. T. and F. P. Stafford (1991). The allocation of time: Empirical findings, behavioral models, and problems of measurement. *Journal of Economic Literature 29*(2), 471–522.

Kleibergen, F. and R. Paap (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics 133*(1), 97 – 126.

Kreindler, G. E. and Y. Miyauchi (2019). Measuring commuting and economic activity inside cities with cell phone records.

Krueger, A. B., D. Kahneman, D. Schkade, N. Schwarz, and A. A. Stone (2009). National time accounting: The currency of life. In *Measuring the subjective well-being of nations: National accounts of time use and well-being*, pp. 9–86. University of Chicago Press.

List, J. A. (2020). Non est disputandum de generalizability? a glimpse into the external validity trial. Technical report, Working paper.

Mäler, K.-G. (1971). *A Method of Estimating Social Benefits from Pollution Control*, pp. 106–118. London: Palgrave Macmillan UK.

Mäler, K.-G. (1974). *Environmental Economics: A Theoretical Inquiry*. Johns Hopkins Press.

Mas, A. and A. Pallais (2017). Valuing alternative work arrangements. *American Economic Review 107*(12), 3722–59.

Mas, A. and A. Pallais (2019). Labor supply and the value of non-work time: Experimental estimates from the field. *American Economic Review: Insights 1*(1), 111–26.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics 3*(4), 303–328.

Mill, J. S. (1848). *Principles of Political Economy*. John W. Parker.

Miller, G. and B. P. Urdinola (2010). Cyclicality, mortality, and the value of time: The case of coffee price fluctuations and child survival in colombia. *Journal of Political Economy 118*(1), 113–155.

Nevo, A. and A. Wong (2015). The elasticity of substitution between time and market goods: Evidence from the great recession. Technical report, National Bureau of Economic Research.

Nordhaus, W. (2009). Measuring real income with leisure and household production. In *Measuring the subjective well-being of nations: National accounts of time use and well-being*, pp. 125–144. University of Chicago Press.

Parry, I. W. and K. A. Small (2009). Should urban transit subsidies be reduced? *American Economic Review 99*(3), 700–724.

Phaneuf, D. J. and V. K. Smith (2005). Recreation demand models. *Handbook of Environmental Economics 2*, 671–761.

Philipson, T. J., G. Becker, D. Goldman, and K. M. Murphy (2010). Terminal care and the value of life near its end. Technical report, National Bureau of Economic Research.

Ramey, V. A. and N. Francis (2009). A century of work and leisure. *American Economic Journal: Macroeconomics 1*(2), 189–224.

Robinson, J. P. and G. Godbey (1999). Time for life: The surprising ways americans use their time. 2nd.

Schrank, D., B. Eisele, and T. Lomax (2012). Tti 2012 urban mobility report. *Texas A&M Transportation Institute. The Texas A&M University System 4*.

Small, K. (2013). *Urban Transportation Economics*. Taylor & Francis.

Small, K. A., E. T. Verhoef, and R. Lindsey (2007). *The economics of urban transportation*. Routledge.

Small, K. A., C. Winston, and J. Yan (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica 73*(4), 1367–1382.

Smith, V. and S. Banzhaf (2007). Quality adjusted price indexes and the willig condition. *Economics Letters 94*(1), 43–48.

Smith, V. K. and C. Mansfield (1998). Buying time: Real and hypothetical offers. *Journal of Environmental Economics and Management 36*(3), 209–224.

Su, Y. (2018). The rising value of time and the origin of urban gentrification.

Sunstein, C. R. (2018). Sludge and ordeals. *Duke Law Journal 68*, 1843.

USDOT (2015). Revised departmental guidance on valuation of travel time in economic analysis. *US Department of Transportation. Washington, DC*.

Van Ommeren, J. and M. Fosgerau (2009). Workers' marginal costs of commuting. *Journal of Urban Economics 65*(1), 38–47.

von Wieser, F. (1876). Über das verhältnis der kosten zum wert (on the relation of cost to value). Technical report, Gesammelte Abhandlungen, pp. 377-404.

Wheaton, W. C. (1977). Income and urban residence: An analysis of consumer demand for location. *American Economic Review 67*(4), 620–631.

Figures

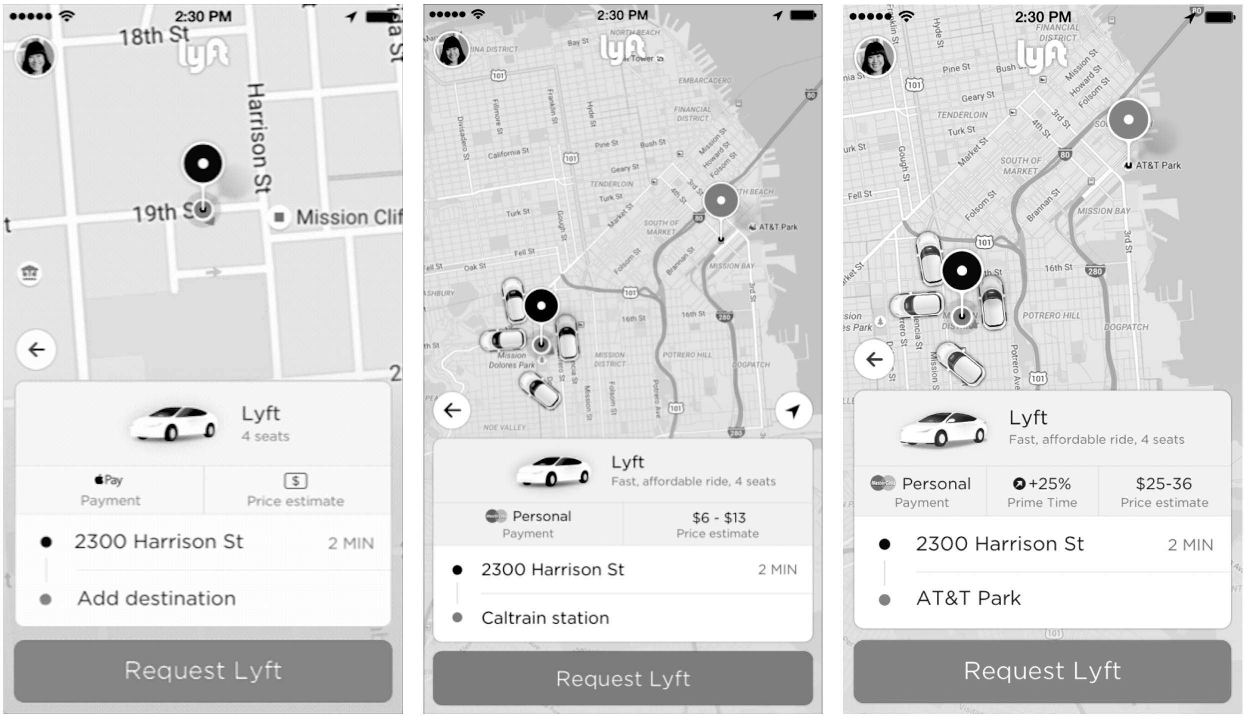


Figure 1
Screenshots of the passenger Lyft app's user interface, at the time of the experiment. Left: no destination entered and no Prime Time active. Middle: destination entered and no Prime Time active. Right: destination entered and +25% Prime Time active.

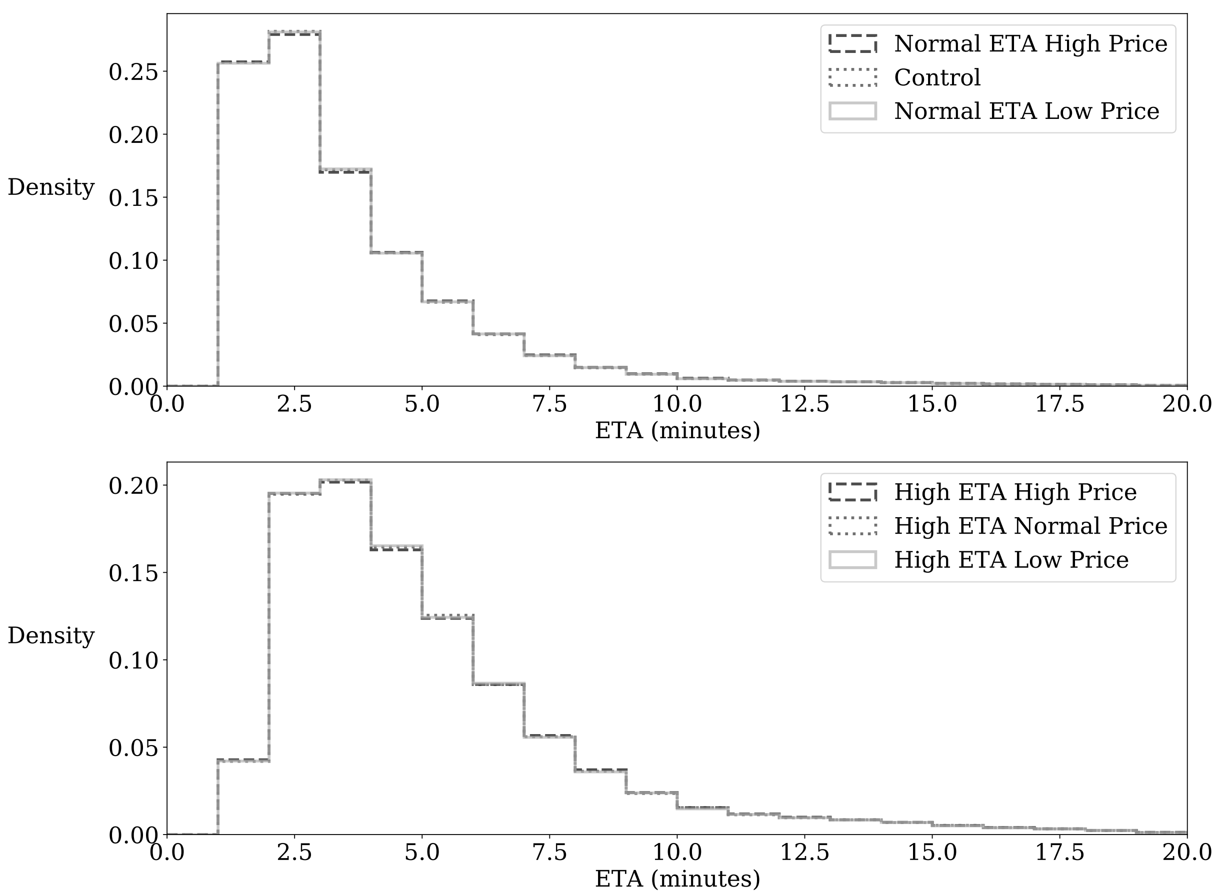


Figure 2
Distribution of ETAs across sessions in each exp group.

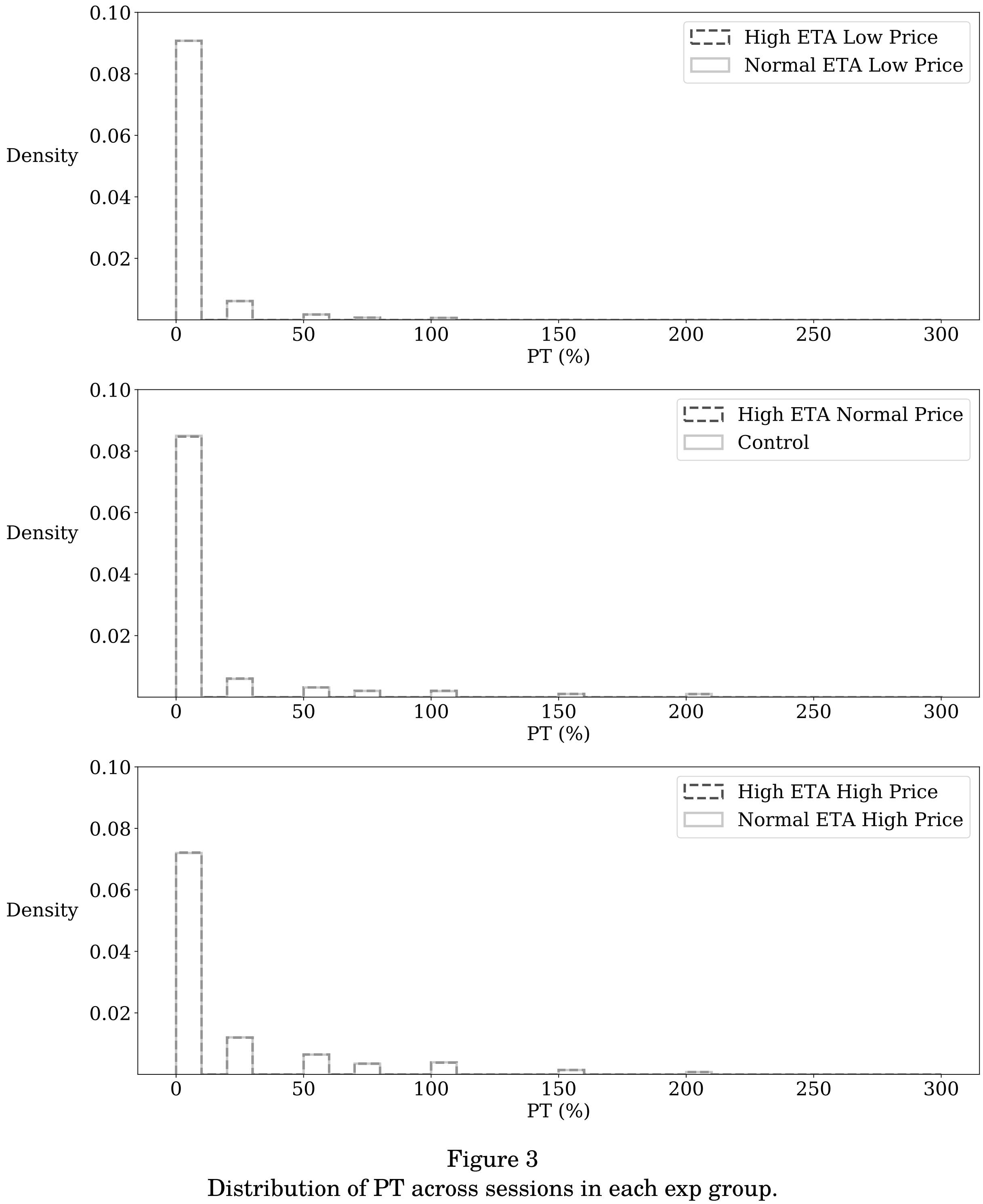


Figure 3
Distribution of PT across sessions in each exp group.

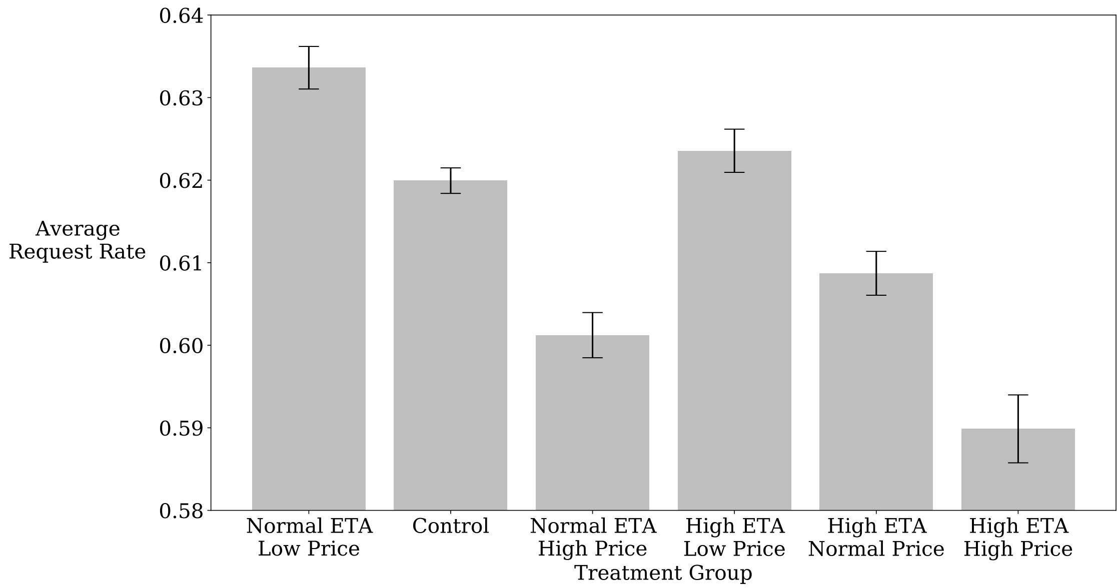


Figure 4
Average request rates across the six experimental groups. Error bars are 95% confidence intervals based on user-level clustered standard errors.

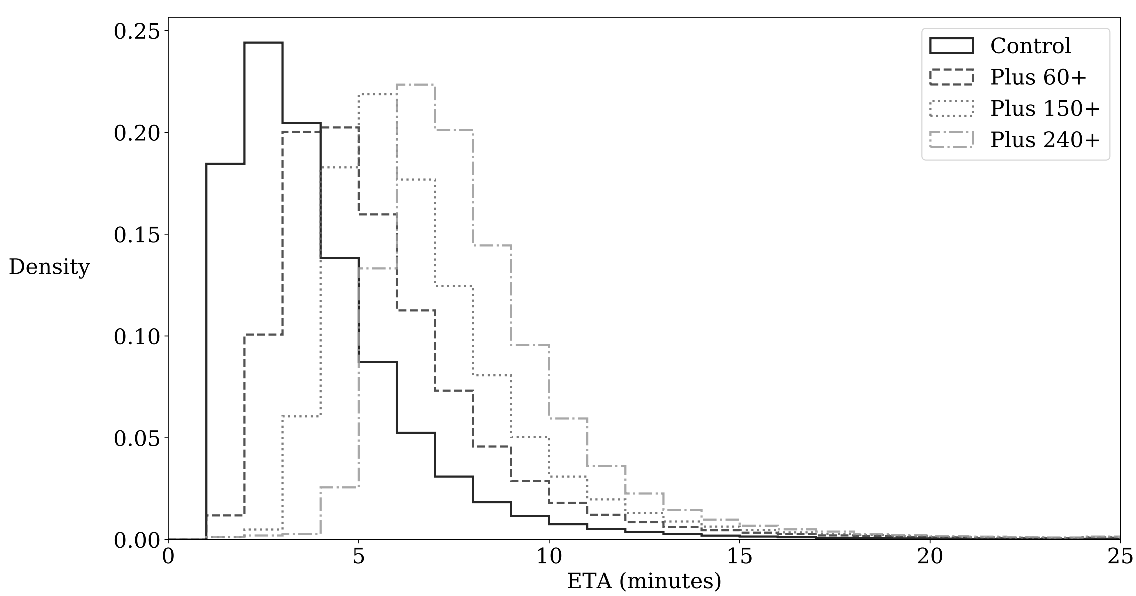


Figure 5
Distribution of ETAs across treatments in experiment two.

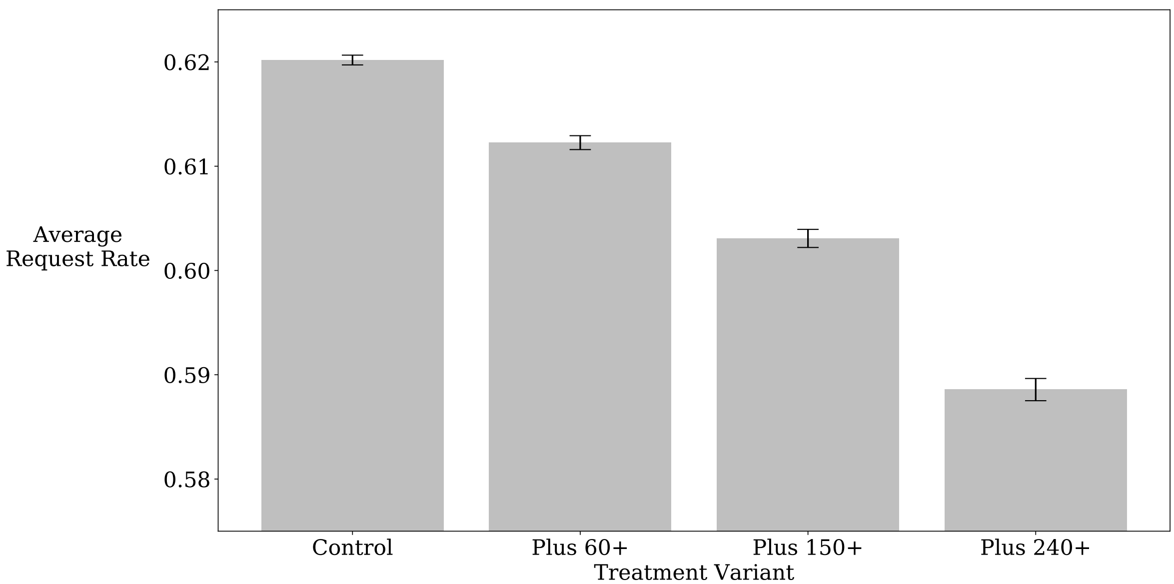


Figure 6
Average demand across the four treatments in experiment two.

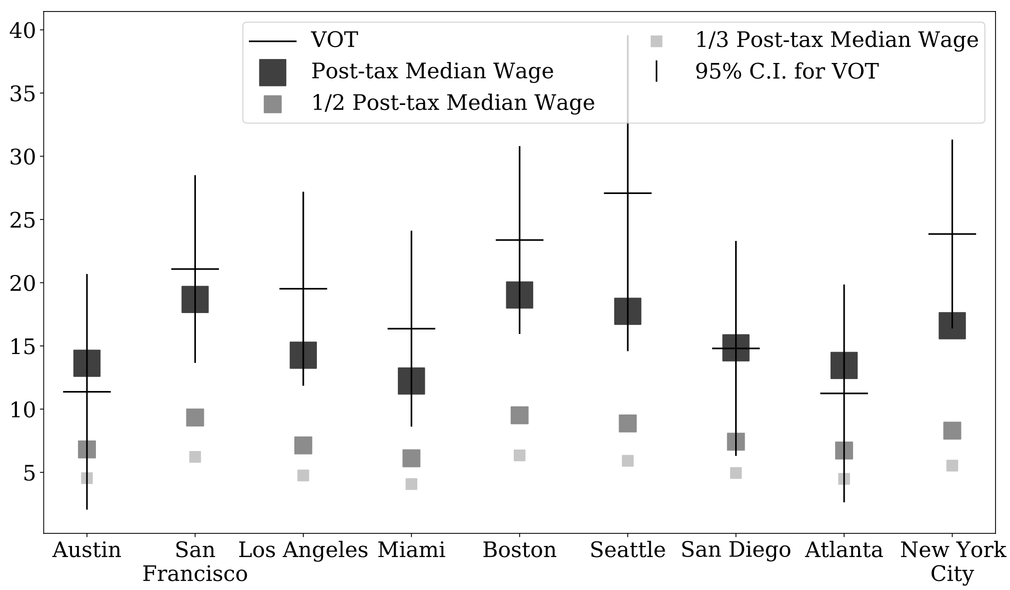


Figure 7
Estimated VOTs vs. mean and median wages for each region.

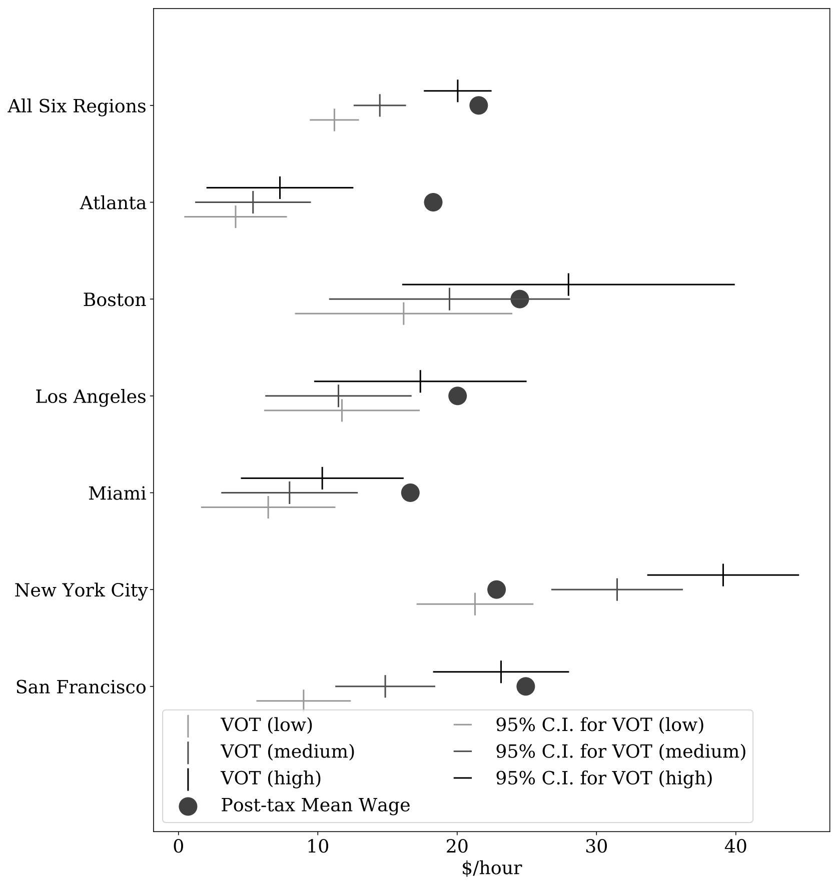


Figure 8
Values of time estimated by three different treatment level contrasts.

Table 1
Distribution of users across the six experimental groups.

	Low Price	**Normal Price**	**High Price**
Normal ETA	5%	15% (control)	5%
High ETA	5%	5%	2%

Notes: The remaining 63% of users were excluded from the experiment.

Table 2
Average session ETA and Prime Time by treatment group.

	Avg. ETA (minutes)	**Avg. PT (%)**	**Avg. Completed Ride Price ($)**	**Number of sessions**
Control	3.082 (0.004)	10.014 (0.032)	13.835 (0.022)	2,123,671
Normal ETA High Price	3.103 (0.008)	16.362 (0.072)	14.204 (0.041)	689,325
Normal ETA Low Price	3.085 (0.008)	3.646 (0.024)	13.544 (0.037)	722,367
High ETA Low Price	4.640 (0.011)	3.647 (0.025)	13.529 (0.038)	693,141
High ETA Normal Price	4.649 (0.011)	10.136 (0.055)	13.788 (0.039)	683,159
High ETA High Price	4.661 (0.017)	16.271 (0.112)	14.062 (0.064)	265,695

Notes: Standard errors in parentheses, clustered at the user level.

Table 3
First-stage regression results for full data sample.

	ln(ETA)	ln(1 + PT)	ln(ETA)	ln(1 + PT)
High ETA High Price	0.474^{***}	0.049^{***}	0.480^{***}	0.049^{***}
	(0.004)	(0.001)	(0.001)	(0.001)
High ETA Normal Price	0.473^{***}	0.001^{**}	0.479^{***}	0.001^{**}
	(0.003)	(0.000)	(0.001)	(0.000)
High ETA Low Price	0.472^{***}	-0.041^{***}	0.479^{***}	-0.042^{***}
	(0.003)	(0.000)	(0.001)	(0.000)
Normal ETA High Price	0.004	0.050^{***}	0.001	0.050^{***}
	(0.003)	(0.001)	(0.001)	(0.000)
Normal ETA Low Price	0.001	-0.041^{***}	0.001	-0.041^{***}
	(0.003)	(0.000)	(0.001)	(0.000)
Controls			x	x
N	5177358	5177358	5177358	5177358
R^2	0.101	0.026	0.534	0.212
F	13646^{***}	12784^{***}	116249^{***}	20455^{***}
Kleibergen and Paap (2006) rk-statistic	40175^{***}		48791^{***}	

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. Independent variables are indicators for each treatment group. Controls include dummy variables for week of year, hour of week, user geohash5, business, airport, and decile of pre-experiment lifetime rides. F statistics are scaled cluster-robust Wald statistics of the null that the coefficients on the treatment indicators are identically 0. Kleibergen and Paap (2006) statistics test underidentification of both endogenous regressors, and follow a $\chi^2(\text{df} = (5-1)(2-1))$ distribution under the null.

Table 4
Second-stage regression results for full data sample.

	(1)	(2)
ln(ETA)	-0.0231^{***}	-0.0264^{***}
	(0.0023)	(0.0018)
$\ln(1+\mathrm{PT})$	-0.3601^{***}	-0.3665^{***}
	(0.0165)	(0.0132)
ETA Elasticity	-0.0374^{***}	-0.0427^{***}
	(0.0025)	(0.0029)
PT Elasticity	-0.5838^{***}	-0.5942^{***}
	(0.0213)	(0.0213)
VOT	17.27^{***}	19.38^{***}
	(1.77)	(1.39)
Control Avg. ETA	3.08	3.08
Control Avg. Price	13.83	13.83
Control Request Rate	0.620	0.620
Controls		x
N	5177358	5177358
R^2	-0.008	0.072

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1 + \mathrm{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include dummy variables for session geohash5, local hour of week and week of year, decile of user lifetime rides, whether the session is at and airport, and whether the user is a business user.

Table 5
Summary of Elasticity and VOT Results: Features Affecting Opportunity Cost

Sub-sample	ETA Elasticity	PT Elasticity	VOT
Full Sample	-0.043^{***}	-0.594^{***}	19.38^{***}
	(0.003)	(0.021)	(1.39)
Day of Week			
Monday	-0.048^{***}	-0.627^{***}	20.08^{***}
	(0.005)	(0.046)	(2.47)
Tuesday	-0.046^{***}	-0.676^{***}	17.95^{***}
	(0.005)	(0.049)	(2.22)
Wednesday	-0.046^{***}	-0.653^{***}	19.34^{***}
	(0.005)	(0.048)	(2.38)
Thursday	-0.041^{***}	-0.585^{***}	19.44^{***}
	(0.005)	(0.035)	(2.41)
Friday	-0.039^{***}	-0.473^{***}	21.26^{***}
	(0.005)	(0.023)	(2.45)
Saturday	-0.039^{***}	-0.596^{***}	18.74^{***}
	(0.004)	(0.031)	(2.04)
Sunday	-0.040^{***}	-0.640^{***}	17.12^{***}
	(0.005)	(0.030)	(2.01)
Time of Day (Weekends)			
6–10 AM	-0.056^{***}	-0.687^{***}	17.40^{***}
	(0.011)	(0.081)	(3.80)
10 AM–4 PM	-0.053^{***}	-0.727^{***}	18.92^{***}
	(0.007)	(0.047)	(2.49)
4 PM–7 PM	-0.048^{***}	-0.616^{***}	20.16^{***}
	(0.007)	(0.039)	(2.92)
7 PM–11 PM	-0.033^{***}	-0.669^{***}	15.04^{***}
	(0.006)	(0.062)	(2.87)
11 PM–6 AM	-0.023^{***}	-0.500^{***}	14.71^{***}
	(0.005)	(0.028)	(3.12)
Time of Day (Weekdays)			
6–10 AM	-0.068^{***}	-0.594^{***}	26.71^{***}
	(0.006)	(0.034)	(2.50)
10 AM–4 PM	-0.047^{***}	-0.704^{***}	18.76^{***}
	(0.006)	(0.061)	(2.64)
4 PM–7 PM	-0.046^{***}	-0.522^{***}	22.12^{***}
	(0.005)	(0.034)	(2.70)
7 PM–11 PM	-0.024^{***}	-0.540^{***}	13.62^{***}
	(0.005)	(0.041)	(2.59)
11 PM–6 AM	-0.040^{***}	-0.539^{***}	19.03^{***}
	(0.006)	(0.031)	(2.66)

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table 6
Summary of Elasticity and VOT Results: Other Trip Characteristics

Sub-sample	ETA Elasticity	PT Elasticity	VOT
Full Sample	-0.043^{***}	-0.594^{***}	19.38^{***}
	(0.003)	(0.021)	(1.39)
Precipitation			
None	-0.042^{***}	-0.597^{***}	19.04^{***}
	(0.003)	(0.024)	(1.50)
Rain	-0.051^{***}	-0.560^{***}	23.10^{***}
	(0.005)	(0.026)	(2.26)
Snow	-0.103^{***}	-0.711^{***}	26.56^{***}
	(0.029)	(0.119)	(7.83)
User Type			
Business	-0.033^{**}	-0.399^{***}	24.42^{**}
	(0.014)	(0.073)	(10.55)
Non-Business	-0.043^{***}	-0.605^{***}	19.22^{***}
	(0.003)	(0.022)	(1.39)

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table 7
Summary of Elasticity and VOT Results: Market Factors

Sub-sample	ETA Elasticity	PT Elasticity	VOT
Full Sample	-0.043^{***}	-0.594^{***}	19.38^{***}
	(0.003)	(0.021)	(1.39)
Region			
Austin	-0.028^{**}	-0.709^{***}	11.37^{**}
	(0.012)	(0.161)	(4.76)
San Francisco	-0.026^{***}	-0.413^{***}	21.08^{***}
	(0.005)	(0.026)	(3.79)
Los Angeles	-0.040^{***}	-0.522^{***}	19.52^{***}
	(0.006)	(0.071)	(3.91)
Miami	-0.052^{***}	-0.595^{***}	16.36^{***}
	(0.010)	(0.103)	(3.95)
Boston	-0.074^{***}	-0.670^{***}	23.37^{***}
	(0.011)	(0.072)	(3.79)
Seattle	-0.053^{***}	-0.516^{***}	27.08^{***}
	(0.011)	(0.077)	(6.38)
San Diego	-0.044^{***}	-0.723^{***}	14.80^{***}
	(0.011)	(0.152)	(4.34)
Atlanta	-0.056^{***}	-0.825^{***}	11.24^{**}
	(0.017)	(0.256)	(4.40)
New York City	-0.056^{***}	-0.956^{***}	23.85^{***}
	(0.009)	(0.046)	(3.81)
Location Type			
Airport	-0.024	-0.553^{***}	27.55
	(0.015)	(0.153)	(18.02)
Non-Airport	-0.043^{***}	-0.595^{***}	18.84^{***}
	(0.003)	(0.021)	(1.35)
Downtown	-0.034^{***}	-0.524^{***}	21.64^{***}
	(0.003)	(0.019)	(1.99)
Non-Downtown	-0.060^{***}	-0.762^{***}	17.88^{***}
	(0.005)	(0.058)	(1.97)
Distance to Nearest Public Transit			
Under 50 Meters	-0.039^{***}	-0.495^{***}	23.81^{***}
	(0.004)	(0.026)	(2.56)
50 to 200 Meters	-0.040^{***}	-0.528^{***}	20.81^{***}
	(0.003)	(0.024)	(1.91)
200 to 800 Meters	-0.051^{***}	-0.802^{***}	16.90^{***}
	(0.006)	(0.049)	(2.03)
Over 800 Meters	-0.058^{***}	-1.162^{***}	10.88^{***}
	(0.014)	(0.161)	(2.82)

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table 8
Summary of Elasticity Results: Heterogeneity (2017 Experiment)

Time of Week	(1)	Trip Characteristics	(2)	Location	(3)
Full Sample	-0.043^{***}	**Full Sample**	-0.043^{***}	**Full Sample**	-0.043^{***}
	(0.001)		(0.001)		(0.001)
Day of Week		**Precipitation**		**Region**	
Monday	-0.044^{***}	None	-0.042^{***}	San Francisco	-0.031^{***}
	(0.002)		(0.001)		(0.001)
Tuesday	-0.047^{***}	Rain	-0.047^{***}	New York City	-0.086^{***}
	(0.002)		(0.002)		(0.002)
Wednesday	-0.044^{***}			Chicago	-0.039^{***}
	(0.002)	**User Type**			(0.002)
Thursday	-0.047^{***}	Business	-0.049^{***}	Washington, D.C.	-0.054^{***}
	(0.002)		(0.002)		(0.002)
Friday	-0.042^{***}	Non-Business	-0.043^{***}	Miami	-0.021^{***}
	(0.002)		(0.001)		(0.002)
Saturday	-0.039^{***}			New Jersey	-0.053^{***}
	(0.002)				(0.004)
Sunday	-0.040^{***}			Boston	-0.055^{***}
	(0.002)				(0.003)
				Philadelphia	-0.040^{***}
Time of Day (Weekends)					(0.003)
6–10 AM	-0.060^{***}			Atlanta	-0.037^{***}
	(0.004)				(0.004)
10 AM–4 PM	-0.048^{***}			Los Angeles	-0.033^{***}
	(0.003)				(0.002)
4 PM–7 PM	-0.044^{***}				
	(0.003)			**Location Type**	
7 PM–11 PM	-0.033^{***}			Downtown	-0.039^{***}
	(0.003)				(0.001)
11 PM–6 AM	-0.026^{***}			Non-Downtown	-0.048^{***}
	(0.002)				(0.001)
Time of Day (Weekdays)				**Distance to Public Transit**	
6–10 AM	-0.061^{***}			Under 50 Meters	-0.042^{***}
	(0.002)				(0.001)
10 AM–4 PM	-0.041^{***}			50 to 200 Meters	-0.038^{***}
	(0.002)				(0.001)
4 PM–7 PM	-0.052^{***}			200 to 800 Meters	-0.049^{***}
	(0.002)				(0.002)
7 PM–11 PM	-0.035^{***}			Over 800 Meters	-0.084^{***}
	(0.002)				(0.004)
11 PM–6 AM	-0.037^{***}				
	(0.002)				

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table 9
Time elasiticities estimated by three different levels of treatment contrasts, across different base ETA levels (minutes) (2017 experiment).

	1	2	3	4	5	6	7	8	9	10
Control — Plus 60	-0.005^{***}	-0.014^{***}	-0.026^{***}	-0.058^{***}	-0.087^{***}	-0.124^{***}	-0.142^{***}	-0.188^{***}	-0.208^{***}	-0.230^{***}
	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)	(0.006)	(0.007)
Control — Plus 150	-0.008^{***}	-0.025^{***}	-0.044^{***}	-0.063^{***}	-0.103^{***}	-0.118^{***}	-0.166^{***}	-0.182^{***}	-0.203^{***}	-0.289^{***}
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.005)	(0.004)
Control — Plus 240	-0.018^{***}	-0.042^{***}	-0.066^{***}	-0.075^{***}	-0.108^{***}	-0.137^{***}	-0.170^{***}	-0.217^{***}	-0.232^{***}	-0.289^{***}
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\text{ETA})$ instrumented by experimental group indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table 10
Time elasiticities estimated by three different treatments.

	(1)
Control — Plus 60	-0.027^{***}
	(0.001)
Control — Plus 150	-0.037^{***}
	(0.001)
Control — Plus 240	-0.053^{***}
	(0.001)

Notes: Clustered standard errors in parentheses. $\ln(\text{ETA})$ instrumented by experimental group indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table 11
Convexity by city.

	San Francisco	New York City	Chicago	D.C.	Miami	New Jersey	Boston	Philadelphia	Atlanta	Los Angeles
Control — Plus 60	-0.015^{***}	-0.055^{***}	-0.017^{***}	-0.026^{***}	-0.015^{***}	-0.041^{***}	-0.034^{***}	-0.033^{***}	-0.023^{***}	-0.026^{***}
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.006)	(0.002)
Control — Plus 150	-0.025^{***}	-0.081^{***}	-0.035^{***}	-0.048^{***}	-0.019^{***}	-0.043^{***}	-0.044^{***}	-0.031^{***}	-0.031^{***}	-0.026^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.002)	(0.004)	(0.005)	(0.002)
Control — Plus 240	-0.039^{***}	-0.101^{***}	-0.048^{***}	-0.071^{***}	-0.025^{***}	-0.061^{***}	-0.067^{***}	-0.050^{***}	-0.047^{***}	-0.040^{***}
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.004)	(0.002)

Notes: Clustered standard errors in parentheses. $\ln(\text{ETA})$ instrumented by experimental group indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table 12
Number of regions for which each hypothesis test rejects the null that the VOT equals the given rule-of-thumb approximation at the 5% level, using post-tax wages and assuming a constant 25% tax rate.

1/3- Rule

	Median	Mean
Log	7	7
Linear	7	7

1/2- Rule

	Median	Mean
Log	7	6
Linear	6	6

100%+ Rule

	Median	Mean
Log	0	0
Linear	0	4

1/2+ Rule

	Median	Mean
Log	0	0
Linear	0	0

Table 13
Number of regions for which each hypothesis test rejects the null that the VOT equals the given rule-of-thumb approximation at the 5% level, using pre-tax wages.

1/3- Rule

	Median	Mean
Log	7	6
Linear	7	6

1/2- Rule

	Median	Mean
Log	6	3
Linear	6	2

100%+ Rule

	Median	Mean
Log	0	7
Linear	4	8

1/2+ Rule

	Median	Mean
Log	0	0
Linear	0	0

Table 14
Regressions of log estimated VOT on log mean wage at the region level.

	OLS	GLS
ln(MeanWage)	1.73^{**}	1.28^{**}
	(0.82)	(0.54)
N	9	9
R^2	0.586	0.534
Adj. R^2	0.527	0.467

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Heteroskedasticity-robust (HC3) standard errors in parentheses. Dependent variable is $\ln(\widehat{VOT})$.

ONLINE APPENDIX

The Value of Time in the United States

Ariel Goldszmidt, John A. List, Robert D. Metcalfe,
Ian Muir, V. Kerry Smith, Jenny Wang
October 12, 2020

Contents

A Derivations 3

B Additional Figures 5

C Additional Tables 15

D Lateness regressions 45

E Demand Effects After End of Experiment 48

F Time to Destination 51

G Removing always- and/or never-requesters 53

H External Validity 62

I Correcting for User Selection into Ride Contexts 68

J Reweighting Elasticity Estimates 79

K Ride Types 86

L Distribution of VOT 89

M Causality with Interaction Effects 91

A Derivations

Equation (2). The utility maximization problem is

$$\begin{aligned}
\max_{Z_1,Z_2} \quad & u(Z_1, Z_2) \\
\text{subject to} \quad & p_1x_1 + p_2x_2 = wT_w + R \\
& x_i = a_iZ_i, i = 1, 2 \\
& T_1 + T_2 + T_w = \bar{T} \\
& T_i = t_iZ_i, i = 1, 2.
\end{aligned}$$

By substituting the second constraint into the first and the fourth into the third, the problem may be rewritten as

$$\begin{aligned}
\max_{Z_1,Z_2} \quad & u(Z_1, Z_2) \\
\text{subject to} \quad & p_1a_1Z_1 + p_2a_2Z_2 = wT_w + R \\
& t_1Z_1 + t_2Z_2 + T_w = \bar{T}.
\end{aligned}$$

Solving for T_w from the second constraint and plugging it into the first, we have

$$\begin{aligned}
\max_{Z_1,Z_2} \quad & u(Z_1, Z_2) \\
\text{subject to} \quad & (p_1a_1 + wt_1)Z_1 + (p_2a_2 + wt_2)Z_2 = w\bar{T} + R.
\end{aligned}$$

The utility maximization problem is now in the standard form, with the (full) prices of Z_1 and Z_2 given by $p_1a_1 + wt_1$ and $p_2a_2 + wt_2$ and the (full) income given by $w\bar{T} + R$. Hence the indirect utility function that results from solving this problem can be generically written as $\overline{V} = \overline{V}(p_1a_1 + wt_1, p_2a_2 + wt_2, w\bar{T} + R)$.

Equation (3). Let $P_i = P_i(w) = p_ia_i + wt_i$ denote the (full) price of consuming one unit of Z_i, which is a function of w, and let $M = M(w, R) = w\bar{T} + R$. Since $\overline{V} = \overline{V}(P_1, P_2, M)$, the chain rule gives

$$\begin{aligned}
\frac{d\overline{V}}{dw} &= \overline{V}_{P_1}(P_1)_w + \overline{V}_{P_2}(P_2)_w + \overline{V}_M M_w \\
&= \overline{V}_{P_1}t_1 + \overline{V}_{P_1}t_2 + \overline{V}_M\bar{T}
\end{aligned}$$

and

$$\frac{d\overline{V}}{dR} = \overline{V}_M M_R = \overline{V}_M.$$

Dividing the first equation by $\overline{V}_M$ on the right and by $\frac{d\overline{V}}{dR}$ on the left, we have

$$\frac{\frac{d\overline{V}}{dw}}{\frac{d\overline{V}}{dR}} = \frac{\overline{V}_{P_1}}{\overline{V}_M} t_1 + \frac{\overline{V}_{P_2}}{\overline{V}_M} t_2 + \bar{T}.$$

Finally, we apply Roy's identity $\frac{\overline{V}_{P_i}}{\overline{V}_M} = -Z_i^*$ to conclude

$$\frac{\frac{d\overline{V}}{dw}}{\frac{d\overline{V}}{dR}} = -Z_1^* t_1 + -Z_2^* t_2 + \bar{T}.$$

Equation (6). Starting from the modified indirect utility function in equation (5), we use the chain rule to calculate

$$\begin{aligned}\frac{dV}{dT_1^a} &= V_1 \cdot (wt_1 + p_1 a_1(T_1^a))_{T_1^a} + V_2 \cdot (wt_2 + p_2 a_2)_{T_1^a} + V_3 \cdot (w(\bar{T} - T_1^a) + R)_{T_1^a} \\ &= V_1 p_1 a_1'(T_1^a) + V_3(-w)\end{aligned}$$

and

$$\begin{aligned}\frac{dV}{dR} &= V_1 \cdot (wt_1 + p_1 a_1(T_1^a))_R + V_2 \cdot (wt_2 + p_2 a_2)_R + V_3 \cdot (w(\bar{T} - T_1^a) + R)_R \\ &= V_3,\end{aligned}$$

where V_i denotes the partial derivative of the function V with respect to its ith argument. We now divide the left-hand side of the first equation by $\frac{dV}{dR}$, the first term on the right-hand side by $\frac{dV}{dR}$, and the second term on the right-hand side by V_3 to conclude

$$\frac{\frac{dV}{dT_1^a}}{\frac{dV}{dR}} = \frac{V_1}{\frac{dV}{dR}} p_1 a_1'(T_1^a) - w.$$

Equation (7). $\pi = -\frac{dV}{dT_1^a} / \frac{dV}{dR}$, the (negative) marginal rate of substitution between waiting time and exogenous income, by definition. From equation (6), we have

$$\pi = -\frac{\frac{dV}{dT_1^a}}{\frac{dV}{dR}} = -\frac{V_1}{\frac{dV}{dR}} p_1 a_1'(T_1^a) + w.$$

Recalling that $\frac{dV}{dR} = V_3$ (from the derivation of (6)) and $V_1/V_3 = -Z_1^*$ (from Roy's identity), we have

$$\pi = -\frac{\frac{dV}{dT_1^a}}{\frac{dV}{dR}} = p_1 Z_1^* a_1'(T_1^a) + w.$$

B Additional Figures

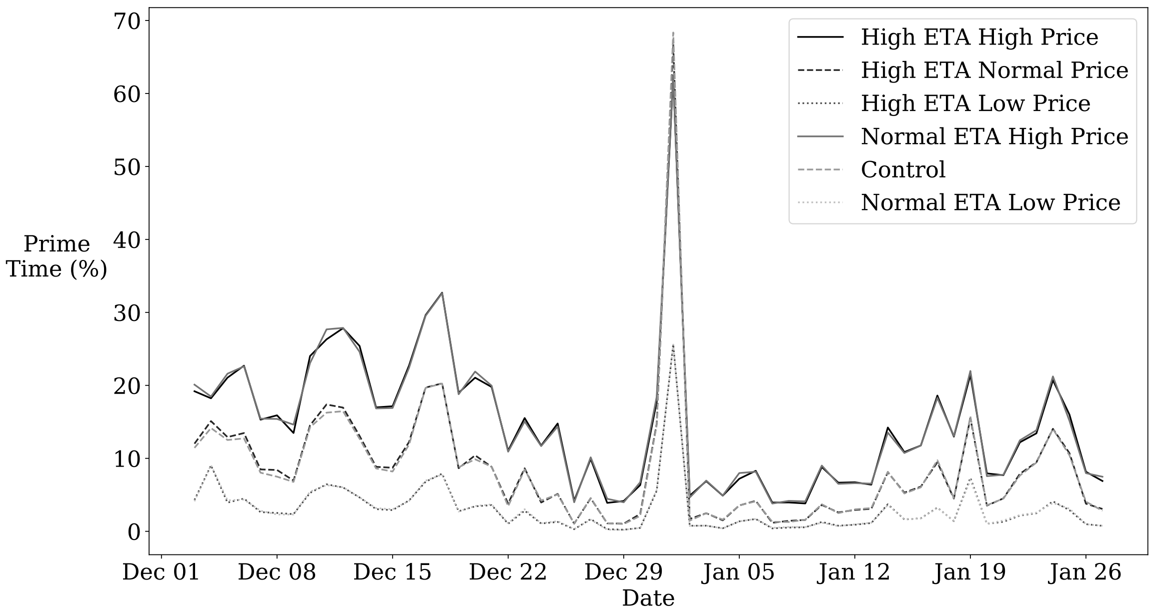


Figure B.1
Mean Prime Time in each treatment group by each day in experiment.

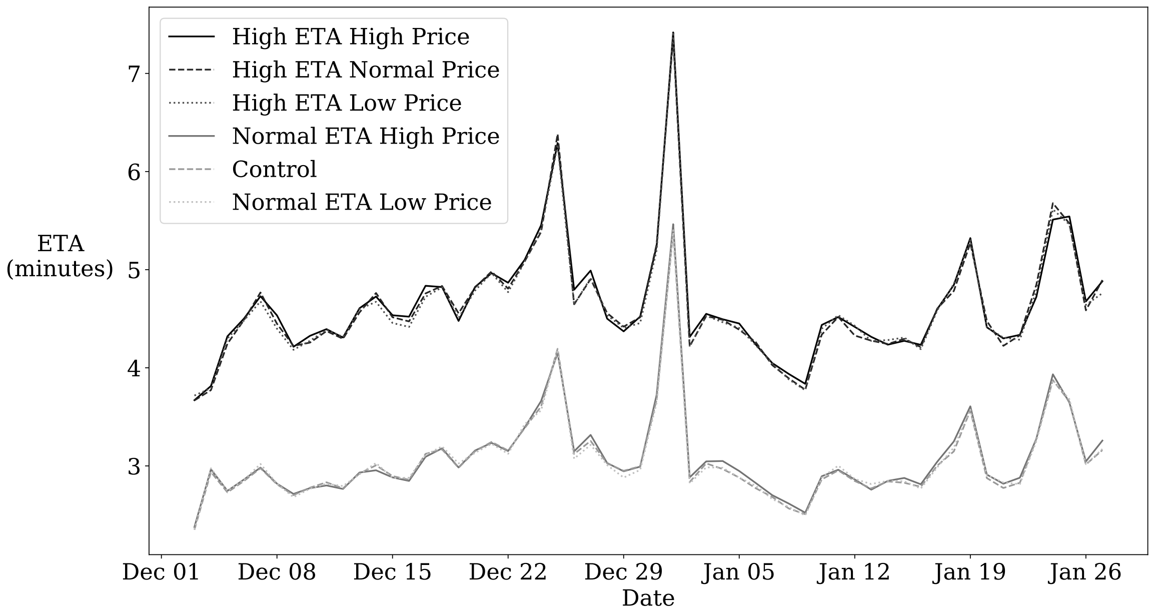


Figure B.2
Mean ETA in each treatment group by each day in experiment.

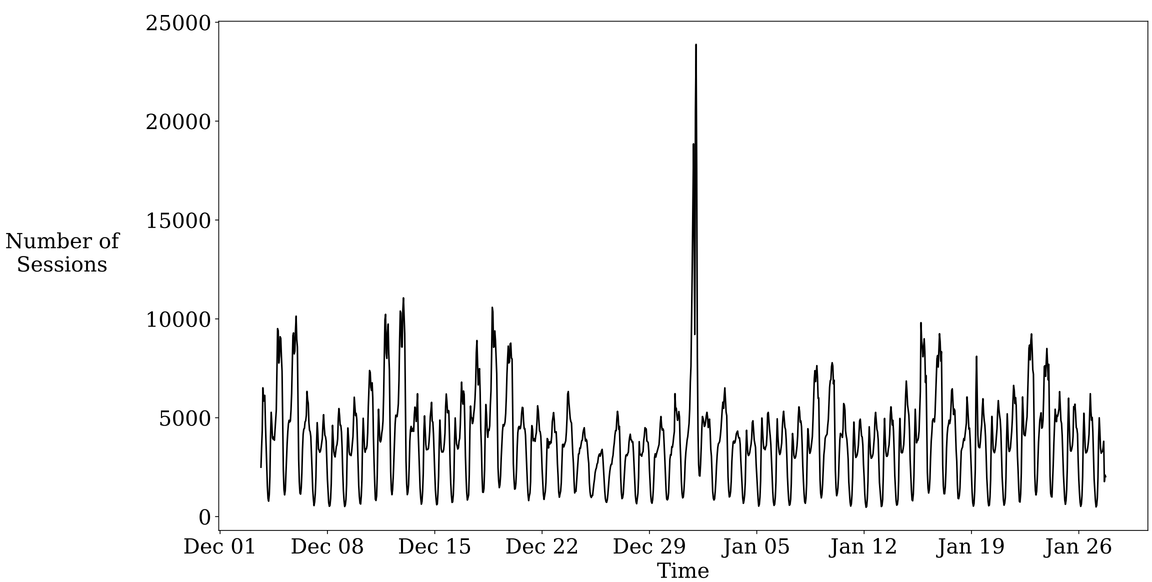


Figure B.3
Total number of sessions in each local hour of the experiment. Note the daily and weekly seasonality, and exceptional behavior near holidays. The large vertical spike is midnight of New Year's Eve.

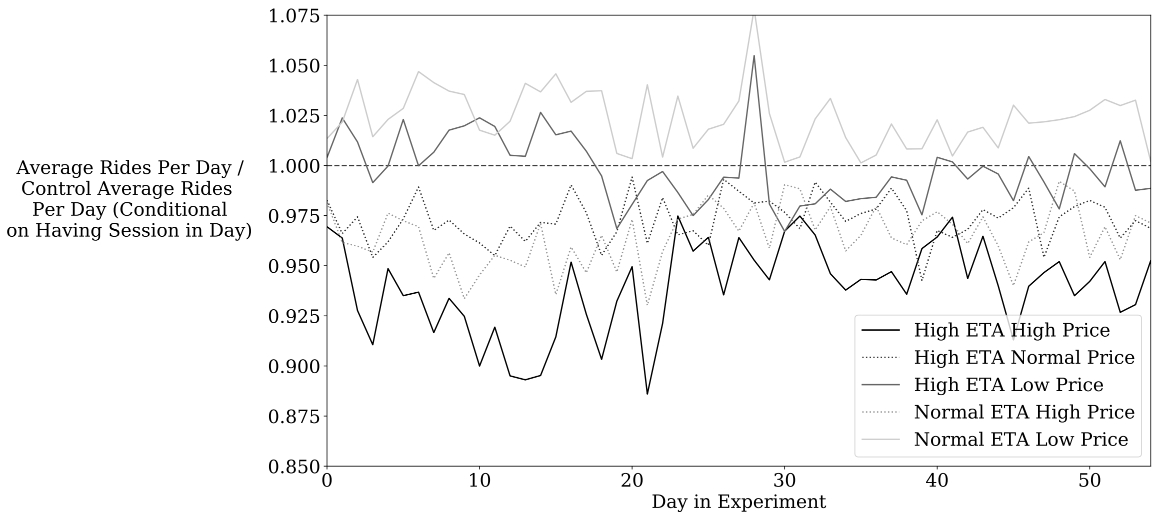


Figure B.4
Each series is the average number of rides taken by a passenger in that treatment group on each day from their first session in the experiment, divided by the average number of rides taken by a control passenger.

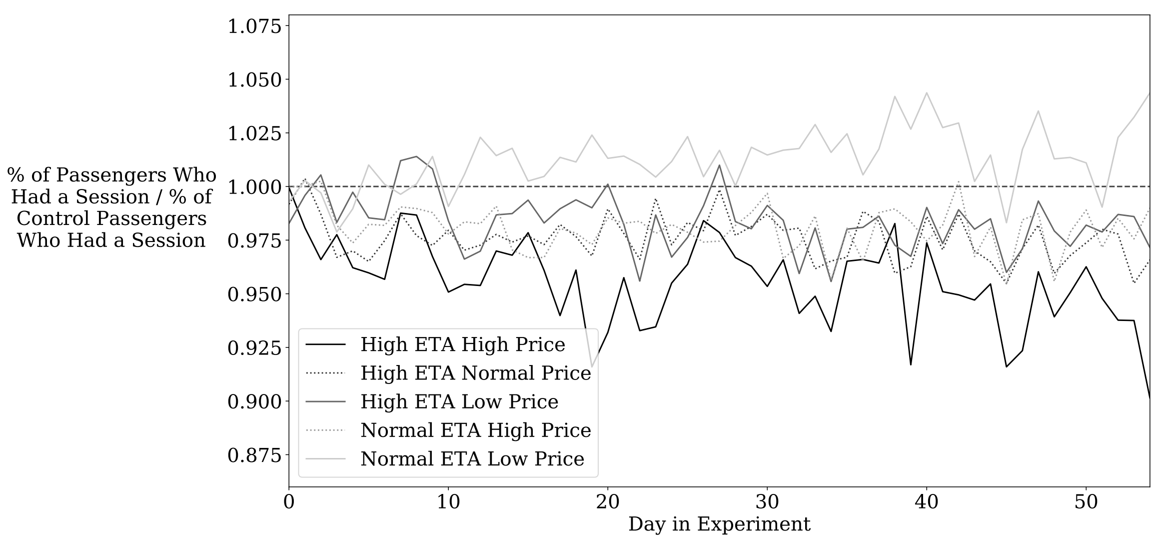
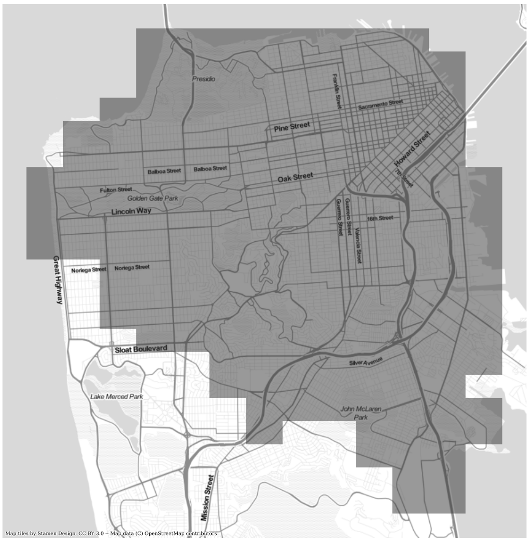
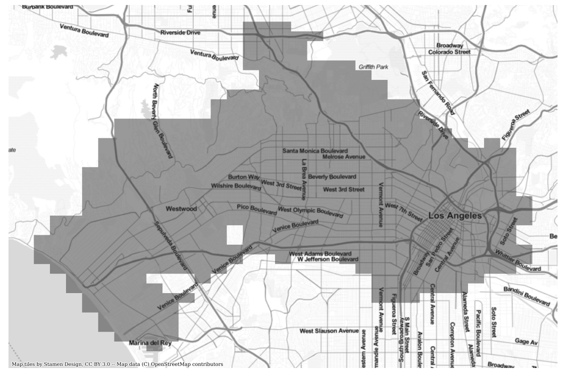


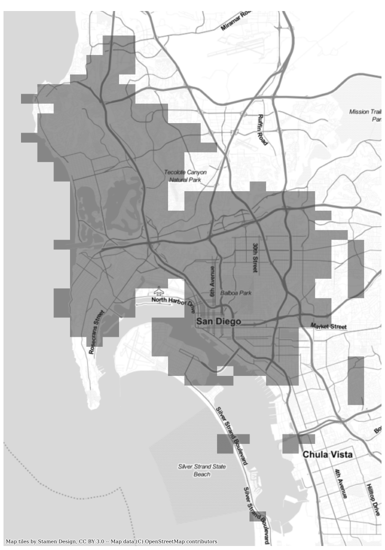
Figure B.5
Each series is the percent of passengers in that treatment group who had a session on that day, divided by the number of control passenger who had a session that day.



San Francisco



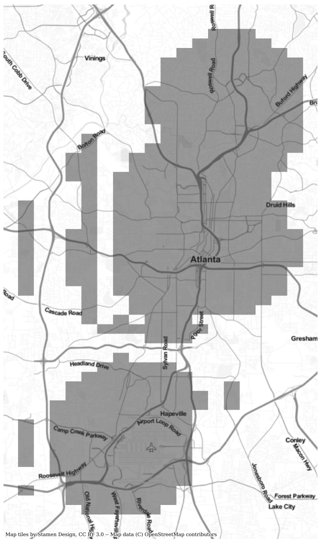
Los Angeles



San Digeo

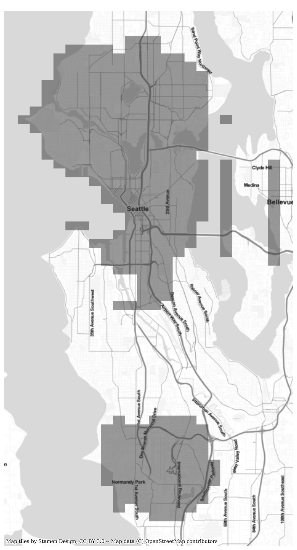
Austin

Figure B.6
Areas classified as "downtown." Maps from OpenStreetMaps.



Atlanta

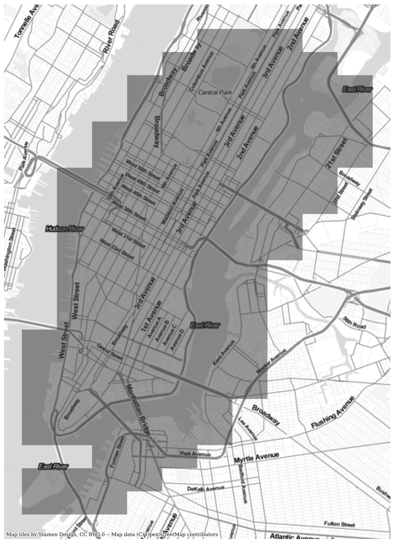
Boston



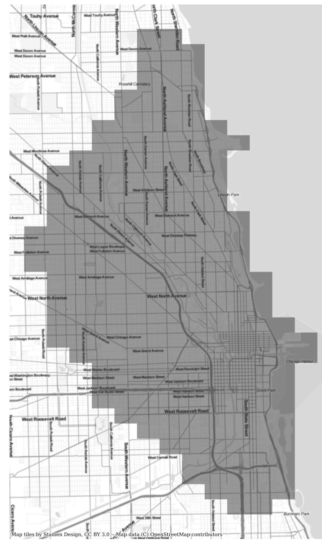
Seattle

Miami

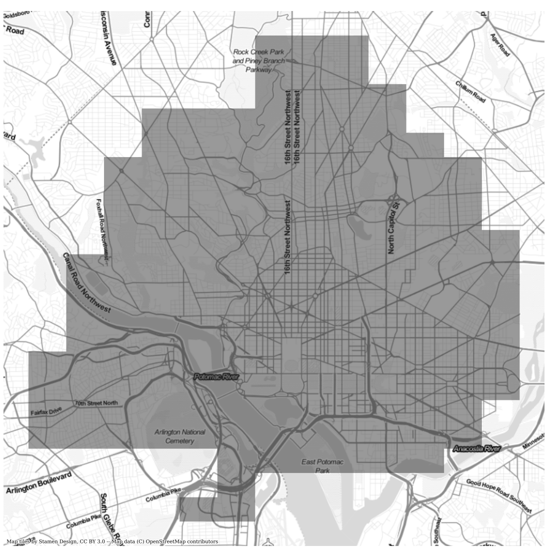
Figure B.6
Areas classified as “downtown.” Maps from OpenStreetMaps.



New York City



Chicago



Washington, D.C.

Philadelphia

Figure B.6
Areas classified as “downtown.” Maps from OpenStreetMaps.

New Jersey

Figure B.6
Areas classified as “downtown.” Maps from OpenStreetMaps.

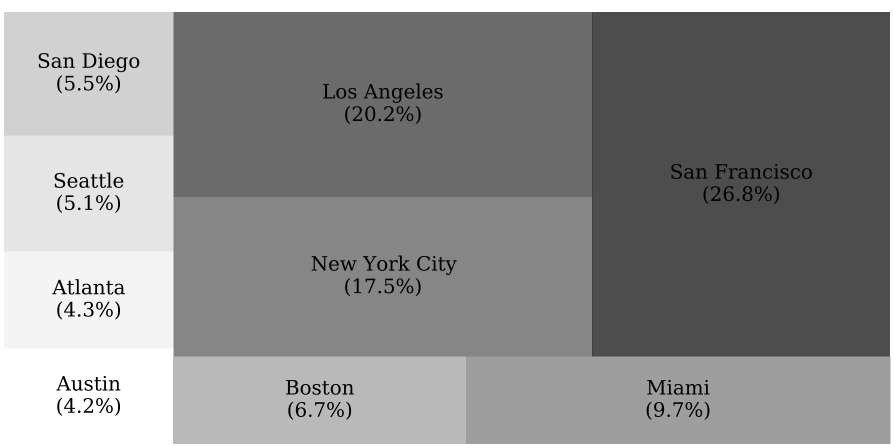


Figure B.7
Distribution of regions across sessions. Area of each square is proportional to the number of sessions in the experiment from that region.

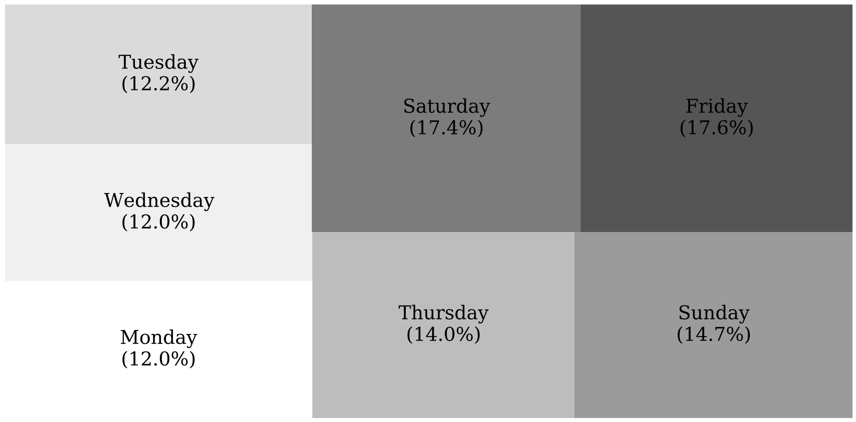


Figure B.8
Distribution of days of the week across sessions. Area of each square is proportional to the number of sessions in the experiment from that day of the week.

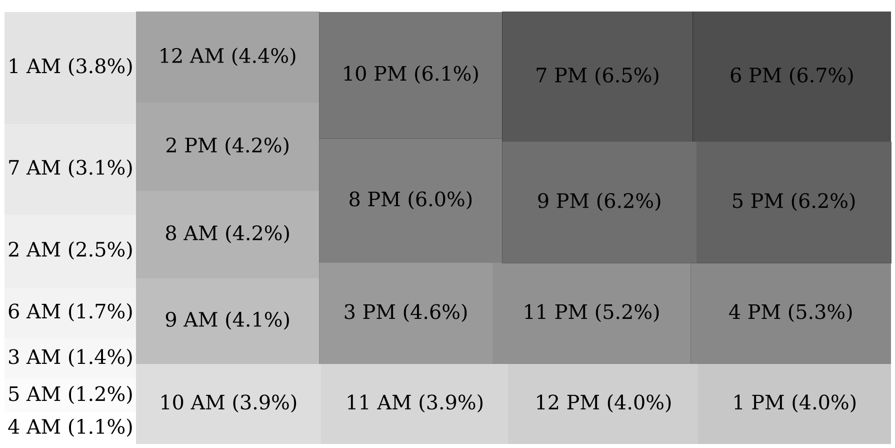


Figure B.9
Distribution of hours of the day across sessions. Area of each square is proportional to the number of sessions in the experiment from that hour of the day.

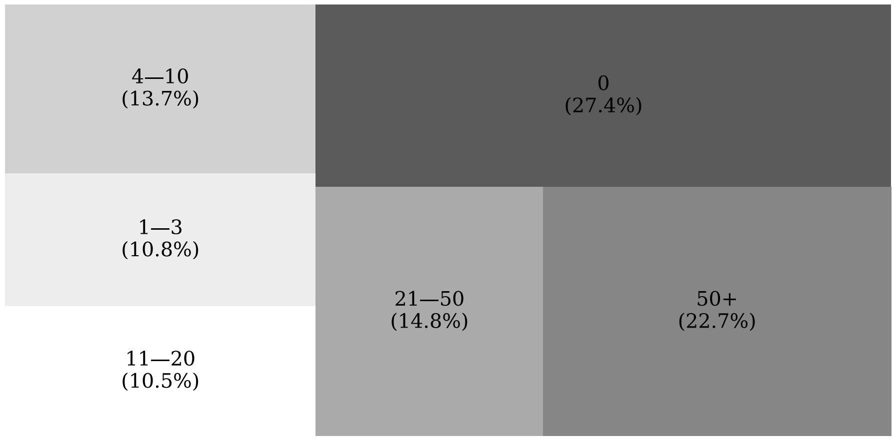


Figure B.10
Distribution of experience levels across sessions. Area of each square is proportional to the number of sessions sessions from passengers with the given number of lifetime rides, before the start of the experiment.

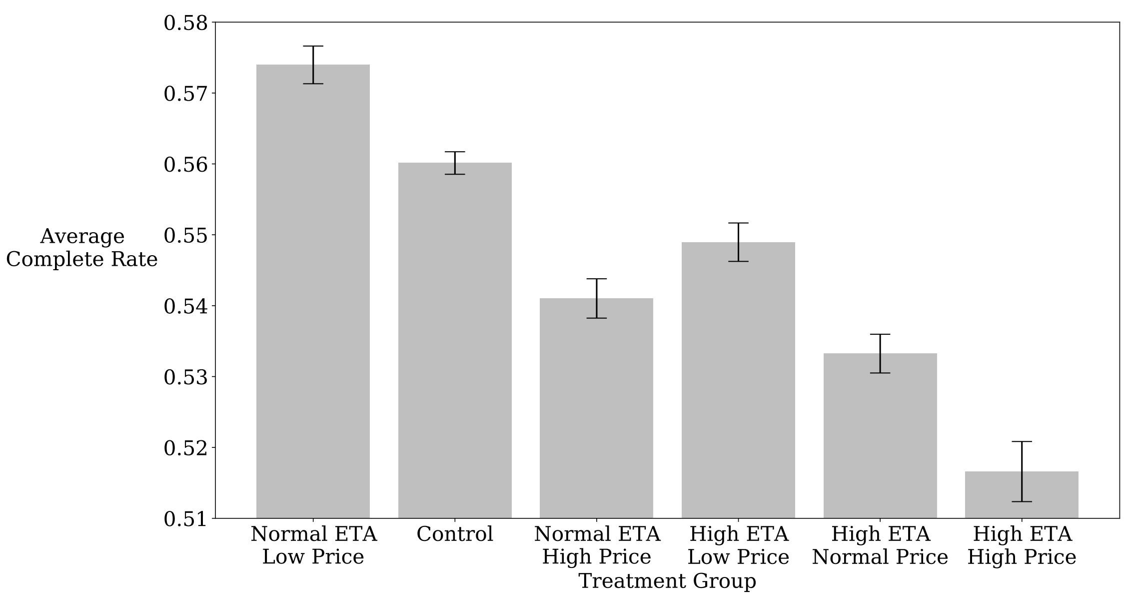


Figure B.11
Average completion rates across the six experimental groups. Error bars are 95% confidence intervals based on user-level clustered standard errors.

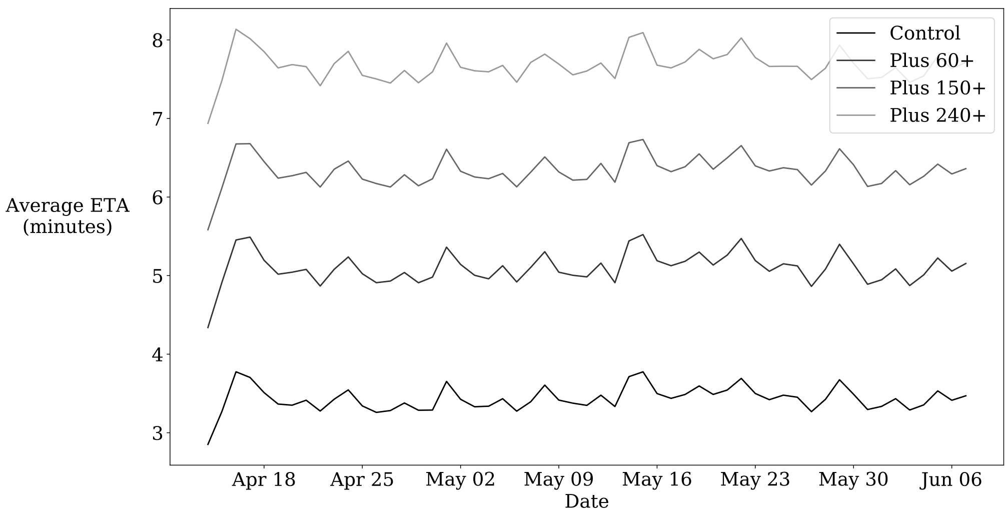


Figure B.12
Timeline of ETAs across experimental treatments in experiment two.

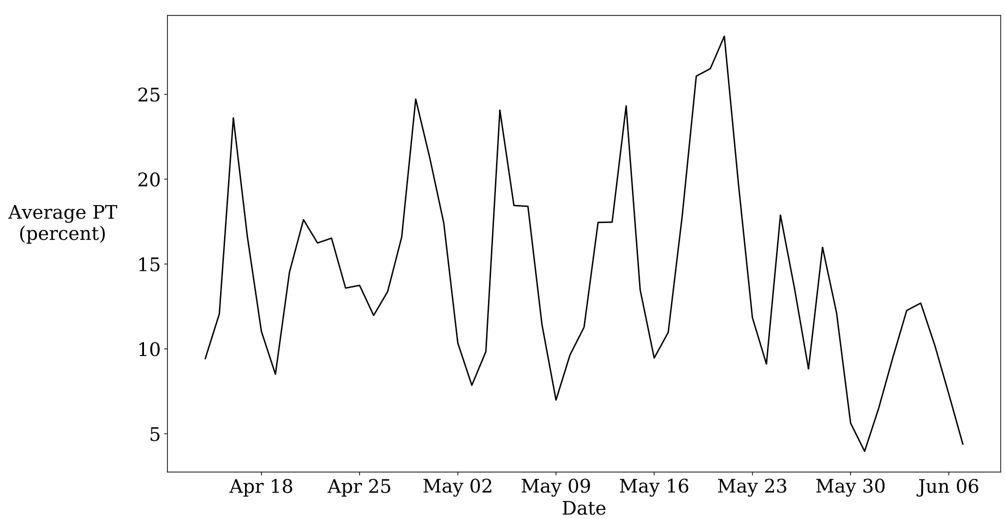


Figure B.13
Timeline of PT in experiment two.

C Additional Tables

Table C.1
Average number of sessions, sessions with requests, and sessions with completed rides per passenger in the past 28 days and in the passenger's lifetime, in each treatment group. Standard errors in parentheses.

	Control	High ETA High Price	High ETA Low Price	High ETA Normal Price	Normal ETA High Price	Normal ETA Low Price
# Sessions (28 days)	3.39 (0.012)	3.37 (0.033)	3.41 (0.021)	3.4 (0.021)	3.42 (0.021)	3.4 (0.021)
# Sessions (lifetime)	27.14 (0.102)	27.29 (0.28)	27.23 (0.179)	27.2 (0.179)	27.44 (0.18)	27.19 (0.179)
# Sessions w/ Request (28 days)	2.2 (0.009)	2.18 (0.025)	2.21 (0.016)	2.21 (0.016)	2.22 (0.016)	2.21 (0.016)
# Sessions w/ Request (lifetime)	17.95 (0.08)	18.01 (0.217)	17.99 (0.14)	17.95 (0.14)	18.15 (0.139)	18.0 (0.141)
# Sessions w/ Complete (28 days)	1.98 (0.008)	1.97 (0.023)	1.99 (0.015)	1.98 (0.014)	1.99 (0.015)	1.98 (0.014)
# Sessions w/ Complete (lifetime)	15.65 (0.073)	15.75 (0.198)	15.7 (0.127)	15.66 (0.127)	15.86 (0.127)	15.69 (0.129)

Table C.2
Average number of sessions, sessions with requests, and sessions with completed rides for each passenger in the experiment, by treatment group. Standard errors in parentheses.

	Control	High ETA High Price	High ETA Low Price	High ETA Normal Price	Normal ETA High Price	Normal ETA Low Price
# Passengers	292025	38674	97254	97051	97185	97870
# Sessions	7.27 (0.019)	6.87 (0.048)	7.13 (0.032)	7.04 (0.032)	7.09 (0.032)	7.38 (0.033)
# Sessions w/ request	4.51 (0.015)	4.05 (0.037)	4.44 (0.025)	4.28 (0.025)	4.26 (0.024)	4.68 (0.026)
# Sessions w/ complete	4.07 (0.014)	3.55 (0.034)	3.91 (0.023)	3.75 (0.023)	3.84 (0.023)	4.24 (0.025)

Table C.3
p-values from pairwise Kolmogorov-Smirnov tests comparing the distributions of of average ETA between users in different experimental treatment groups.

	Normal ETA Low Price	High ETA Low Price	Control	High ETA Normal Price	Normal ETA High Price	High ETA High Price
Normal ETA Low Price	1.0000	0.0000	0.4192	0.0000	0.0629	0.0000
High ETA Low Price	0.0000	1.0000	0.0000	0.7931	0.0000	0.9990
Control	0.4192	0.0000	1.0000	0.0000	0.2579	0.0000
High ETA Normal Price	0.0000	0.7931	0.0000	1.0000	0.0000	0.6178
Normal ETA High Price	0.0629	0.0000	0.2579	0.0000	1.0000	0.0000
High ETA High Price	0.0000	0.9990	0.0000	0.6178	0.0000	1.0000

Table C.4
p-values from pairwise Kolmogorov-Smirnov tests comparing the distributions of average PT between users in different experimental treatment groups.

	Normal ETA Low Price	**High ETA Low Price**	**Control**	**High ETA Normal Price**	**Normal ETA High Price**	**High ETA High Price**
Normal ETA Low Price	1.0000	0.0392	0.0000	0.0000	0.0000	0.0000
High ETA Low Price	0.0392	1.0000	0.0000	0.0000	0.0000	0.0000
Control	0.0000	0.0000	1.0000	0.1199	0.0000	0.0000
High ETA Normal Price	0.0000	0.0000	0.1199	1.0000	0.0000	0.0000
Normal ETA High Price	0.0000	0.0000	0.0000	0.0000	1.0000	0.1351
High ETA High Price	0.0000	0.0000	0.0000	0.0000	0.1351	1.0000

Table C.5
Distribution of Prime Time multipliers across sessions for each experimental treatment group.

	Prime Time Level (%)								
	0%	**25%**	**50%**	**75%**	**100%**	**150%**	**200%**	**250%**	**300%**
Control	84.981	5.932	3.132	2.005	1.941	1.000	1.004	0.003	0.001
High ETA High Price	72.089	11.984	6.451	3.481	3.844	1.413	0.737	0.000	0.000
High ETA Low Price	90.755	6.108	1.743	0.708	0.644	0.026	0.018	0.000	0.000
High ETA Normal Price	84.789	6.005	3.138	2.032	2.025	1.021	0.984	0.004	0.002
Normal ETA High Price	71.984	12.031	6.416	3.497	3.914	1.412	0.746	0.000	0.000
Normal ETA Low Price	90.750	6.108	1.754	0.702	0.642	0.025	0.018	0.000	0.000

Table C.6
OLS regression results for full data sample.

	(1)	(2)
ln(ETA)	-0.104^{***}	-0.032^{***}
	(0.001)	(0.001)
$\ln(1+\text{PT})$	0.014^{***}	-0.124^{***}
	(0.002)	(0.002)
Controls		x
N	5177358	5177358
R^2	0.022	0.079

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. Controls include dummy variables for local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.7
2SLS regressions for each day of the week.

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
ln(ETA)	-0.0289^{***}	-0.0285^{***}	-0.0285^{***}	-0.0251^{***}	-0.0245^{***}	-0.0249^{***}	-0.0243^{***}
	(0.0031)	(0.0031)	(0.0031)	(0.0030)	(0.0028)	(0.0026)	(0.0028)
$\ln(1+\mathrm{PT})$	-0.3776^{***}	-0.4208^{***}	-0.4087^{***}	-0.3557^{***}	-0.2948^{***}	-0.3763^{***}	-0.3837^{***}
	(0.0277)	(0.0305)	(0.0300)	(0.0214)	(0.0145)	(0.0193)	(0.0180)
ETA Elasticity	-0.0480^{***}	-0.0458^{***}	-0.0456^{***}	-0.0413^{***}	-0.0393^{***}	-0.0395^{***}	-0.0405^{***}
	(0.0052)	(0.0050)	(0.0050)	(0.0049)	(0.0045)	(0.0041)	(0.0046)
PT Elasticity	-0.6274^{***}	-0.6763^{***}	-0.6530^{***}	-0.5846^{***}	-0.4729^{***}	-0.5956^{***}	-0.6400^{***}
	(0.0460)	(0.0490)	(0.0479)	(0.0352)	(0.0233)	(0.0306)	(0.0300)
VOT	20.08^{***}	17.95^{***}	19.34^{***}	19.44^{***}	21.26^{***}	18.74^{***}	17.12^{***}
	(2.47)	(2.22)	(2.38)	(2.41)	(2.45)	(2.04)	(2.01)
Control Avg. ETA	3.13	3.03	2.93	3.07	3.38	2.86	3.12
Control Avg. Price	13.70	13.38	13.55	14.08	14.41	13.50	14.07
Control Req. Rate	0.606	0.627	0.630	0.610	0.625	0.634	0.604
Controls	x	x	x	x	x	x	x
N	621203	632247	623344	722306	913007	902130	763121
R^2	0.082	0.077	0.077	0.071	0.065	0.077	0.086

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\mathrm{ETA})$ and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.8
2SLS regressions by time of day.

	Weekdays				
	6–10 AM	10 AM–4 PM	4–7 PM	7–11 PM	11 PM–6 AM
ln(ETA)	-0.0448^{***}	-0.0271^{***}	-0.0279^{***}	-0.0154^{***}	-0.0251^{***}
	(0.0038)	(0.0033)	(0.0032)	(0.0029)	(0.0035)
$\ln(1+\mathrm{PT})$	-0.3899^{***}	-0.4095^{***}	-0.3186^{***}	-0.3398^{***}	-0.3333^{***}
	(0.0225)	(0.0353)	(0.0208)	(0.0255)	(0.0193)
ETA Elasticity	-0.0682^{***}	-0.0466^{***}	-0.0457^{***}	-0.0244^{***}	-0.0405^{***}
	(0.0057)	(0.0056)	(0.0052)	(0.0045)	(0.0057)
PT Elasticity	-0.5936^{***}	-0.7037^{***}	-0.5224^{***}	-0.5398^{***}	-0.5386^{***}
	(0.0342)	(0.0607)	(0.0340)	(0.0405)	(0.0313)
VOT	26.71^{***}	18.76^{***}	22.12^{***}	13.62^{***}	19.03^{***}
	(2.50)	(2.64)	(2.70)	(2.59)	(2.66)
Control Avg. ETA	3.54	2.95	3.16	2.60	3.81
Control Avg. Price	13.72	13.93	13.33	13.04	16.07
Control Req. Rate	0.662	0.587	0.614	0.631	0.619
Controls	x	x	x	x	x
N	545313	859760	659754	888423	558857
R^2	0.085	0.070	0.084	0.086	0.062

	Weekends				
	6–10 AM	10 AM–4 PM	4–7 PM	7–11 PM	11 PM–6 AM
ln(ETA)	-0.0307^{***}	-0.0303^{***}	-0.0285^{***}	-0.0211^{***}	-0.0159^{***}
	(0.0061)	(0.0038)	(0.0040)	(0.0038)	(0.0034)
$\ln(1+\mathrm{PT})$	-0.3791^{***}	-0.4138^{***}	-0.3664^{***}	-0.4217^{***}	-0.3412^{***}
	(0.0448)	(0.0265)	(0.0234)	(0.0389)	(0.0194)
ETA Elasticity	-0.0556^{***}	-0.0532^{***}	-0.0478^{***}	-0.0334^{***}	-0.0233^{***}
	(0.0110)	(0.0066)	(0.0067)	(0.0061)	(0.0050)
PT Elasticity	-0.6865^{***}	-0.7272^{***}	-0.6156^{***}	-0.6692^{***}	-0.4998^{***}
	(0.0811)	(0.0465)	(0.0393)	(0.0618)	(0.0285)
VOT	17.40^{***}	18.92^{***}	20.16^{***}	15.04^{***}	14.71^{***}
	(3.80)	(2.49)	(2.92)	(2.87)	(3.12)
Control Avg. ETA	4.13	3.09	3.14	2.59	2.76
Control Avg. Price	14.79	13.32	13.59	13.01	14.54
Control Req. Rate	0.557	0.574	0.600	0.633	0.683
Controls	x	x	x	x	x
N	135429	412620	280795	386487	449920
R^2	0.078	0.082	0.092	0.091	0.062

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.9
2SLS regressions by precipitation type.

	No Precipitation	Rain	Snow
$\ln(\text{ETA})$	-0.0256^{***}	-0.0318^{***}	-0.0516^{***}
	(0.0019)	(0.0030)	(0.0144)
$\ln(1+\text{PT})$	-0.3676^{***}	-0.3483^{***}	-0.3551^{***}
	(0.0145)	(0.0163)	(0.0596)
ETA Elasticity	-0.0416^{***}	-0.0510^{***}	-0.1033^{***}
	(0.0031)	(0.0048)	(0.0288)
PT Elasticity	-0.5974^{***}	-0.5598^{***}	-0.7113^{***}
	(0.0236)	(0.0263)	(0.1195)
VOT	19.04^{***}	23.10^{***}	26.56^{***}
	(1.50)	(2.26)	(7.83)
Control Avg. ETA	3.04	3.29	5.31
Control Avg. Price	13.84	13.88	16.18
Control Req. Rate	0.618	0.626	0.500
Controls	x	x	x
N	4174595	749874	29693
R^2	0.074	0.075	0.064

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\text{ETA})$ and $\ln(1+\text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.10
2SLS regressions for business and non-business users.

	Non-Business	Business
ln(ETA)	-0.0265^{***}	-0.0233^{**}
	(0.0018)	(0.0098)
$\ln(1+\text{PT})$	-0.3709^{***}	-0.2834^{***}
	(0.0135)	(0.0522)
ETA Elasticity	-0.0433^{***}	-0.0327^{**}
	(0.0029)	(0.0138)
PT Elasticity	-0.6053^{***}	-0.3985^{***}
	(0.0221)	(0.0734)
VOT	19.22^{***}	24.42^{**}
	(1.39)	(10.55)
Control Avg. ETA	3.11	2.30
Control Avg. Price	13.96	11.38
Control Req. Rate	0.616	0.714
Controls	x	x
N	4960178	217180
R^2	0.071	0.096

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1+\text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.11
2SLS regressions on each region in the experiment.

	Austin	San Francisco	Los Angeles	Miami	Boston	Seattle	San Diego	Atlanta	New York City
ln(ETA)	-0.0181^{**}	-0.0187^{***}	-0.0250^{***}	-0.0300^{***}	-0.0434^{***}	-0.0331^{***}	-0.0271^{***}	-0.0309^{***}	-0.0285^{***}
	(0.0078)	(0.0034)	(0.0039)	(0.0060)	(0.0062)	(0.0068)	(0.0067)	(0.0092)	(0.0045)
$\ln(1+\text{PT})$	-0.4593^{***}	-0.2932^{***}	-0.3241^{***}	-0.3431^{***}	-0.3922^{***}	-0.3200^{***}	-0.4419^{***}	-0.4565^{***}	-0.4896^{***}
	(0.1040)	(0.0183)	(0.0441)	(0.0596)	(0.0421)	(0.0477)	(0.0928)	(0.1418)	(0.0238)
ETA Elasticity	-0.0280^{**}	-0.0263^{***}	-0.0403^{***}	-0.0519^{***}	-0.0742^{***}	-0.0534^{***}	-0.0444^{***}	-0.0558^{***}	-0.0557^{***}
	(0.0120)	(0.0047)	(0.0063)	(0.0103)	(0.0107)	(0.0109)	(0.0109)	(0.0166)	(0.0088)
PT Elasticity	-0.7088^{***}	-0.4127^{***}	-0.5217^{***}	-0.5948^{***}	-0.6704^{***}	-0.5163^{***}	-0.7232^{***}	-0.8246^{***}	-0.9564^{***}
	(0.1605)	(0.0257)	(0.0711)	(0.1033)	(0.0721)	(0.0770)	(0.1520)	(0.2561)	(0.0465)
VOT	11.37^{**}	21.08^{***}	19.52^{***}	16.36^{***}	23.37^{***}	27.08^{***}	14.80^{***}	11.24^{**}	23.85^{***}
	(4.76)	(3.79)	(3.91)	(3.95)	(3.79)	(6.38)	(4.34)	(4.40)	(3.81)
Median Wage	18.17	24.90	19.02	16.30	25.37	23.65	19.80	17.94	22.13
Mean Wage	24.44	33.23	26.71	22.19	32.66	30.43	26.68	24.38	30.44
Control Avg. ETA	2.79	2.12	3.02	4.39	3.65	3.41	3.43	4.92	3.09
Control Avg. Price	13.40	11.65	12.71	13.73	12.86	14.85	13.81	13.62	21.08
Control Req. Rate	0.650	0.711	0.627	0.581	0.593	0.626	0.612	0.555	0.515
Controls	x	x	x	x	x	x	x	x	x
N	218666	1387300	1044273	501297	345974	266222	282393	222981	908252
R^2	0.060	0.077	0.068	0.041	0.054	0.073	0.073	0.050	0.029

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1+\text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.
Wage data are the "All Occupations" median and mean hourly wages for the metropolitan area most closely corresponding to each Lyft region, as reported in the May 2016 Occupational Employment Statistics (U.S. Bureau of Labor Statistics, 2016).

Table C.12
VOT in airport and non-airport sessions.

	Non-Airport	Airport
ln(ETA)	-0.0266^{***}	-0.0143
	(0.0018)	(0.0092)
$\ln(1+\mathrm{PT})$	-0.3673^{***}	-0.3348^{***}
	(0.0132)	(0.0924)
ETA Elasticity	-0.0431^{***}	-0.0236
	(0.0029)	(0.0152)
PT Elasticity	-0.5951^{***}	-0.5534^{***}
	(0.0214)	(0.1525)
VOT	18.84^{***}	27.55
	(1.35)	(18.02)
Control Avg. ETA	3.09	2.92
Control Avg. Price	13.37	31.40
Control Req. Rate	0.620	0.606
Controls	x	x
N	5048270	129088
R^2	0.074	0.080

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and ln(1+PT) instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.13
VOT in downtown and non-downtown sessions.

	Non-downtown	Downtown
$\ln(\mathrm{ETA})$	-0.0316^{***}	-0.0231^{***}
	(0.0028)	(0.0021)
$\ln(1+\mathrm{PT})$	-0.4031^{***}	-0.3520^{***}
	(0.0304)	(0.0128)
ETA Elasticity	-0.0598^{***}	-0.0344^{***}
	(0.0054)	(0.0031)
PT Elasticity	-0.7624^{***}	-0.5241^{***}
	(0.0575)	(0.0192)
VOT	17.88^{***}	21.64^{***}
	(1.97)	(1.99)
Control Avg. ETA	4.36	2.28
Control Avg. Price	16.57	12.54
Control Req. Rate	0.533	0.674
Controls	x	x
N	1986393	3190965
R^2	0.049	0.059

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\mathrm{ETA})$ and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.14
2SLS regressions by distance to nearest public transit stop.

	Under 50 Meters	50 to 200 Meters	200 to 800 Meters	Over 800 Meters
ln(ETA)	-0.0269^{***}	-0.0251^{***}	-0.0281^{***}	-0.0279^{***}
	(0.0028)	(0.0022)	(0.0032)	(0.0066)
$\ln(1+\mathrm{PT})$	-0.3381^{***}	-0.3357^{***}	-0.4460^{***}	-0.5633^{***}
	(0.0174)	(0.0154)	(0.0274)	(0.0780)
ETA Elasticity	-0.0393^{***}	-0.0395^{***}	-0.0505^{***}	-0.0575^{***}
	(0.0041)	(0.0034)	(0.0057)	(0.0137)
PT Elasticity	-0.4951^{***}	-0.5283^{***}	-0.8021^{***}	-1.1624^{***}
	(0.0256)	(0.0242)	(0.0493)	(0.1610)
VOT	23.81^{***}	20.81^{***}	16.90^{***}	10.88^{***}
	(2.56)	(1.91)	(2.03)	(2.82)
Control Avg. ETA	2.47	2.82	3.60	5.29
Control Avg. Price	12.34	13.07	16.08	19.39
Control Req. Rate	0.686	0.638	0.560	0.487
Controls	x	x	x	x
N	1054335	2554328	1269849	298846
R^2	0.062	0.071	0.058	0.059

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\mathrm{ETA})$ and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.15
2SLS by passenger lifetime rides prior to the start of the experiment.

	0	1–3	4–10	11–20	20–50	Over 50
ln(ETA)	-0.0172^{***}	-0.0345^{***}	-0.0322^{***}	-0.0288^{***}	-0.0333^{***}	-0.0221^{***}
	(0.0031)	(0.0053)	(0.0048)	(0.0056)	(0.0046)	(0.0037)
$\ln(1+\mathrm{PT})$	-0.2602^{***}	-0.4464^{***}	-0.4103^{***}	-0.4212^{***}	-0.4530^{***}	-0.3226^{***}
	(0.0280)	(0.0370)	(0.0362)	(0.0391)	(0.0319)	(0.0235)
ETA Elasticity	-0.0320^{***}	-0.0727^{***}	-0.0591^{***}	-0.0469^{***}	-0.0488^{***}	-0.0281^{***}
	(0.0057)	(0.0111)	(0.0089)	(0.0092)	(0.0067)	(0.0048)
PT Elasticity	-0.4844^{***}	-0.9417^{***}	-0.7527^{***}	-0.6863^{***}	-0.6648^{***}	-0.4116^{***}
	(0.0520)	(0.0780)	(0.0666)	(0.0637)	(0.0469)	(0.0299)
VOT	17.62^{***}	22.34^{***}	22.23^{***}	19.06^{***}	20.41^{***}	19.38^{***}
	(3.39)	(3.66)	(3.68)	(3.89)	(2.96)	(3.35)
Control Avg. ETA	3.59	3.50	3.28	3.07	2.81	2.31
Control Avg. Price	15.98	16.89	15.48	14.29	13.02	10.92
Control Req. Rate	0.539	0.479	0.551	0.617	0.687	0.787
Controls	x	x	x	x	x	x
N	1418383	561670	709230	545632	767179	1175264
R^2	0.029	0.034	0.036	0.033	0.027	0.025

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1 + \mathrm{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.16
2SLS by passenger rides in the 28 days prior to the start of the experiment.

	0	1–2	3–8	Over 8
ln(ETA)	-0.0260^{***}	-0.0297^{***}	-0.0281^{***}	-0.0222^{***}
	(0.0025)	(0.0045)	(0.0038)	(0.0039)
$\ln(1+\text{PT})$	-0.3231^{***}	-0.4043^{***}	-0.3975^{***}	-0.3869^{***}
	(0.0206)	(0.0295)	(0.0272)	(0.0263)
ETA Elasticity	-0.0490^{***}	-0.0526^{***}	-0.0435^{***}	-0.0286^{***}
	(0.0047)	(0.0079)	(0.0058)	(0.0050)
PT Elasticity	-0.6081^{***}	-0.7163^{***}	-0.6158^{***}	-0.4988^{***}
	(0.0388)	(0.0523)	(0.0422)	(0.0340)
VOT	22.11^{***}	21.25^{***}	19.75^{***}	15.22^{***}
	(2.34)	(3.35)	(2.83)	(2.71)
Control Avg. ETA	3.48	3.13	2.86	2.53
Control Avg. Price	15.94	15.10	13.34	11.18
Control Req. Rate	0.534	0.568	0.650	0.780
Controls	x	x	x	x
N	2133364	737501	1118059	1188434
R^2	0.037	0.055	0.051	0.035

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1+\text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.17
Results of Durbin-Wu-Hausman augmented regression test of exogeneity.

	Requested
$\ln(\mathrm{ETA})$	-0.0264^{***}
	(0.0018)
$\ln(1+\mathrm{PT})$	-0.3665^{***}
	(0.0132)
$\ln(\mathrm{ETA})$ Residuals	-0.0094^{***}
	(0.0018)
$\ln(1+\mathrm{PT})$ Residuals	0.2542^{***}
	(0.0133)
Controls	x
N	5177358
F	213^{***}

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. The Durbin-Wu-Hausman augmented regression test of exogeneity regresses the independent variable on the suspected endogenous variables and the residuals from the first stage regressions of the suspected endogenous variables on the exogenous instruments.
t-tests suggest that the coefficients on the residuals in this regression differ significantly from 0, and the F statistic for the joint hypothesis that both coefficients are 0 is 213 ($p < 0.0001$). This result supports the conclusion that $\ln(\mathrm{ETA})$ and $\ln(1+\mathrm{PT})$ are endogeneous in equation (13), so that their coefficients cannot be consistently estimated by OLS.
Controls include dummy variables for local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.18
Test of overidentifying restrictions.

	2SLS Residuals
High ETA High Price	-0.0019
	(0.0018)
High ETA Low Price	0.0001
	(0.0013)
High ETA Normal Price	-0.0009
	(0.0012)
Normal ETA High Price	-0.0014
	(0.0013)
Normal ETA Low Price	-0.0018
	(0.0012)
Controls	x
Hansen J	3.769

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. Independent variables are indicators for each treatment group.
The Hansen J statistic is the GMM criterion function evaluated at $\hat{\beta}_{2SLS}$, using a cluster-robust optimal weighting matrix. The statistic tests the null that the excluded instruments are orthogonal to the second-stage error terms, as required for 2SLS consistency. If follows an asymptotic $\chi^2(\mathrm{df} = 5 - 2)$ distributions in our setup, and is robust to clustering of the second stage errors.

Table C.19
Values of time by week in experiment 1.

	1	2	3	4	5	6	7	8
ln(ETA)	-0.0212^{***}	-0.0278^{***}	-0.0283^{***}	-0.0243^{***}	-0.0195^{***}	-0.0245^{***}	-0.0269^{***}	-0.0331^{***}
	(0.0033)	(0.0029)	(0.0031)	(0.0037)	(0.0034)	(0.0034)	(0.0032)	(0.0033)
$\ln(1+\mathrm{PT})$	-0.3360^{***}	-0.3534^{***}	-0.3433^{***}	-0.4000^{***}	-0.2738^{***}	-0.5261^{***}	-0.3935^{***}	-0.4213^{***}
	(0.0207)	(0.0151)	(0.0167)	(0.0445)	(0.0221)	(0.0568)	(0.0272)	(0.0272)
ETA Elasticity	-0.0337^{***}	-0.0447^{***}	-0.0466^{***}	-0.0423^{***}	-0.0329^{***}	-0.0386^{***}	-0.0425^{***}	-0.0522^{***}
	(0.0052)	(0.0047)	(0.0050)	(0.0065)	(0.0058)	(0.0054)	(0.0050)	(0.0052)
PT Elasticity	-0.5348^{***}	-0.5683^{***}	-0.5657^{***}	-0.6953^{***}	-0.4634^{***}	-0.8272^{***}	-0.6216^{***}	-0.6629^{***}
	(0.0329)	(0.0243)	(0.0275)	(0.0773)	(0.0374)	(0.0893)	(0.0430)	(0.0429)
VOT	18.73^{***}	22.93^{***}	22.83^{***}	16.92^{***}	18.07^{***}	13.37^{***}	17.40^{***}	18.07^{***}
	(2.96)	(2.46)	(2.56)	(3.04)	(3.20)	(2.20)	(2.21)	(1.97)
Control Avg. ETA	2.81	2.88	3.17	3.26	3.60	2.73	2.99	3.24
Control Avg. Price	13.93	14.00	14.64	15.12	15.25	13.04	12.69	12.41
Control Req. Rate	0.630	0.625	0.611	0.580	0.591	0.640	0.637	0.639
Controls	x	x	x	x	x	x	x	x
N	667477	737270	709946	496447	677050	596913	693419	598836
R^2	0.083	0.080	0.081	0.076	0.075	0.084	0.070	0.082

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\mathrm{ETA})$ and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.20
Values of time by session in experiment one.

	1	2	3	4	5	6	7	8
ln(ETA)	-0.0250^{***}	-0.0224^{***}	-0.0255^{***}	-0.0254^{***}	-0.0285^{***}	-0.0234^{***}	-0.0255^{***}	-0.0290^{***}
	(0.0026)	(0.0030)	(0.0033)	(0.0036)	(0.0039)	(0.0041)	(0.0044)	(0.0047)
$\ln(1+\text{PT})$	-0.3366^{***}	-0.2817^{***}	-0.3095^{***}	-0.3326^{***}	-0.3181^{***}	-0.3031^{***}	-0.2688^{***}	-0.3364^{***}
	(0.0177)	(0.0205)	(0.0228)	(0.0253)	(0.0276)	(0.0298)	(0.0319)	(0.0342)
ETA Elasticity	-0.0424^{***}	-0.0407^{***}	-0.0448^{***}	-0.0438^{***}	-0.0482^{***}	-0.0393^{***}	-0.0425^{***}	-0.0481^{***}
	(0.0045)	(0.0055)	(0.0057)	(0.0062)	(0.0065)	(0.0069)	(0.0074)	(0.0079)
PT Elasticity	-0.5717^{***}	-0.5121^{***}	-0.5446^{***}	-0.5740^{***}	-0.5385^{***}	-0.5088^{***}	-0.4485^{***}	-0.5584^{***}
	(0.0302)	(0.0373)	(0.0401)	(0.0437)	(0.0467)	(0.0501)	(0.0532)	(0.0567)
VOT	22.72^{***}	22.81^{***}	23.09^{***}	21.16^{***}	24.62^{***}	20.88^{***}	25.54^{***}	23.03^{***}
	(2.50)	(3.21)	(3.18)	(3.14)	(3.65)	(3.90)	(4.97)	(4.09)
Control Avg. ETA	3.36	3.24	3.20	3.16	3.13	3.11	3.09	3.07
Control Avg. Price	17.14	15.51	14.96	14.62	14.31	14.03	13.88	13.69
Control Req. Rate	0.592	0.553	0.572	0.582	0.594	0.599	0.603	0.605
Controls	x	x	x	x	x	x	x	x
N	720059	548960	441707	367616	312175	268928	233632	203882
R^2	0.070	0.066	0.064	0.063	0.067	0.069	0.073	0.076

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1+\text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.21
Values of time by session in experiment one (passengers with at least three prior rides).

	1	2	3	4	5	6	7	8
ln(ETA)	-0.0322^{***}	-0.0308^{***}	-0.0306^{***}	-0.0286^{***}	-0.0340^{***}	-0.0305^{***}	-0.0289^{***}	-0.0253^{***}
	(0.0038)	(0.0041)	(0.0043)	(0.0046)	(0.0049)	(0.0052)	(0.0054)	(0.0057)
$\ln(1+\text{PT})$	-0.3986^{***}	-0.3194^{***}	-0.3089^{***}	-0.3528^{***}	-0.3491^{***}	-0.3467^{***}	-0.3218^{***}	-0.3352^{***}
	(0.0228)	(0.0248)	(0.0266)	(0.0283)	(0.0302)	(0.0329)	(0.0341)	(0.0365)
ETA Elasticity	-0.0535^{***}	-0.0510^{***}	-0.0494^{***}	-0.0456^{***}	-0.0534^{***}	-0.0471^{***}	-0.0444^{***}	-0.0384^{***}
	(0.0064)	(0.0068)	(0.0070)	(0.0074)	(0.0076)	(0.0080)	(0.0083)	(0.0087)
PT Elasticity	-0.6626^{***}	-0.5280^{***}	-0.4993^{***}	-0.5618^{***}	-0.5475^{***}	-0.5356^{***}	-0.4937^{***}	-0.5090^{***}
	(0.0379)	(0.0411)	(0.0430)	(0.0451)	(0.0474)	(0.0508)	(0.0524)	(0.0554)
VOT	24.09^{***}	28.52^{***}	28.98^{***}	23.60^{***}	28.09^{***}	24.88^{***}	25.40^{***}	21.24^{***}
	(3.00)	(4.05)	(4.43)	(3.96)	(4.35)	(4.43)	(5.06)	(4.95)
Control Avg. ETA	3.00	2.92	2.91	2.90	2.87	2.87	2.83	2.81
Control Avg. Price	14.92	14.40	14.22	14.05	13.80	13.53	13.35	13.22
Control Req. Rate	0.606	0.609	0.624	0.632	0.642	0.653	0.657	0.661
Controls	x	x	x	x	x	x	x	x
N	337881	282903	243117	212205	187175	166323	148726	133426
R^2	0.080	0.069	0.067	0.065	0.068	0.066	0.070	0.070

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1+\text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.22
Checks of robustness to data subsample.

	No Holidays	Standard Only	Considered Shared
ln(ETA)	-0.0266^{***}	-0.0336^{***}	-0.0137^{***}
	(0.0018)	(0.0021)	(0.0029)
$\ln(1 + \text{PT})$	-0.3781^{***}	-0.4182^{***}	-0.2982^{***}
	(0.0140)	(0.0181)	(0.0169)
ETA Elasticity	-0.0428^{***}	-0.0548^{***}	-0.0217^{***}
	(0.0029)	(0.0034)	(0.0046)
PT Elasticity	-0.6079^{***}	-0.6830^{***}	-0.4727^{***}
	(0.0226)	(0.0296)	(0.0268)
VOT	19.34^{***}	21.44^{***}	12.97^{***}
	(1.41)	(1.49)	(2.73)
Control Avg. ETA	2.97	3.23	2.62
Control Avg. Price	13.58	14.36	12.34
Control Req. Rate	0.625	0.616	0.633
Controls	x	x	x
N	4750614	3232416	1804771
R^2	0.076	0.068	0.089

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1 + \text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.23
Checks of robustness to model specification.

	Completed	Avg ETA	Linear ETA	Squared ETA
ln(ETA)	-0.0580^{***}	-0.0272^{***}	—	—-
	(0.0018)	(0.0018)		
$\ln(1+\mathrm{PT})$	-0.3690^{***}	-0.3664^{***}	-0.3664^{***}	-0.3661^{***}
	(0.0131)	(0.0132)	(0.0131)	(0.0131)
ETA	—	—	-0.4770^{***}	—
			(0.0321)	
ETA^2	—	—	—	-2.915^{***}
				(0.196)
VOT	42.31^{***}	19.90^{***}	18.01^{***}	11.32^{***}
	(1.83)	(1.42)	(1.29)	(0.81)
Controls	x	x	x	x
N	5177358	5177358	5177358	5177358
R^2	0.092	0.072	0.072	0.070

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\mathrm{ETA})$ and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include dummy variables for user geohash5, local hour of week and week of year, decile of user lifetime rides, an indicator of whether the user is a business user, and an indicator of whether the session is at an airport
In column (1), the dependent variable is an indicator of whether a session had a completed ride, while for the other three the dependent variable is an indicator of whether a session had a ride request. In column (2), the session-average ETA is used in place of the last-in-session ETA. In columns (3) and (4), ETA enters the equation linearly and squared, respectively, rather in logarithm. For both of these columns, ETA is in units of hours.

Table C.24
Checks of robustness to estimation strategy.

	IV-Probit (MLE)
$\ln(\text{ETA})$	-0.061^{***}
	(0.006)
$\ln(1+\text{PT})$	-0.934^{***}
	(0.042)
AME of $\ln(\text{ETA})$	-0.023^{***}
	(0.002)
AME of $\ln(1+\text{PT})$	-0.351^{***}
	(0.015)
VOT	17.53^{***}
	(1.78)
Controls	
N	5177358
Log Pseudolikelihood	-7060668.5

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\text{ETA})$ and $\ln(1+\text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour.

Table C.25
2SLS results by region, using linear functional form (experiment 1).

	Austin	San Francisco	Los Angeles	Miami	Boston	Seattle	San Diego	Atlanta	New York City
ETA	-0.350^{**}	-0.469^{***}	-0.476^{***}	-0.390^{***}	-0.637^{***}	-0.545^{***}	-0.424^{***}	-0.348^{***}	-0.525^{***}
	(0.150)	(0.084)	(0.075)	(0.078)	(0.092)	(0.112)	(0.104)	(0.103)	(0.083)
$\ln(1+\text{PT})$	-0.460^{***}	-0.293^{***}	-0.324^{***}	-0.343^{***}	-0.392^{***}	-0.322^{***}	-0.442^{***}	-0.457^{***}	-0.489^{***}
	(0.104)	(0.018)	(0.044)	(0.060)	(0.042)	(0.048)	(0.093)	(0.142)	(0.024)
VOT	10.18^{**}	18.65^{***}	18.68^{***}	15.61^{***}	20.89^{***}	25.18^{***}	13.23^{***}	10.38^{**}	22.60^{***}
	(4.27)	(3.35)	(3.75)	(3.77)	(3.39)	(5.90)	(3.87)	(4.05)	(3.61)
Control Avg. ETA	2.79	2.12	3.02	4.39	3.65	3.41	3.43	4.92	3.09
Control Avg. Price	13.40	11.65	12.71	13.73	12.86	14.85	13.81	13.62	21.08
Control Req. Rate	0.650	0.711	0.627	0.581	0.593	0.626	0.612	0.555	0.515
Controls	x	x	x	x	x	x	x	x	x
N	218666	1387300	1044273	501297	345974	266222	282393	222981	908252
R^2	0.059	0.077	0.067	0.041	0.054	0.073	0.073	0.050	0.029

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\text{ETA})$ and $\ln(1+\text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.26
Summary statistics for sessions with and without destinations entered.

	Without Destination	With Destination
Mean ETA (minutes)	3.722 (0.004)	3.364 (0.004)
Mean PT (%)	8.143 (0.022)	11.472 (0.33)
Requested	0.510 (0.001)	0.782 (0.001)

Table C.27
2SLS results by whether session had a destination entered.

	Without Destination	With Destination
ln(ETA)	-0.0367^{***}	-0.0105^{***}
	(0.0024)	(0.0022)
$\ln(1 + \mathrm{PT})$	-0.4554^{***}	-0.3644^{***}
	(0.0201)	(0.0137)
$\ln(Price)$	—	-0.1043^{***}
		(0.0014)
ETA Elasticity	-0.0721^{***}	-0.0134^{***}
	(0.0046)	(0.0028)
PT Elasticity	-0.8938^{***}	-0.4661^{***}
	(0.0394)	(0.0175)
VOT	21.07^{***}	8.20^{***}
	(1.53)	(1.68)
Control Avg. ETA	3.20	2.90
Control Avg. Price	13.93	13.73
Control Req. Rate	0.514	0.639
Controls	x	x
N	3136534	2040824
R^2	0.112	0.055

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1 + \mathrm{PT})$ instrumented by experimental group indicators. Price is midpoint of estimated fare range. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.28
Effect of treatments on probability of entering destination.

	Whether Destination Entered
High ETA High Price	0.010^{***}
	(0.002)
High ETA Normal Price	-0.000
	(0.002)
High ETA Low Price	-0.005^{***}
	(0.002)
Normal ETA High Price	0.010^{***}
	(0.002)
Normal ETA Low Price	-0.005^{***}
	(0.002)
Controls	x
N	5177358
R^2	0.103
F	21^{***}

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. F test is of the null that all the coefficients on the instruments equal 0. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.29
First stage for $\ln(\text{ETA})$ by region (2015 experiment).

	Dependent variable: $\ln(\text{ETA})$								
	Austin	San Francisco	Los Angeles	Miami	Boston	Seattle	San Diego	Atlanta	New York City
High ETA High Price	0.502^{***}	0.484^{***}	0.492^{***}	0.494^{***}	0.504^{***}	0.515^{***}	0.495^{***}	0.434^{***}	0.432^{***}
	(0.005)	(0.003)	(0.003)	(0.003)	(0.005)	(0.006)	(0.005)	(0.004)	(0.002)
High ETA Normal Price	0.488^{***}	0.485^{***}	0.491^{***}	0.491^{***}	0.508^{***}	0.517^{***}	0.492^{***}	0.428^{***}	0.435^{***}
	(0.004)	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)	(0.003)	(0.002)
High ETA Low Price	0.496^{***}	0.486^{***}	0.491^{***}	0.492^{***}	0.507^{***}	0.520^{***}	0.495^{***}	0.429^{***}	0.432^{***}
	(0.004)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)	(0.002)
Normal ETA High Price	0.000	0.001	-0.000	0.001	0.006	0.006	0.006	-0.002	-0.003
	(0.004)	(0.002)	(0.002)	(0.003)	(0.004)	(0.005)	(0.004)	(0.004)	(0.002)
Normal ETA Low Price	0.000	0.002	-0.000	0.000	0.000	0.009^{*}	0.002	-0.002	-0.001
	(0.004)	(0.002)	(0.002)	(0.003)	(0.004)	(0.005)	(0.004)	(0.004)	(0.002)
Controls	x	x	x	x	x	x	x	x	x
N	218666	1387300	1044273	501297	345974	266222	282393	222981	908252
R^2	0.515	0.478	0.497	0.516	0.493	0.507	0.519	0.481	0.457
F	7131^{***}	28858^{***}	24306^{***}	16708^{***}	8919^{***}	6078^{***}	6957^{***}	8216^{***}	29963^{***}

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. F test is of the null that all the coefficients on the instruments equal 0. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.30
First stage for $\ln(1 + \mathrm{PT})$ by region (2015 experiment).

	Dependent variable: ln(1 + PT)								
	Austin	San Francisco	Los Angeles	Miami	Boston	Seattle	San Diego	Atlanta	New York City
High ETA High Price	0.003	0.062^{***}	0.042^{***}	0.030^{***}	0.061^{***}	0.055^{***}	0.018^{***}	0.015^{***}	0.072^{***}
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
High ETA Normal Price	-0.002^{*}	0.001	0.001^{**}	0.001	−0.001	0.003^{**}	−0.000	0.001	0.000
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
High ETA Low Price	-0.035^{***}	-0.057^{***}	-0.023^{***}	-0.039^{***}	-0.043^{***}	-0.049^{***}	-0.027^{***}	-0.024^{***}	-0.047^{***}
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Normal ETA High Price	0.009^{***}	0.065^{***}	0.041^{***}	0.027^{***}	0.060^{***}	0.055^{***}	0.020^{***}	0.013^{***}	0.072^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Normal ETA Low Price	-0.037^{***}	-0.057^{***}	-0.023^{***}	-0.039^{***}	-0.044^{***}	-0.048^{***}	-0.027^{***}	-0.023^{***}	-0.047^{***}
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Controls	x	x	x	x	x	x	x	x	x
N	218666	1387300	1044273	501297	345974	266222	282393	222981	908252
R^2	0.406	0.254	0.300	0.283	0.219	0.297	0.345	0.221	0.227
F	636^{***}	7822^{***}	3889^{***}	1803^{***}	2319^{***}	1569^{***}	822^{***}	785^{***}	6891^{***}

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. F test is of the null that all the coefficients on the instruments equal 0. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table C.31
Counts of geohash7s per region in experiment 2.

Region	Observed geohash7s	Total geohash7s
Atlanta	84046	130643
New York City	41951	62148
Boston	44413	633542
Chicago	80633	883088
Washington, D.C.	67422	364094
Los Angeles	122626	314864
Miami	82108	763292
New Jersey	103103	1269121
Philadelphia	43829	235351
San Francisco	60933	293467

Notes: A geohash7 is considered "observed" if it has at least one session in the experiment sample. Total geohash7 counts include geohash7s which are entirely covered by water.

Table C.32
ETAs and sample sizes in experiment two.

	Average ETA (minutes)	**Number of location-hour blocks**	**Sessions**
Control	3.433 (0.002)	3562926	4861532
Plus 60+ seconds	5.099 (0.003)	1771883	2410230
Plus 150+ seconds	6.332 (0.003)	1060521	1441839
Plus 240+ seconds	7.673 (0.004)	703933	955219

Notes: Standard errors clustered at the geohash7-hour level in parentheses.

Table C.33

First stage regression of $\ln(\mathrm{ETA})$ for experiment 2.

	(1)
Plus 60+	0.4827^{***}
	(0.0004)
Plus 150+	0.7522^{***}
	(0.0004)
Plus 240+	0.9699^{***}
	(0.0005)
Controls	x
N	9668820
R^2	0.608
F	1644577.2^{***}

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.34

Regression of $\ln(1 + \mathrm{PT})$ on experimental treatment indicators for experiment 2.

	(1)
Plus 60+	-0.0003
	(0.0003)
Plus 150+	0.0001
	(0.0004)
Plus 240+	0.0009
	(0.0006)
Controls	x
N	9668820
R^2	0.144
F	1.4

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.35
Time elasticities (2017 experiment).

	Full sample	Destination	No Destination	2017 users in 2015	2015 users in 2017
ln(ETA)	-0.0265^{***}	-0.0150^{***}	-0.0268^{***}	-0.0303^{***}	-0.0312^{***}
	(0.0004)	(0.0005)	(0.0007)	(0.0029)	(0.0014)
$\ln(1+\mathrm{PT})$	-0.2851^{***}	-0.2002^{***}	-0.1393^{***}	-0.3785^{***}	-0.3169^{***}
	(0.0011)	(0.0013)	(0.0013)	(0.0183)	(0.0026)
$\ln(Price)$	—	-0.1677^{***}	—	—	—
		(0.0003)			
ETA Elasticity	-0.0432^{***}	-0.0202^{***}	-0.1292^{***}	-0.0460^{***}	-0.0467^{***}
	(0.0007)	(0.0003)	(0.0034)	(0.0044)	(0.0021)
VOT	—	22.15^{***}	—	22.71^{***}	—
		(0.67)		(2.25)	
Control Avg. ETA	3.43	3.33	3.76	2.72	2.82
Control Avg. Price	11.96	13.85	19.53	12.90	11.37
Control Req. Rate	0.620	0.742	0.228	0.661	0.678
N	9668820	7395984	2272839	2098259	711857
R^2	0.075	0.110	0.228	0.082	0.106

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses, with clustering for the 2017 data at the hour-geohash7-level and clustering for the 2015 data at the user-level. ln(ETA) instrumented by experimental group indicators; $\ln(1+\mathrm{PT})$ is instrumented using experimental group indicators. Including controls for region, geohash5, local hour of week, local week of year, business user, and decile of user lifetime rides. Control for airports included for 2015 data. $Price$ is the upfront price shown to passengers who enter a destination in the 2017 experiment. VOT for sessions with a destination entered is computed using the coefficient on $\ln(Price)$ as the price semi-elasticity and the average upfront price shown as the average price. VOT for 2015 users in 2015 is computed using the 2015 PT multiplier semi-elasticity.
2017 prices deflated to December 2015 USD using the Consumer Price Index for All Urban Consumers (U.S. Bureau of Labor Statistics, 2020).
"2017 users in 2015" refers to the sample of sessions in the 2015 experiment from users who appear in both experiments. "2015 users in 2017" refers to the sample of sessions in the 2017 experiment from users who appear in both experiments.

Table C.36
Time elasticities by day of week in experiment two.

	Mon	Tues	Wed	Thurs	Fri	Sat	Sun
ln(ETA)	-0.0261^{***}	-0.0291^{***}	-0.0278^{***}	-0.0295^{***}	-0.0264^{***}	-0.0238^{***}	-0.0237^{***}
	(0.0013)	(0.0012)	(0.0012)	(0.0012)	(0.0011)	(0.0010)	(0.0012)
$\ln(1+\mathrm{PT})$	-0.2978^{***}	-0.2883^{***}	-0.3087^{***}	-0.3151^{***}	-0.3065^{***}	-0.2633^{***}	-0.2664^{***}
	(0.0022)	(0.0026)	(0.0029)	(0.0022)	(0.0018)	(0.0031)	(0.0020)
ETA Elasticity	-0.0442^{***}	-0.0474^{***}	-0.0440^{***}	-0.0470^{***}	-0.0423^{***}	-0.0391^{***}	-0.0402^{***}
	(0.0021)	(0.0020)	(0.0019)	(0.0019)	(0.0017)	(0.0017)	(0.0020)
Control Avg. ETA	3.65	3.45	3.37	3.38	3.43	3.31	3.49
Control Req. Rate	0.599	0.623	0.640	0.635	0.632	0.616	0.597
Controls	x	x	x	x	x	x	x
N	1246509	1270040	1242952	1228361	1555551	1734742	1390665
R^2	0.086	0.080	0.074	0.077	0.079	0.070	0.082

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.37
Time elasticities by time of day (weekdays) in experiment two.

	6–10 AM	10 AM–4 PM	4 PM–7 PM	7 PM–11 PM	11 PM–6 AM
ln(ETA)	-0.0391^{***}	-0.0248^{***}	-0.0320^{***}	-0.0220^{***}	-0.0226^{***}
	(0.0012)	(0.0010)	(0.0013)	(0.0010)	(0.0015)
ln(1 + PT)	-0.3450^{***}	-0.2972^{***}	-0.3106^{***}	-0.2832^{***}	-0.2503^{***}
	(0.0018)	(0.0022)	(0.0022)	(0.0024)	(0.0029)
ETA Elasticity	-0.0607^{***}	-0.0411^{***}	-0.0521^{***}	-0.0353^{***}	-0.0368^{***}
	(0.0019)	(0.0017)	(0.0021)	(0.0017)	(0.0025)
Control Avg. ETA	3.61	3.51	3.66	2.96	3.69
Control Req. Rate	0.655	0.610	0.622	0.628	0.620
Controls	x	x	x	x	x
N	1215158	1831566	1228901	1472841	794947
R^2	0.092	0.068	0.079	0.091	0.080

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.38
Time elasticities by time of day (weekends) in experiment two.

	6–10 AM	10 AM–4 PM	4 PM–7 PM	7 PM–11 PM	11 PM–6 AM
ln(ETA)	-0.0342^{***}	-0.0278^{***}	-0.0268^{***}	-0.0198^{***}	-0.0169^{***}
	(0.0026)	(0.0015)	(0.0019)	(0.0016)	(0.0016)
ln(1 + PT)	-0.3083^{***}	-0.3124^{***}	-0.2974^{***}	-0.2559^{***}	-0.2296^{***}
	(0.0038)	(0.0025)	(0.0034)	(0.0035)	(0.0029)
ETA Elasticity	-0.0599^{***}	-0.0484^{***}	-0.0445^{***}	-0.0327^{***}	-0.0262^{***}
	(0.0045)	(0.0026)	(0.0032)	(0.0026)	(0.0024)
Control Avg. ETA	3.93	3.57	3.42	3.31	2.97
Control Req. Rate	0.580	0.582	0.610	0.611	0.650
Controls	x	x	x	x	x
N	320032	901858	507669	691344	704504
R^2	0.079	0.073	0.079	0.090	0.073

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.39
Time elasticities by precipitation type in experiment two.

	No Precipitation	Rain
ln(ETA)	-0.0261^{***}	-0.0277^{***}
	(0.0005)	(0.0010)
$\ln(1+\text{PT})$	-0.2783^{***}	-0.3072^{***}
	(0.0015)	(0.0015)
ETA Elasticity	-0.0421^{***}	-0.0473^{***}
	(0.0008)	(0.0017)
Control Avg. ETA	3.38	3.61
Control Req. Rate	0.628	0.595
Controls	x	x
N	7586012	2061855
R^2	0.074	0.071

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.40
Time elasticities by whether business user in experiment two.

	Not Business User	Business User
ln(ETA)	-0.0259^{***}	-0.0334^{***}
	(0.0005)	(0.0015)
$\ln(1+\text{PT})$	-0.2853^{***}	-0.2850^{***}
	(0.0012)	(0.0028)
ETA Elasticity	-0.0426^{***}	-0.0491^{***}
	(0.0008)	(0.0022)
Control Avg. ETA	3.47	2.88
Control Req. Rate	0.615	0.689
Controls	x	x
N	9002541	666279
R^2	0.074	0.089

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.41
Time elasticities by region in experiment two.

	San Francisco	New York City	Chicago	D.C.	Miami	New Jersey	Boston	Philadelphia	Atlanta	Los Angeles
ln(ETA)	-0.0221^{***}	-0.0445^{***}	-0.0251^{***}	-0.0328^{***}	-0.0136^{***}	-0.0279^{***}	-0.0326^{***}	-0.0234^{***}	-0.0215^{***}	-0.0207^{***}
	(0.0010)	(0.0012)	(0.0012)	(0.0015)	(0.0015)	(0.0020)	(0.0016)	(0.0020)	(0.0025)	(0.0011)
$\ln(1+\text{PT})$	-0.2496^{***}	-0.3741^{***}	-0.2844^{***}	-0.2923^{***}	-0.1366^{***}	-0.2978^{***}	-0.3523^{***}	-0.2579^{***}	-0.1908^{***}	-0.2379^{***}
	(0.0024)	(0.0022)	(0.0024)	(0.0021)	(0.0057)	(0.0028)	(0.0021)	(0.0034)	(0.0045)	(0.0039)
ETA Elasticity	-0.0307^{***}	-0.0856^{***}	-0.0394^{***}	-0.0543^{***}	-0.0211^{***}	-0.0530^{***}	-0.0552^{***}	-0.0397^{***}	-0.0367^{***}	-0.0330^{***}
	(0.0014)	(0.0024)	(0.0018)	(0.0025)	(0.0023)	(0.0038)	(0.0027)	(0.0034)	(0.0043)	(0.0017)
Control Avg. ETA	2.70	3.24	2.96	3.54	3.23	5.11	3.47	3.68	4.99	3.34
Control Req. Rate	0.727	0.532	0.645	0.613	0.647	0.534	0.599	0.597	0.591	0.633
Controls	x	x	x	x	x	x	x	x	x	x
N	1346530	1289094	1117826	908843	812250	651194	690288	482477	496835	1873483
R^2	0.075	0.067	0.070	0.059	0.051	0.069	0.084	0.065	0.060	0.060

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.42
Time elasticities by whether session downtown in experiment two.

	Non-Downtown	Downtown
ln(ETA)	-0.0271^{***}	-0.0258^{***}
	(0.0008)	(0.0005)
$\ln(1+\text{PT})$	-0.2707^{***}	-0.2940^{***}
	(0.0028)	(0.0010)
ETA Elasticity	-0.0477^{***}	-0.0392^{***}
	(0.0014)	(0.0008)
Control Avg. ETA	4.35	2.38
Control Req. Rate	0.576	0.668
Controls	x	x
N	4319737	4969830
R^2	0.062	0.070

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.43
Time elasticities by distance to public transit in experiment two.

	Under 50 Meters	50 to 200 Meters	200 to 800 Meters	Over 800 Meters
ln(ETA)	-0.0274^{***}	-0.0241^{***}	-0.0281^{***}	-0.0416^{***}
	(0.0008)	(0.0006)	(0.0011)	(0.0026)
$\ln(1+\mathrm{PT})$	-0.2863^{***}	-0.2883^{***}	-0.2853^{***}	-0.2505^{***}
	(0.0016)	(0.0011)	(0.0036)	(0.0045)
ETA Elasticity	-0.0417^{***}	-0.0383^{***}	-0.0494^{***}	-0.0841^{***}
	(0.0013)	(0.0009)	(0.0019)	(0.0052)
Control Avg. ETA	2.65	3.07	4.11	6.63
Control Req. Rate	0.666	0.636	0.576	0.503
Controls	x	x	x	x
N	1984715	4837325	2292367	554413
R^2	0.069	0.071	0.073	0.080

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.44
Summary statistics for sessions with and without destination entered in experiment 2. Standard errors in parentheses.

	Without destination	**With destination**
Average ETA (minutes)	5.114 (0.003)	4.572 (0.001)
Average PT (%)	13.283 (0.037)	15.143 (0.041)
Request Rate (%)	20.729 (0.030)	73.707 (0.017)

Table C.45
Time elasticities by whether destination entered in experiment 2

	Without Destination	With Destination
ln(ETA)	-0.0268^{***}	-0.0148^{***}
	(0.0007)	(0.0005)
$\ln(1+\mathrm{PT})$	-0.1393^{***}	-0.1990^{***}
	(0.0013)	(0.0014)
$\ln(Price)$	—	-0.1668^{***}
		(0.0005)
ETA Elasticity	-0.1292^{***}	-0.0202^{***}
	(0.0034)	(0.0007)
VOT	—	22.66^{***}
		(0.77)
Control Avg. ETA	3.76	3.34
Control Req. Rate	0.216	0.742
Control Avg. Price	—	14.22
Controls	x	x
N	2272839	7422265
R^2	0.228	0.073

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides. Value of time calculated using the semi-elasticity of price.

Table C.46
Convexity by whether session was downtown.

	Non-downtown	Downtown
Control — Plus 60	-0.034^{***}	-0.022^{***}
	(0.002)	(0.001)
Control — Plus 150	-0.046^{***}	-0.031^{***}
	(0.001)	(0.001)
Control — Plus 240	-0.055^{***}	-0.051^{***}
	(0.001)	(0.001)

Notes: Clustered standard errors in parentheses. ln(ETA) instrumented by experimental group indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.47
Convexity by day of week.

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Control — Plus 60	-0.020^{***}	-0.030^{***}	-0.026^{***}	-0.033^{***}	-0.028^{***}	-0.025^{***}	-0.027^{***}
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Control — Plus 150	-0.037^{***}	-0.039^{***}	-0.037^{***}	-0.040^{***}	-0.038^{***}	-0.034^{***}	-0.035^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Control — Plus 240	-0.058^{***}	-0.060^{***}	-0.057^{***}	-0.057^{***}	-0.050^{***}	-0.045^{***}	-0.050^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)

Notes: Clustered standard errors in parentheses. $\ln(\text{ETA})$ instrumented by experimental group indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.48
Convexity by time of day (weekdays).

	6–10 AM	10 AM–4 PM	4 PM–7 PM	7 PM–11 PM	11 PM–6 AM
Control — Plus 60	-0.039^{***}	-0.025^{***}	-0.030^{***}	-0.022^{***}	-0.021^{***}
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Control — Plus 150	-0.050^{***}	-0.038^{***}	-0.044^{***}	-0.026^{***}	-0.034^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Control — Plus 240	-0.076^{***}	-0.052^{***}	-0.066^{***}	-0.044^{***}	-0.042^{***}
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)

Notes: Clustered standard errors in parentheses. $\ln(\text{ETA})$ instrumented by experimental group indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.49
Convexity by time of day (weekends).

	6–10 AM	10 AM–4 PM	4 PM–7 PM	7 PM–11 PM	11 PM–6 AM
Control — Plus 60	-0.035^{***}	-0.028^{***}	-0.032^{***}	-0.024^{***}	-0.019^{***}
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Control — Plus 150	-0.057^{***}	-0.042^{***}	-0.040^{***}	-0.024^{***}	-0.024^{***}
	(0.004)	(0.002)	(0.003)	(0.003)	(0.003)
Control — Plus 240	-0.071^{***}	-0.058^{***}	-0.054^{***}	-0.041^{***}	-0.029^{***}
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)

Notes: Clustered standard errors in parentheses. $\ln(\text{ETA})$ instrumented by experimental group indicators. Controls include local week of year, local hour of week, user geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.50
VOT estimates at different ETA treatment levels.

	All Six Regions	San Francisco	Los Angeles	Miami	Boston	Atlanta	New York City
ETA Elasticity (Plus 60+)	-0.028^{***}	-0.015^{***}	-0.027^{***}	-0.015^{***}	-0.038^{***}	-0.025^{***}	-0.054^{***}
	(0.002)	(0.003)	(0.003)	(0.004)	(0.005)	(0.007)	(0.004)
ETA Elasticity (Plus 150+)	-0.036^{***}	-0.025^{***}	-0.026^{***}	-0.019^{***}	-0.046^{***}	-0.032^{***}	-0.080^{***}
	(0.001)	(0.002)	(0.002)	(0.003)	(0.004)	(0.006)	(0.003)
ETA Elasticity (Plus 240+)	-0.050^{***}	-0.039^{***}	-0.040^{***}	-0.024^{***}	-0.067^{***}	-0.044^{***}	-0.099^{***}
	(0.001)	(0.002)	(0.002)	(0.003)	(0.004)	(0.006)	(0.003)
PT Elasticity (2015)	-0.609^{***}	-0.469^{***}	-0.482^{***}	-0.482^{***}	-0.531^{***}	-0.901^{***}	-0.968^{***}
	(0.035)	(0.046)	(0.105)	(0.127)	(0.112)	(0.313)	(0.062)
VOT (Plus 60+)	11.19^{***}	8.98^{***}	11.73^{***}	6.44^{***}	16.16^{***}	4.10^{**}	21.28^{***}
	(0.90)	(1.73)	(2.85)	(2.46)	(3.98)	(1.88)	(2.14)
VOT (Plus 150+)	14.45^{***}	14.84^{***}	11.48^{***}	7.97^{***}	19.45^{***}	5.35^{**}	31.48^{***}
	(0.96)	(1.83)	(2.68)	(2.50)	(4.41)	(2.12)	(2.41)
VOT (Plus 240+)	20.04^{***}	23.15^{***}	17.36^{***}	10.32^{***}	27.99^{***}	7.28^{***}	39.09^{***}
	(1.24)	(2.49)	(3.89)	(2.98)	(6.09)	(2.69)	(2.78)
Base Avg. ETA (2017)	3.31	2.70	3.34	3.23	3.47	4.99	3.24
Control Avg. Price (2015)	12.50	11.45	10.90	9.65	12.23	11.53	19.90
Controls	x	x	x	x	x	x	x

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\text{ETA})$ and $\ln(1 + \text{PT})$ instrumented by experimental treatment indicators. Controls include local hour of week and week of year, session geohash5, business user, and decile of pre-experiment lifetime rides.

Table C.51

Summary of cost-benefit analyses and VoT measures used.

					Original VoT	Adjusted VoT	
Year	Location	Description of Project	Source of VoT	Measure	VoT used: 50% of avg/med wage	Simple rule: 75% of avg/med wage	Region-specific point estimate
2014	San Francisco, CA	SF MUNI public transit system	1/2 of average/median wage (following US DOT 2014 report)	VoT personal	$17.90	$26.85	$20.82
				VoT business/trucks	$35.50		
				% benefits from VoT	13%	17%	14%
				Total benefits	$1,897,700,000	$2,004,690,602	$1,932,448,621
				Total costs	$651,791,000	$651,791,000	$651,791,000
				Benefit-cost ratio	2.91	3.08	2.96
2002	Seattle, WA	Monorail “green line” in downtown Seattle	1/2 the average regional wage rate (Washington Employment Security Department	VoT personal	$10.10	$15.15	$21.31
				VoT business/trucks			
				% benefits from VoT	57%	66%	74%
				Total benefits	$2,067,264,491	$2,549,752,571	$3,140,301,358
				Total costs	$1,650,000,000	$1,650,000,000	$1,650,000,000
				Benefit-cost ratio	1.25	1.55	1.90
1998	San Francisco, CA	Electronic toll collector system in Bay Area	1/2 of average/median wage (following US DOT 1996 report)	VoT personal	$12.75	$19.13	$14.33
				VoT business/trucks	$33.41		
				% benefits from VoT	92%	92%	92%
				Total benefits	$13,663,608	$19,877,980	$15,212,255
				Total costs	$2,884,584	$2,884,584	$2,884,584
				Benefit-cost ratio	4.74	6.89	5.27
2000	Los Angeles, CA	Improved bus service, pedestrian/bike facilities, roads	1/2 statewide average hourly wage for personal vehicles	VoT personal	$13.25	$19.88	$14.47
				VoT business/trucks	$31.05		
				% benefits from VoT	63%	72%	65%
				Total benefits	$599,000,000	$789,000,000	$633,851,466
				Total costs	$255,000,000	$255,000,000	$255,000,000
				Benefit-cost ratio	2.35	3.09	2.49
2000	Los Angeles, CA	Freeway tunnel construction, single bore tunnel with tolls	1/2 statewide average hourly wage for personal vehicles	VoT personal	$13.25	$19.88	$14.47
				VoT business/trucks	$31.05		
				% benefits from VoT	75%	82%	77%
				Total benefits	$3,503,000,000	$4,815,500,000	$3,743,750,255
				Total costs	$1,979,000,000	$1,979,000,000	$1,979,000,000
				Benefit-cost ratio	1.77	2.43	1.89
2017	Atlanta, GA	Pedestrian/cyclist roadway construction	1/2 of average/median wage (following US DOT 2018 grant guidance report)	VoT personal	$14.80	$22.20	$18.55
				VoT business/trucks			
				% benefits from VoT	8%	11%	10%
				Total benefits	$143,040,000	$148,740,000	$145,931,073
				Total costs	$86,500,000	$86,500,000	$86,500,000
				Benefit-cost ratio	1.65	1.72	1.69

*All values of time (VoT) and measures of cost and benefit are given in units of USD in year of project (see first column). When both the business VoT and proportion of business trips are given in the original cost-benefit report, we leave benefit estimation for business travel unchanged and only impute our estimate for the personal value of time. Otherwise, (when not given a proportion for business travel trips) total benefit from business VoT is assumed to increase/decrease proportionally with the personal VoT (i.e. If we adjust the personal travel VoT by 10%, we also adjust the business travel VoT by 10%). This likely does not change results by much, since in analyses that report business travel trip counts, the proportion of travel that is business-related is generally a small proportion of total travel (between 5-9%).

D Lateness regressions

In the regressions in Table D.1, we add a control for the lateness that passengers experienced on their most recent trip with Lyft. The effect we aim to capture is the impact of reliability on future demand. In the first specification (Column 2), the lateness term is the difference between actual and expected arrival times. The difference between actual and expected time to arrival enters linearly; early arrivals correspond to negative values of the "Lateness" variable. In the second specification (Column 3), the lateness term is the difference between actual and expected times to arrival, conditional on being late, and zero otherwise. In this specification, all early arrivals are treated the same (earlier arrivals do not move the "lateness" variable more). Finally, in the third specification (Column 4), the lateness term is zero for arrivals that weren't late, and the logged difference otherwise.

In all the specifications, we find that the coefficient on lateness is negative, though very close to zero, suggesting that lateness may slightly decrease future demand, but by a relatively small amount. For example, in the first specification, we find a coefficient of -0.0001 on the difference between actual and expected times to arrival. This translates to a decrease in request rate of -0.01 percentage points per minute increase in prior lateness.

One feature to note is that these analyses only include the 62% of sessions in our data that had previous rides associated with that passenger. The request rate increases a fair amount when we drop the users who hadn't had previous rides (from 64% to 71%. Furthermore, since passengers without sessions do not show up in our data, we are only considering the effect of lateness, conditional on the passenger returning to the platform after their experience.

Table D.1
Second stage regression results, with controls for lateness of previous ride.

With controls for lateness of last experience (experiment 1)				
	No control	$(ATA - ETA)$	$\mathbb{1}_{late}(ATA - ETA)$	$\mathbb{1}_{late}\ln(ATA - ETA)$
ln(ETA)	-0.022^{***}	-0.022^{***}	-0.022^{***}	-0.022^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
$\ln(1 + \mathrm{PT})$	-0.320^{***}	-0.320^{***}	-0.320^{***}	-0.320^{***}
	(0.016)	(0.016)	(0.016)	(0.016)
Lateness	—	-0.0001^{***}	-0.0001^{**}	-0.0033^{***}
		(0.0000)	(0.0000)	(0.0003)
ETA Elasticity	-0.032^{***}	-0.032^{***}	-0.032^{***}	-0.031^{***}
	(0.003)	(0.003)	(0.003)	(0.003)
PT Elasticity	-0.459^{***}	-0.458^{***}	-0.458^{***}	-0.458^{***}
	(0.023)	(0.023)	(0.023)	(0.023)
VOT	19.28^{***}	19.28^{***}	19.27^{***}	19.01^{***}
	(1.79)	(1.79)	(1.79)	(1.79)
Control Avg. ETA	2.91	2.91	2.91	2.91
Control Avg. Price	13.44	13.44	13.44	13.44
Control Req. Rate	0.699	0.699	0.699	0.699
Controls	x	x	x	x
N	3103899	3103899	3103899	3103899
R^2	0.058	0.058	0.058	0.058

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\mathrm{ETA})$ and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Including controls for region, geohash5, local hour of week, local week of year, business user, airport session, and decile of user lifetime rides. The results in this table are from a regression of the form: $Request = \beta_0\beta_1 \ln(\mathrm{ETA}) + \beta_2 ln(1 + PT)\beta_3 Lateness + \sum_k \gamma_k X_k$, where the measure of lateness is from the most recent trip that the rider took prior to that session and X are the controls. The measures of lateness that we examine include (1) the difference between ATA and ETA for the last ride (in minutes), (2) the difference between ATA and ETA for the last ride if the arrival was late, and (3) the logged difference between ATA and ETA if the arrival was late. Arrivals were tagged as late if $ATA - ETA > 0$. We subset to sessions where the passenger had prior ride experience; this consists of 62% of the full sample.

Similarly, in the 2017 experiment, we add the controls for lateness and find that the elasticity estimates are unchanged. Removing the sessions where the user did not have a previous ride (and thus no previous lateness experience), we limit our sample to 64% of the original sample. Again, the request rate in this group is a bit higher than the request rate in the full 2017 sample.

Table D.2
Second stage regression results (2017 experiment), with controls for lateness of previous ride.

	With controls for lateness of last experience			
	No control	$(ATA - ETA)$	$\mathbb{1}_{late}(ATA - ETA)$	$\mathbb{1}_{late}\ln(ATA - ETA)$
ln(ETA)	-0.036^{***}	-0.036^{***}	-0.036^{***}	-0.036^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
$\ln(1 + \text{PT})$	-0.293^{***}	-0.292^{***}	-0.292^{***}	-0.292^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Lateness	—	-0.0000^{***}	-0.0024^{***}	-0.0069^{***}
		(0.0000)	(0.0002)	(0.0002)
ETA Elasticity	-0.051^{***}	-0.051^{***}	-0.051^{***}	-0.051^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Control Avg. ETA	3.25	3.25	3.25	3.25
Control Avg. Price	11.81	11.81	11.81	11.81
Control Req. Rate	0.714	0.714	0.714	0.714
Controls	x	x	x	x
N	6176908	6176908	6176908	6176908
R^2	0.052	0.052	0.052	0.052

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental group indicators. Controls include local hour of week and week of year, session geohash5, business user, and decile of pre-experiment lifetime rides.

E Demand Effects After End of Experiment

We next turn to analyzing whether and how exposure to the ETA and price treatments during the experiment affects passenger behavior after the experiment's conclusion. This is an important issue because to understand long-term welfare consequences from these medium-run changes in price and time changes. We use the following reduced-form approach: for each week after the experiment's conclusion, we calculate the number of sessions and the number of sessions with a request that passenger had in that week, and that passenger's average request rate (if they had at least one session). For each week after the experiment, we regress each passenger's session count, request count, and request rate on the indicators of exposure to each of the experimental treatments. If there are no lasting effects of the treatments on demand, then these regressions should find no significant differences in demand behavior across the treatment groups. Moreover, comparing results across multiple weeks of the experiment gives a sense of the rate with which lasting demand effects decay.

First, we confirm that the experimental treatments stop affecting ETA and PT after the end of experiment, by regressing $\ln(\mathrm{ETA})$ and $\ln(1 + \mathrm{PT})$ on the experimental indicators, split up across weeks. Tables E.1 and E.2 give the results of these regressions. While there are some significant treatment effects on ETA and PT in the first two weeks after the experiment,[64] by the third week after the experiment, there remains little evidence of any statistically significant effects, and what evidence does remain appears economically negligible.

Next, Tables E.3, E.4, and E.5 show the effects of the experimental treatments on users' numbers of sessions, numbers of sessions with requests, and request rates in each week after the experiment, respectively. We do not find any meaningful demand effects even twelve weeks after the end of the experiment. Passengers exposed to the high-ETA-high-price treatment have about 0.0138 fewer sessions and sessions with request per week than the control group. However, the results in Table E.5 suggests that there are no effects on demand conditional on opening up the app. Overall, while passengers exposed to higher ETAs and prices return to Lyft slightly less frequent in the immediate future, once they open the app again, their past experience does not appear to significantly affect their probability of requesting a ride.

[64] These effects were caused by a technical glitch that arose when disabling the experiment.

Table E.1
Regressions of session $\ln(\mathrm{ETA})$ on the experimental treatment indicators by week. Week 0 is first week after the end of the experiment.

	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11
High ETA High Price	0.478^{***}	0.477^{***}	-0.002	0.405^{***}	0.001	-0.000	0.007^{*}	0.004	-0.000	-0.000	0.004	0.002	0.007^{*}	0.001
	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
High ETA Normal Price	0.478^{***}	0.480^{***}	-0.003	0.397^{***}	-0.002	0.001	0.000	-0.003	0.004	-0.000	0.002	0.006^{**}	0.006^{**}	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
High ETA Low Price	0.478^{***}	0.480^{***}	-0.003	0.397^{***}	0.002	0.001	0.000	-0.002	0.004^{*}	0.001	0.006^{**}	0.001	0.001	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Normal ETA High Price	0.002	-0.000	0.245^{***}	0.003	0.002	-0.002	0.000	-0.003	0.000	0.002	0.002	-0.001	0.000	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Normal ETA Low Price	0.002	-0.001	0.244^{***}	0.001	0.002	0.002	-0.002	-0.002	0.002	0.001	0.000	0.001	0.003	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Controls	x	x	x	x	x	x	x	x	x	x	x	x	x	x
N	705870	690470	618570	614975	656352	587628	600938	600625	609180	626015	554877	500449	521414	525224
R^2	0.556	0.569	0.515	0.544	0.450	0.452	0.458	0.449	0.445	0.448	0.461	0.443	0.435	0.416
F	29676^{***}	26397^{***}	5024^{***}	15808^{***}	1	0	1	1	1	0	1	1	2	0

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. F test is of the null that all the coefficients on the instruments equal 0. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table E.2
Regressions of session $\ln(1 + \mathrm{PT})$ on the experimental treatment indicators by week. Week 0 is first week after the end of the experiment.

	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11
High ETA High Price	0.047^{***}	0.038^{***}	0.043^{***}	0.021^{***}	-0.001	0.000	0.000	0.000	-0.001	-0.001	-0.001	-0.002	-0.002	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
High ETA Normal Price	-0.001	-0.000	0.000	0.000	0.000	-0.000	-0.000	-0.000	0.002^{*}	0.001	-0.001	-0.000	-0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
High ETA Low Price	-0.031^{***}	-0.040^{***}	-0.036^{***}	-0.055^{***}	-0.000	-0.000	-0.001	0.001	-0.000	0.000	0.000	-0.000	0.001	0.000
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Normal ETA High Price	0.047^{***}	0.038^{***}	0.044^{***}	0.021^{***}	0.001	-0.001	-0.000	-0.001	0.001	0.001	-0.001	-0.001	-0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Normal ETA Low Price	-0.031^{***}	-0.039^{***}	-0.036^{***}	-0.056^{***}	0.000	0.001	-0.001	-0.000	0.001	0.000	-0.001	-0.000	-0.000	0.001
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls	x	x	x	x	x	x	x	x	x	x	x	x	x	x
N	705870	690470	618570	614975	656352	587628	600938	600625	609180	626015	554877	500449	521414	525224
R^2	0.251	0.184	0.187	0.189	0.309	0.305	0.218	0.261	0.296	0.307	0.207	0.224	0.200	0.183
F	3895^{***}	4813^{***}	4638^{***}	3760^{***}	0	0	0	1	1	1	1	1	1	0

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. F test is of the null that all the coefficients on the instruments equal 0. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table E.3
Regressions of number of sessions per user on the experimental treatment indicators by week. Week 0 is first week after the end of the experiment.

	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11
High ETA High Price	-0.069^{***}	-0.065^{***}	-0.066^{***}	-0.100^{***}	-0.072^{***}	-0.061^{***}	-0.049^{***}	-0.059^{***}	-0.055^{***}	-0.055^{***}	-0.042^{***}	-0.044^{***}	-0.045^{***}	-0.041^{***}
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.009)
High ETA Normal Price	-0.040^{***}	-0.031^{***}	-0.039^{***}	-0.044^{***}	-0.043^{***}	-0.037^{***}	-0.034^{***}	-0.032^{***}	-0.030^{***}	-0.033^{***}	-0.030^{***}	-0.018^{***}	-0.014^{**}	-0.015^{**}
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
High ETA Low Price	-0.026^{***}	-0.026^{***}	-0.018^{***}	-0.024^{***}	-0.019^{***}	-0.023^{***}	-0.012^{*}	-0.020^{***}	-0.022^{***}	-0.022^{***}	-0.020^{***}	-0.015^{**}	-0.011^{*}	-0.016^{***}
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
Normal ETA High Price	-0.032^{***}	-0.027^{***}	-0.028^{***}	-0.039^{***}	-0.034^{***}	-0.024^{***}	-0.029^{***}	-0.020^{***}	-0.022^{***}	-0.026^{***}	-0.024^{***}	-0.013^{**}	-0.016^{***}	-0.012^{*}
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
Normal ETA Low Price	0.016^{**}	0.016^{**}	0.020^{***}	0.003	0.016^{**}	0.008	0.008	0.007	0.015^{**}	0.009	0.002	0.008	0.006	0.004
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
Controls	x	x	x	x	x	x	x	x	x	x	x	x	x	x
N	720059	720059	720059	720059	720059	720059	720059	720059	720059	720059	720059	720059	720059	720059
R^2	0.080	0.078	0.097	0.111	0.111	0.116	0.116	0.120	0.119	0.111	0.108	0.104	0.106	0.102
F	19^{***}	16^{***}	20^{***}	31^{***}	21^{***}	16^{***}	12^{***}	12^{***}	13^{***}	12^{***}	9^{***}	9^{***}	7^{***}	6^{***}

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Heteroskedasticity-robust standard errors in parentheses. F test is of the null that all the coefficients on the instruments equal 0. Controls include decile of lifetime rides and indicator of business user.

Table E.4

Regressions of number of sessions with request per user on the experimental treatment indicators by week. Week 0 is first week after the end of the experiment.

	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11
High ETA High Price	-0.070^{***}	-0.071^{***}	-0.062^{***}	-0.085^{***}	-0.054^{***}	-0.043^{***}	-0.041^{***}	-0.045^{***}	-0.038^{***}	-0.040^{***}	-0.031^{***}	-0.037^{***}	-0.030^{***}	-0.029^{***}
	(0.008)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
High ETA Normal Price	-0.034^{***}	-0.033^{***}	-0.030^{***}	-0.040^{***}	-0.030^{***}	-0.024^{***}	-0.025^{***}	-0.024^{***}	-0.022^{***}	-0.024^{***}	-0.020^{***}	-0.014^{***}	-0.008^{*}	-0.011^{**}
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
High ETA Low Price	-0.018^{***}	-0.016^{***}	-0.001	-0.009^{*}	-0.010^{*}	-0.011^{**}	-0.008	-0.012^{**}	-0.014^{**}	-0.019^{***}	-0.016^{***}	-0.009^{*}	-0.005	-0.011^{**}
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Normal ETA High Price	-0.040^{***}	-0.034^{***}	-0.045^{***}	-0.042^{***}	-0.028^{***}	-0.018^{***}	-0.022^{***}	-0.009^{*}	-0.015^{***}	-0.021^{***}	-0.018^{***}	-0.013^{***}	-0.011^{**}	-0.011^{**}
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Normal ETA Low Price	0.019^{***}	0.027^{***}	0.019^{***}	0.020^{***}	0.010^{*}	0.008	0.003	0.004	0.011^{**}	0.005	-0.001	0.005	0.006	0.003
	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Controls	x	x	x	x	x	x	x	x	x	x	x	x	x	x
N	720059	720059	720059	720059	720059	720059	720059	720059	720059	720059	720059	720059	720059	720059
R^2	0.107	0.106	0.121	0.129	0.127	0.129	0.129	0.131	0.128	0.121	0.116	0.111	0.113	0.107
F	38^{***}	40^{***}	40^{***}	54^{***}	21^{***}	13^{***}	12^{***}	11^{***}	11^{***}	11^{***}	7^{***}	9^{***}	6^{***}	6^{***}

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Heteroskedasticity-robust standard errors in parentheses. F test is of the null that all the coefficients on the instruments equal 0. Controls include decile of lifetime rides and indicator of business user.

Table E.5

Regressions of average request rate per user on the experimental treatment indicators by week, for all users with at least one session in that week. Week 0 is first week after the end of the experiment.

	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11
High ETA High Price	-0.036^{***}	-0.036^{***}	-0.029^{***}	-0.038^{***}	-0.007^{*}	-0.005	-0.017^{***}	-0.012^{***}	-0.003	-0.006	0.001	-0.007^{*}	-0.003	-0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
High ETA Normal Price	-0.013^{***}	-0.015^{***}	-0.006^{**}	-0.018^{***}	-0.005^{*}	-0.002	-0.007^{**}	-0.005^{*}	-0.006^{**}	-0.004	-0.002	-0.003	0.001	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
High ETA Low Price	-0.003	0.002	0.012^{***}	0.004	-0.001	0.000	-0.003	-0.002	-0.003	-0.007^{***}	-0.006^{**}	-0.004	0.000	-0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Normal ETA High Price	-0.022^{***}	-0.018^{***}	-0.033^{***}	-0.020^{***}	-0.007^{***}	-0.005^{*}	-0.002	0.003	-0.001	-0.006^{**}	-0.002	-0.006^{**}	-0.004	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Normal ETA Low Price	0.011^{***}	0.018^{***}	0.007^{***}	0.023^{***}	-0.001	0.005^{*}	0.000	-0.002	0.003	0.001	0.001	0.001	-0.001	0.004
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Controls	x	x	x	x	x	x	x	x	x	x	x	x	x	x
N	246021	241777	211156	209537	214997	197686	199483	197628	197285	200569	186401	179396	184503	185825
R^2	0.061	0.065	0.072	0.070	0.057	0.059	0.058	0.054	0.051	0.048	0.050	0.044	0.044	0.041
F	44^{***}	54^{***}	55^{***}	64^{***}	2^{**}	2^{**}	4^{***}	3^{**}	2	3^{**}	1	2	1	1

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Heteroskedasticity-robust standard errors in parentheses. F test is of the null that all the coefficients on the instruments equal 0. Controls include decile of lifetime rides and indicator of business user.

F Time to Destination

In this section, we consider the potential bias that may arise in our estimate of the price equivalent of waiting time (PEWT) by failing to account for the effect of in-vehicle time on demand. Suppose the utility a passenger receives from taking a Lyft trip at Prime Time level PT, waiting time (ETA) WT, and in-vehicle time IVT can be written

$$U = \beta_0 + \beta_1 \ln(1 + \mathrm{PT}) + \beta_2 \ln(WT) + \beta_3 \ln(IVT) + controls + \varepsilon, \tag{22}$$

but we ignore the $\ln(IVT)$ term. Note that Lyft did not show IVT estimates during the experiments, so $\ln(IVT)$ captures the passenger's perceived waiting time. Our estimate of β_2 will suffer from omitted variable bias, converging in probability to

$$\beta_2 + \delta\beta_3 \tag{23}$$

where δ is the coefficient on $\ln(WT)$ in a regression of $\ln(IVT)$ on $\ln(WT)$ and controls. Instrumenting $\ln(WT)$ with our experimental treatment indicators will *not* remove this bias if changes in quoted waiting time affect a passenger's estimate of the in-vehicle time.[65]

Using existing estimates from the literature on the relative values of waiting and in-vehicle time, we can assess the asymptotic bias of our estimate of the price-equivalent of waiting time (PEWT). In particular, given the ommitted variable bias from $\ln(IVT)$, our estimator of $PEWT$ is also biased in the limit:

$$\begin{aligned}
\widehat{PEWT} &= \frac{\hat{\beta}_2 \overline{P}}{\hat{\beta}_1 \overline{WT}} && (24)\\
&\xrightarrow{p} \frac{(\beta_2 + \delta\beta_3)\overline{P}}{\beta_1 \overline{WT}} && (25)\\
&= \frac{\beta_2 \overline{P}}{\beta_1 \overline{WT}} + \delta \frac{\overline{IVT}}{\overline{WT}} \frac{\beta_3 \overline{P}}{\beta_1 \overline{IVT}} && (26)\\
&= PEWT + \delta \frac{\overline{IVT}}{\overline{WT}} PEIVT, && (27)
\end{aligned}$$

where $PEIVT$ is the price-equivalent of in-vehicle time. Letting $\kappa = \frac{PEIVT}{PEWT}$, we conclude

$$\widehat{PEWT} \xrightarrow{p} \left(1 + \delta\kappa \frac{\overline{IVT}}{\overline{WT}}\right) PEWT, \tag{28}$$

so that our estimate of the price equivalent of waiting time is biased upwards by a factor of $\left(1 + \delta\kappa \frac{\overline{IVT}}{\overline{WT}}\right)$ in the limit.

[65] By design of the experiment, the treatments only impact waiting time and not in-vehicle time. Nonetheless, a long waiting time may lead a passenger to believe that the in-vehicle time will also be longer, by changing the passenger's perception of local traffic conditions.

$\frac{\overline{IVT}}{\overline{WT}}$ can be directly computed for completed rides in our sample, and has a value of about 5. δ is a parameter reflecting passenger expectations about the relationship between waiting time and in-vehicle time, and so cannot be directly estimated; we approximate it by running a cross-sectional regression of $\ln(IVT)$ on $\ln(WT)$ and $controls$ for completed rides in our sample, which should capture the (endogenous) relationship between waiting time and in-vehicle time arising from traffic conditions.[66] Using data on completed rides from the 2017 experiment, we estimate this value to be 0.1, which we may interpret to mean that a 10% increase in waiting time is associated with a 1% increase in expected in-vehicle time. We take $\kappa = 1/2$.

Plugging these numbers into equation (28), we conclude that the omitted variable bias from in-vehicle time inflates our PEWT estimate by about 25%. Thus, if one believes the purported relationship between quoted waiting time and passenger-estimated in-vehicle time (which, unfortunately, cannot be tested empirically), one should adjust our PEWT estimates downward by a factor of 1.25, to arrive at an overall PEWT of $15/hour rather than $19/hour.

For the 2017 experiment, we observe estimated in-vehicle time for all sessions that entered a destination, and can use this data to directly evaluate whether our estimated waiting time effects also reflect the value of in-vehicle time. Table F.1 gives ETA elasticity estimates broken up by quintile of estimated time to drop off. We do not find any clear pattern in the elasticity estimates, supporting the conclusion that our experiments are primarily estimating the value of pre-pickup waiting time, and not a combination of waiting and in-vehicle time.

Table F.1
ETA elasticities by quintile of estimated trip duration (2017 experiment).

	Q1	Q2	Q3	Q4	Q5
ln(ETA)	-0.0190^{***}	-0.0146^{***}	-0.0159^{***}	-0.0163^{***}	-0.0133^{***}
	(0.0008)	(0.0008)	(0.0009)	(0.0010)	(0.0011)
$\ln(1+\text{PT})$	-0.2667^{***}	-0.3242^{***}	-0.3771^{***}	-0.3880^{***}	-0.3587^{***}
	(0.0020)	(0.0019)	(0.0019)	(0.0019)	(0.0021)
ETA Elasticity	-0.0233^{***}	-0.0181^{***}	-0.0210^{***}	-0.0236^{***}	-0.0233^{***}
	(0.0010)	(0.0010)	(0.0012)	(0.0015)	(0.0020)
Control Avg. ETA	3.16	3.20	3.22	3.33	3.80
Control Req. Rate	0.819	0.810	0.762	0.693	0.575
Controls	x	x	x	x	x
N	1598111	1596340	1593377	1592525	1594672
R^2	0.073	0.077	0.085	0.081	0.070

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides. Bins are quintiles of estimated in-vehicle time, in seconds. The ordered quintiles are 561, 772, 1032, 1455.

[66] The inclusion of location and time controls should partial out some of the endogenous relationship unrelated to traffic conditions, arising, for example, due to the fact that passengers in less driver-dense regions may take longer rides in general.

G Removing always- and/or never-requesters

In this section, we consider how the estimates of interest change when excluding always-requesters, excluding never-requesters, and excluding both. Our identification strategy of the value of time rests on the assumption that users are cognisant of and internalizing the prices and times displayed in making their decision to request a ride. One concern, however, may be that these estimates include users who do *not* make a decision based on displayed times and prices. While our data does not allow us to precisely determine what enters our users' decision-making process, we can explore how robust our results are to narrowing our focus on passengers who sometimes request and sometimes do not.

Notably, the proportion of always- and/or never-requesters will differ by treatment group, but not by drastic proportions, with most treatments consisting of about 20-22% always-requesters and 19-21% never-requesters.

Table G.1
Proportion of always- and never-requesters by treatment group (2015 experiment).

Treatment group	Number of users	Prop. always requesters	Prop. never requesters	Prop. always/never requesters
Control	292025	0.21 (0.001)	0.2 (0.001)	0.41 (0.001)
High ETA High Price	38674	0.19 (0.002)	0.22 (0.002)	0.42 (0.003)
High ETA Low Price	97254	0.21 (0.001)	0.2 (0.001)	0.41 (0.002)
High ETA Normal Price	97051	0.2 (0.001)	0.21 (0.001)	0.41 (0.002)
Normal ETA High Price	97185	0.2 (0.001)	0.21 (0.001)	0.41 (0.002)
Normal ETA Low Price	97870	0.21 (0.001)	0.19 (0.001)	0.41 (0.002)

Similarly, we can look at how these proportions differ across regions. There is a fair amount of variation across regions. For example, the proportion of never-requesters ranges from 15.8% (San Francisco) to 28% (Atlanta), reflecting differences in request patterns and Lyft app usage across various locations.

Table G.2
Proportion of always- and never-requesters by region.

Region	Number of users	Prop. always requesters	Prop. never requesters	Prop. always/never requesters
Atlanta	37238	0.19 (0.002)	0.27 (0.002)	0.47 (0.003)
Austin	34828	0.25 (0.002)	0.18 (0.002)	0.43 (0.003)
New York City	118762	0.18 (0.001)	0.24 (0.001)	0.42 (0.001)
Boston	51332	0.19 (0.002)	0.21 (0.002)	0.4 (0.002)
Los Angeles	169031	0.22 (0.001)	0.2 (0.001)	0.42 (0.001)
Miami	75468	0.18 (0.001)	0.21 (0.001)	0.38 (0.002)
San Diego	47769	0.25 (0.002)	0.23 (0.002)	0.47 (0.002)
Seattle	36003	0.21 (0.002)	0.19 (0.002)	0.4 (0.003)
San Francisco	149628	0.21 (0.001)	0.15 (0.001)	0.36 (0.001)

In the following tables, we remove all sessions of always-requesters (users with a request rate of 100%), of never-requesters (users with a request rate of 0%), and both.

We perform this analysis for various subsets of our population that may be of particular interest and find similar patterns in changes to time and price elasticities and values of time across most of these subsets. Removing the set of sessions of always-requesters tends to yield more elastic estimates of both price and time and a similar value of time, while removing the sessions of never-requesters tends to yield less elastic estimates of price and time, with price elasticity changing much more than the time elasticity, thus decreasing the value of time.

Table G.3
Second-stage regressions with and without always-/never-requesters.

	Full sample	Without never and always	Without never	Without always
$\ln(\text{ETA})$	-0.026^{***}	-0.027^{***}	-0.022^{***}	-0.023^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
$\ln(1+\text{PT})$	-0.367^{***}	-0.365^{***}	-0.332^{***}	-0.334^{***}
	(0.013)	(0.014)	(0.013)	(0.014)
ETA Elasticity	-0.043^{***}	-0.044^{***}	-0.035^{***}	-0.037^{***}
	(0.003)	(0.003)	(0.003)	(0.003)
PT Elasticity	-0.594^{***}	-0.592^{***}	-0.539^{***}	-0.541^{***}
	(0.021)	(0.023)	(0.021)	(0.022)
VOT	19.38^{***}	20.11^{***}	17.73^{***}	18.48^{***}
	(1.39)	(1.48)	(1.53)	(1.60)
Control Avg. ETA	3.08	3.08	3.08	3.08
Control Avg. Price	13.83	13.83	13.83	13.83
Control Req. Rate	0.620	0.620	0.620	0.620
Controls	x	x	x	x
N	5177358	4806027	4921221	4549890
R^2	0.072	0.084	0.056	0.065

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\text{ETA})$ and $\ln(1+\text{PT})$ instrumented by experimental group indicators. Always- (never-) requesters are passengers who have a request rate of 100% (0%) during experimental period. Including controls for geohash5, local hour of week, local week of year, business user, airport session, and decile of user lifetime rides.

Table G.4
Second-stage regressions with and without always-/never-requesters, for business user sessions.

	Full sample	Without never and always	Without never	Without always
ln(ETA)	-0.023^{**}	-0.019^{*}	-0.020^{**}	-0.016
	(0.010)	(0.010)	(0.010)	(0.010)
$\ln(1 + \text{PT})$	-0.283^{***}	-0.282^{***}	-0.274^{***}	-0.273^{***}
	(0.052)	(0.053)	(0.052)	(0.053)
ETA Elasticity	-0.033^{**}	-0.027^{*}	-0.028^{**}	-0.023
	(0.014)	(0.014)	(0.014)	(0.014)
PT Elasticity	-0.398^{***}	-0.397^{***}	-0.385^{***}	-0.384^{***}
	(0.073)	(0.076)	(0.072)	(0.075)
VOT	24.42^{**}	20.37^{*}	21.62^{**}	17.78
	(10.55)	(10.71)	(10.85)	(11.01)
Control Avg. ETA	2.30	2.30	2.30	2.30
Control Avg. Price	11.38	11.38	11.38	11.38
Control Req. Rate	0.714	0.714	0.714	0.714
Controls	x	x	x	x
N	217180	205970	213224	202014
R^2	0.096	0.108	0.081	0.090

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1 + \text{PT})$ instrumented by experimental group indicators. Always- (never-) requesters are passengers who have a request rate of 100% (0%) during experimental period. Including controls for geohash5, local hour of week, local week of year, airport session, and decile of user lifetime rides.

Table G.5
Second-stage regressions with and without always-/never-requesters, for airport sessions.

	Full sample	Without never and always	Without never	Without always
ln(ETA)	-0.014	-0.020^{**}	-0.006	-0.011
	(0.009)	(0.010)	(0.010)	(0.010)
ln(1 + PT)	-0.335^{***}	-0.283^{***}	-0.248^{**}	-0.204^{**}
	(0.092)	(0.096)	(0.097)	(0.102)
ETA Elasticity	-0.024	-0.033^{*}	-0.009	-0.018
	(0.015)	(0.018)	(0.015)	(0.017)
PT Elasticity	-0.553^{***}	-0.468^{***}	-0.409^{***}	-0.337^{**}
	(0.153)	(0.175)	(0.149)	(0.171)
VOT	27.55	45.57^{*}	14.91	34.91
	(18.02)	(24.68)	(24.52)	(33.70)
Control Avg. ETA	2.92	2.92	2.92	2.92
Control Avg. Price	31.40	31.40	31.40	31.40
Control Req. Rate	0.606	0.606	0.606	0.606
Controls	x	x	x	x
N	129088	112222	119819	102953
R^2	0.080	0.087	0.077	0.081

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\text{ETA})$ and $\ln(1+\text{PT})$ instrumented by experimental group indicators. Always- (never-) requesters are passengers who have a request rate of 100% (0%) during experimental period. Including controls for region, geohash5, local hour of week, local week of year, business user, and decile of user lifetime rides.

Table G.6
Second-stage regression results without always- and never-requesters, by region.

	San Francisco	Miami	Los Angeles	Boston	Seattle	San Diego	Austin	Atlanta	New York City
ln(ETA)	-0.0170^{***}	-0.0219^{***}	-0.0236^{***}	-0.0381^{***}	-0.0266^{***}	-0.0223^{***}	-0.0166^{**}	-0.0179^{*}	-0.0265^{***}
	(0.0035)	(0.0063)	(0.0042)	(0.0066)	(0.0071)	(0.0072)	(0.0083)	(0.0098)	(0.0048)
$\ln(1+\mathrm{PT})$	-0.2785^{***}	-0.2310^{***}	-0.2939^{***}	-0.3399^{***}	-0.2886^{***}	-0.3581^{***}	-0.4752^{***}	-0.4262^{***}	-0.4579^{***}
	(0.0189)	(0.0641)	(0.0485)	(0.0444)	(0.0507)	(0.1030)	(0.1161)	(0.1582)	(0.0256)
ETA Elasticity	-0.0239^{***}	-0.0374^{***}	-0.0378^{***}	-0.0646^{***}	-0.0429^{***}	-0.0365^{***}	-0.0259^{**}	-0.0315^{*}	-0.0512^{***}
	(0.0050)	(0.0108)	(0.0067)	(0.0111)	(0.0114)	(0.0118)	(0.0130)	(0.0172)	(0.0092)
PT Elasticity	-0.3921^{***}	-0.3947^{***}	-0.4708^{***}	-0.5765^{***}	-0.4648^{***}	-0.5866^{***}	-0.7410^{***}	-0.7497^{***}	-0.8830^{***}
	(0.0266)	(0.1094)	(0.0777)	(0.0754)	(0.0817)	(0.1689)	(0.1811)	(0.2783)	(0.0494)
VOT	20.06^{***}	17.74^{***}	20.13^{***}	23.64^{***}	24.14^{***}	14.98^{**}	10.03^{**}	6.92^{*}	23.19^{***}
	(4.16)	(6.40)	(4.75)	(4.65)	(7.24)	(5.91)	(4.93)	(4.19)	(4.22)
Control Avg. ETA	2.08	4.35	2.97	3.61	3.36	3.40	2.77	4.92	3.09
Control Avg. Price	11.39	13.56	12.42	12.70	14.66	13.64	13.25	13.53	20.59
Control Req. Rate	0.710	0.588	0.629	0.596	0.627	0.612	0.644	0.569	0.522
Controls	x	x	x	x	x	x	x	x	x
N	1247246	441748	893844	303445	236524	234723	185909	190384	816067
R^2	0.069	0.039	0.059	0.049	0.065	0.062	0.050	0.041	0.029

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\mathrm{ETA})$ and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Always- (never-) requesters are passengers who have a request rate of 100% (0%) during experimental period. Including controls for user geohash5, local hour of week, local week of year, business user, airport session, and decile of user lifetime rides.

Table G.7
Second-stage regression results without always- and never-requesters, by local day of week.

	Mon	Tues	Wed	Thu	Fri	Sat	Sun
ln(ETA)	-0.0248^{***}	-0.0262^{***}	-0.0262^{***}	-0.0191^{***}	-0.0212^{***}	-0.0208^{***}	-0.0227^{***}
	(0.0033)	(0.0033)	(0.0033)	(0.0031)	(0.0030)	(0.0028)	(0.0029)
$\ln(1+\mathrm{PT})$	-0.3491^{***}	-0.3838^{***}	-0.3618^{***}	-0.3112^{***}	-0.2791^{***}	-0.3343^{***}	-0.3600^{***}
	(0.0289)	(0.0318)	(0.0312)	(0.0226)	(0.0155)	(0.0204)	(0.0191)
ETA Elasticity	-0.0409^{***}	-0.0417^{***}	-0.0416^{***}	-0.0310^{***}	-0.0338^{***}	-0.0329^{***}	-0.0380^{***}
	(0.0055)	(0.0052)	(0.0052)	(0.0051)	(0.0047)	(0.0044)	(0.0049)
PT Elasticity	-0.5750^{***}	-0.6116^{***}	-0.5733^{***}	-0.5055^{***}	-0.4453^{***}	-0.5290^{***}	-0.6011^{***}
	(0.0476)	(0.0507)	(0.0495)	(0.0367)	(0.0247)	(0.0323)	(0.0318)
VOT	18.52^{***}	17.92^{***}	20.01^{***}	16.91^{***}	19.43^{***}	17.56^{***}	16.91^{***}
	(2.73)	(2.54)	(2.83)	(2.85)	(2.72)	(2.40)	(2.25)
Control Avg. ETA	3.10	3.00	2.90	3.02	3.32	2.84	3.09
Control Avg. Price	13.45	13.14	13.35	13.88	14.15	13.34	13.79
Control Req. Rate	0.610	0.631	0.634	0.616	0.627	0.634	0.603
Controls	x	x	x	x	x	x	x
N	549814	561522	553726	635104	796294	785743	667687
R^2	0.075	0.070	0.070	0.063	0.058	0.070	0.078

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\mathrm{ETA})$ and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Always- (never-) requesters are passengers who have a request rate of 100% (0%) during experimental period. Including controls for region, user geohash5, local hour of week, local week of year, business user, airport session, and decile of user lifetime rides.

Table G.8
Second-stage regression results without always- and never-requesters, by time of day (weekdays).

	11PM-6AM	6AM-10AM	10AM-4PM	4PM-7PM	7PM-11PM
ln(ETA)	-0.0419^{***}	-0.0224^{***}	-0.0230^{***}	-0.0131^{***}	-0.0214^{***}
	(0.0040)	(0.0035)	(0.0034)	(0.0030)	(0.0038)
$\ln(1+\mathrm{PT})$	-0.3681^{***}	-0.3654^{***}	-0.2814^{***}	-0.3006^{***}	-0.3173^{***}
	(0.0234)	(0.0374)	(0.0220)	(0.0272)	(0.0214)
ETA Elasticity	-0.0631^{***}	-0.0380^{***}	-0.0373^{***}	-0.0206^{***}	-0.0345^{***}
	(0.0060)	(0.0059)	(0.0055)	(0.0048)	(0.0061)
PT Elasticity	-0.5549^{***}	-0.6208^{***}	-0.4560^{***}	-0.4747^{***}	-0.5129^{***}
	(0.0353)	(0.0636)	(0.0356)	(0.0430)	(0.0347)
VOT	26.34^{***}	17.19^{***}	20.56^{***}	13.14^{***}	17.09^{***}
	(2.77)	(3.04)	(3.19)	(3.11)	(2.98)
Control Avg. ETA	3.51	2.92	3.13	2.55	3.72
Control Avg. Price	13.55	13.68	13.10	12.85	15.74
Control Req. Rate	0.667	0.592	0.620	0.634	0.618
Controls	x	x	x	x	x
N	491568	754771	581334	781610	487177
R^2	0.079	0.062	0.076	0.077	0.057

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Always- (never-) requesters are passengers who have a request rate of 100% (0%) during experimental period. Including controls for user geohash5, region, local week of year, business user, airport session, and decile of user lifetime rides.

Table G.9
Second-stage regression results without always- and never-requesters, by time of day (weekends).

	11PM-6AM	6AM-10AM	10AM-4PM	4PM-7PM	7PM-11PM
ln(ETA)	-0.0284^{***}	-0.0262^{***}	-0.0246^{***}	-0.0194^{***}	-0.0136^{***}
	(0.0065)	(0.0040)	(0.0043)	(0.0041)	(0.0037)
$\ln(1+\mathrm{PT})$	-0.3554^{***}	-0.3831^{***}	-0.3455^{***}	-0.3647^{***}	-0.3068^{***}
	(0.0470)	(0.0278)	(0.0249)	(0.0414)	(0.0209)
ETA Elasticity	-0.0515^{***}	-0.0459^{***}	-0.0411^{***}	-0.0307^{***}	-0.0200^{***}
	(0.0117)	(0.0070)	(0.0071)	(0.0065)	(0.0054)
PT Elasticity	-0.6442^{***}	-0.6706^{***}	-0.5764^{***}	-0.5784^{***}	-0.4520^{***}
	(0.0852)	(0.0487)	(0.0416)	(0.0656)	(0.0308)
VOT	16.99^{***}	17.53^{***}	18.35^{***}	16.01^{***}	13.88^{***}
	(4.23)	(2.80)	(3.26)	(3.59)	(3.70)
Control Avg. ETA	4.10	3.07	3.10	2.56	2.74
Control Avg. Price	14.51	13.10	13.32	12.84	14.35
Control Req. Rate	0.556	0.575	0.603	0.633	0.679
Controls	x	x	x	x	x
N	120709	363839	244162	335387	389333
R^2	0.077	0.076	0.083	0.082	0.058

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Always- (never-) requesters are passengers who have a request rate of 100% (0%) during experimental period. Including controls for user geohash5, region, local week of year, business user, airport session, and decile of user lifetime rides.

Table G.10
Second stage regression results, with controls for lateness of previous ride, excluding always-/never-requesters.

	With controls for lateness of last experience			
	No control	$(ATA - ETA)$	$\mathbb{1}_{late}(ATA - ETA)$	$\mathbb{1}_{late}log(ATA - ETA)$
ln(ETA)	-0.023^{***}	-0.023^{***}	-0.023^{***}	-0.023^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
$\ln(1 + \mathrm{PT})$	-0.324^{***}	-0.324^{***}	-0.324^{***}	-0.324^{***}
	(0.016)	(0.016)	(0.016)	(0.016)
Lateness	—	-0.0001^{***}	-0.0001^{***}	-0.0035^{***}
		(0.0000)	(0.0000)	(0.0003)
ETA Elasticity	-0.034^{***}	-0.034^{***}	-0.034^{***}	-0.034^{***}
	(0.003)	(0.003)	(0.003)	(0.003)
PT Elasticity	-0.481^{***}	-0.481^{***}	-0.481^{***}	-0.481^{***}
	(0.024)	(0.024)	(0.024)	(0.024)
VOT	19.46^{***}	19.46^{***}	19.45^{***}	19.18^{***}
	(1.83)	(1.83)	(1.83)	(1.83)
Control Avg. ETA	2.93	2.93	2.93	2.93
Control Avg. Price	13.35	13.35	13.35	13.35
Control Req. Rate	0.676	0.676	0.676	0.676
Controls	x	x	x	x
N	2874921	2874921	2874921	2874921
R^2	0.062	0.062	0.062	0.062

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\mathrm{ETA})$ and $\ln(1 + \mathrm{PT})$ instrumented by experimental group indicators. Always- (never-) requesters are passengers who have a request rate of 100% (0%) during experimental period. Including controls for user geohash5, region, local week of year, business user, airport session, and decile of user lifetime rides. The results in this table are from a regression of the form: $Request = \beta_0 + \beta_1 \ln(\mathrm{ETA}) + \beta_2 \ln(1+\mathrm{PT})\beta_3 Lateness + \sum_k \gamma_k X_k$, where the measure of lateness is from the most recent trip that the rider took prior to that session and X are the controls. The measures of lateness that we examine include (1) the difference between ATA and ETA for the last ride (in seconds), (2) the difference between ATA and ETA for the last ride if the arrival was late, and (3) the logged difference between ATA and ETA if the arrival was late. Arrivals were tagged as late if $ATA - ETA > 0$. We subset to sessions where the passenger had prior ride experience; this consists of 62% of the full sample.

H External Validity

One concern with our elasticity and VOT estimates is that they are based on a particular subpopulation—Lyft passengers in the sampled cities—and hence may not generalize to the broader population.

To assess the degree of this external validity problem, we use person-level data from the 2017 National Household Travel Survey (Federal Highway Administration, 2017). This survey reports person-level data on a number of a demographic characteristics, including race, gender, income, and education; person-level weights, which can be used to extrapolate the survey results to the full U.S. population; and detailed information on travel patterns, including the answer to this question about use of rideshare services: *"In the past 30 days, how many times have you purchased a ride with a smartphone rideshare app (e.g. Uber, Lyft, Sidecar)?"*

We define "Rideshare Users" to be all persons in the survey who reported a positive number to this question. In Figure H.1, we subset the NHTS data to persons who live in core-based statistical areas (CBSAs) corresponding to the eight regions in our first experiment, and compare "Rideshare Users" to the general population in the survey along a number of demographic dimensions. Throughout this analysis, we use the person-level weights provided by the NHTS to account for sample selection of the survey itself.

We note from the graphs that while rideshare users have similar race and gender distributions to the full population, they tend to skew toward the upper end of the income and education distributions, bunch up in the 20–40 years old portion of the age distribution, and skew toward the lower end of the household size distribution. Rideshare users are also more likely to reside in urban areas.

This evidence suggests that the population of rideshare users does differ systematically from the general population, implying that our value of time and elasticity estimates may not readily generalize to non-rideshare users.

To measure the extent of the bias resulting from this selection, we consider reweighting our data by the inverse of each passenger's propensity to use rideshare. Specifically, we proceed as follows:

1. Use the NHTS person-level data to estimate a probit model of an individual's propensity to use rideshare in the regions in our first experiment, as a function of their CBSA, income level, and home Census tract population density.[67]

2. For each user in the first experiment, we impute a home location (either a latitude-longitude pair or a ZIP Code Tabulation Area), and use this home location to estimate their CBSA, Census tract, and household income.[68]

[67] The NHTS person-level weights are used in the estimation of the probit model.

[68] Passenger home locations are inferred through a combination of billing addresses, in-app passenger-set shortcuts, and frequent pickup and destination locations. Income is imputed as the median household income in the home Census block group (when precise home coordinates are available) or ZCTA (when only a home ZIP is available) using data from the 2017 American Community Survey (ACS). ZIP codes were mapped to CBSAs and Census tracts using the HUD-USPS ZIP Code Crosswalk data.

3. Using the model estimated in Step 1 and the demographic features imputed in Step 2, we calculate the propensity to use rideshare of each user in our first experiment.

4. We run our main specification, using inverse propensity weights based on the propensities computed in Step 3 and subsetting to each user's first session in the experiment (to account for the fact our weights are at the user-level rather than the session level).

Table H.1 shows the results, both with and without the inverse probability weights. Some users were dropped, as median income data for their modal Census tracts was unavailable.

Table H.1
2SLS regression results for each passenger's first session in the experiment, both with and without inverse propensity weighting by propensity to use rideshare services.

	Unweighted	Weighted
$\ln(\text{ETA})$	-0.0264^{***}	-0.0230^{***}
	(0.0027)	(0.0046)
$\ln(1+\text{PT})$	-0.3439^{***}	-0.3098^{***}
	(0.0180)	(0.0353)
ETA Elasticity	-0.0447^{***}	-0.0390^{***}
	(0.0046)	(0.0079)
PT Elasticity	-0.5833^{***}	-0.5253^{***}
	(0.0305)	(0.0600)
VOT	23.39^{***}	22.84^{***}
	(2.50)	(4.89)
Control Avg. ETA	3.35	3.53
Control Avg. Price	17.05	18.12
Control Req. Rate	0.593	0.594
Controls	x	x
N	694703	694703
R^2	0.070	0.073

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\text{ETA})$ and $\ln(1+\text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

The inverse propensity weighting decreases both elasticity estimates. At the same time, in the reweighted sample, control average ETAs are slightly higher and control average prices slightly lower. These effects together result in an estimated VOT about 2% smaller than that which is estimated without weighting. This result suggests that our estimates are not significantly biased by the selective nature of the subpopulation of rideshare users. Some caveats must be kept in mind: our demographic features are imputed rather than actual, and our propensity score model estimates the propensity to use any rideshare services within a month, which may differ from the propensity to use Lyft over a ten-week period.

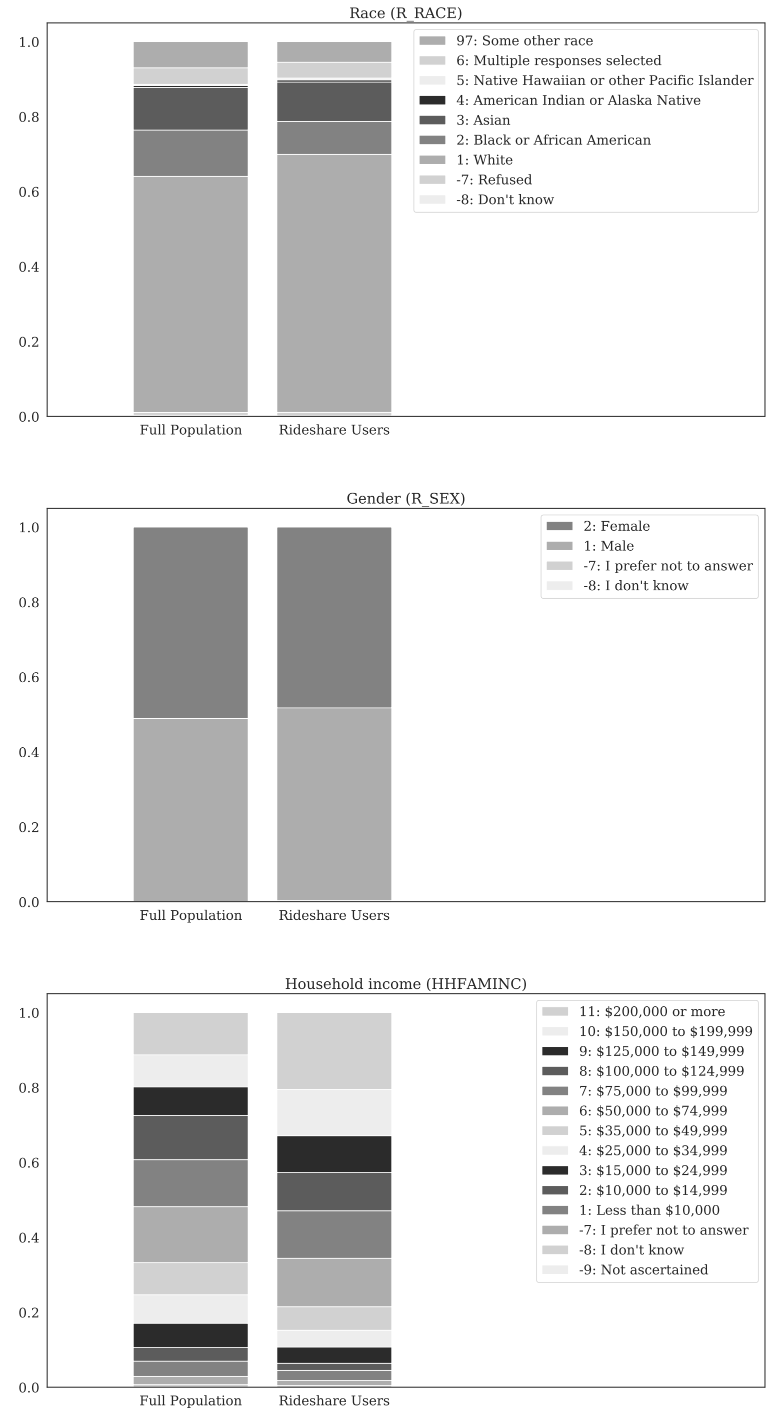
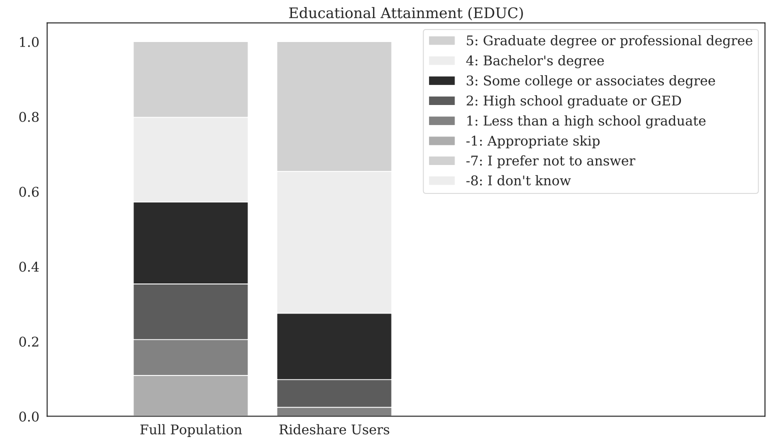
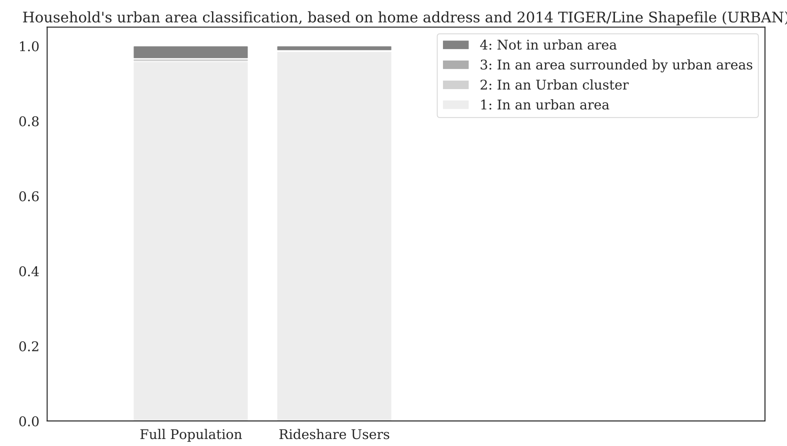


Figure H.1
Comparison of demographic characteristics of rideshare users and the general population in the regions sampled for experiment 1, based on data from the National Household Travel Survey (Federal Highway Administration, 2017).





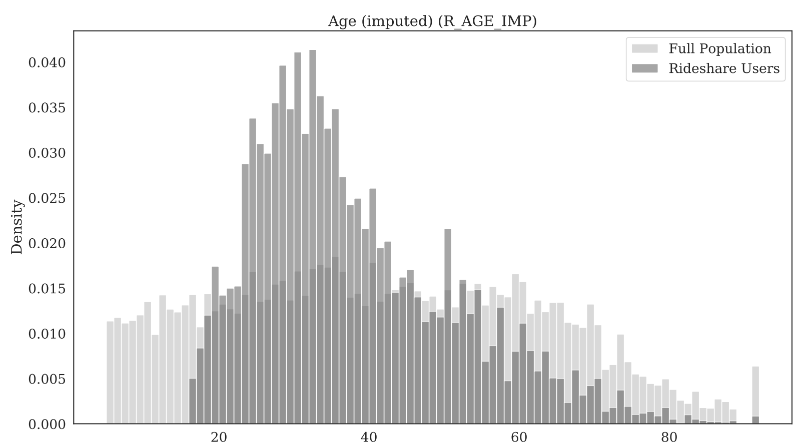


Figure H.1
Comparison of demographic characteristics of rideshare users and the general population in the regions sampled for experiment 1, based on data from the National Household Travel Survey (Federal Highway Administration, 2017) (continued).

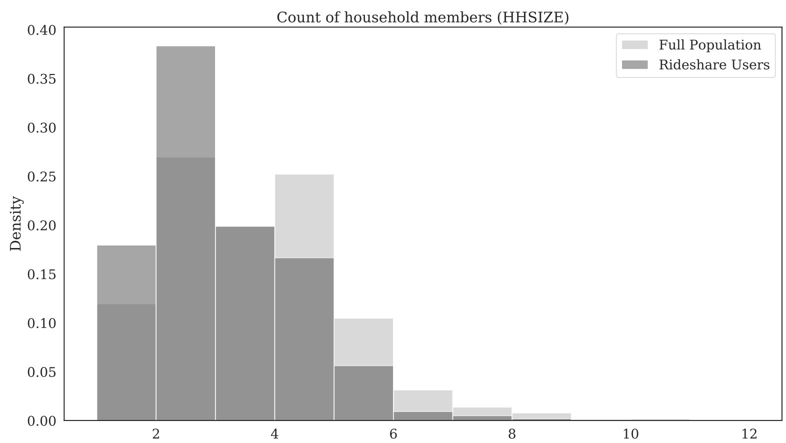


Figure H.1
Comparison of demographic characteristics of rideshare users and the general population in the regions sampled for experiment 1, based on data from the National Household Travel Survey (Federal Highway Administration, 2017) (continued).

I Correcting for User Selection into Ride Contexts

The patterns of heterogeneity we document across contexts do not take into account user self-selection into these contexts. As a consequence, though our heterogeneity analysis suggests, for example, that the VOT is higher in passenger sessions occurring during commuting times than other times of the week, it does not let us say whether this context "causes" higher VOTs—it may instead be the case that passengers with higher VOTs are more likely to have sessions during commuting times.

The purpose of this analysis is to separate contextual and selection effects in the observed heterogeneity, by weighting each user i's sessions in context x by the inverse of an estimate of user i's propensity to have a session in context x, $\hat{p}_{i,x}$.

We can estimate these propensities using pre-experiment data as the observed fraction of user i's sessions that occur in context x. The propensity is estimated using $n_{i,x}$, the number of user i's pre-experiment sessions that occurred in context x:

$$\hat{p}_{x,i} = \frac{n_{x,i}}{\sum_{x' \in X} n_{x',i}}. \tag{29}$$

These estimates have high variances. The estimated propensities often equal 0 (in which case their inverses are undefined), and do not exist for users with no pre-experiment sessions.

To address these issues, we instead adopt an empirical Bayes estimation strategy. For each range of contexts X (for example, the different days of the week, or {downtown, non-downtown}), we assume that the set of vectors $\{(p_{i,x})_{x \in X}\}$ is composed of independent draws from a Dirichlet distribution with parameter vector $\boldsymbol{\alpha}_X$. $\boldsymbol{\alpha}_X$ can be estimated from the data. Conditional on $\boldsymbol{\alpha}_X$, we note that $\{(n_{x,i})_{x \in X}\}$ are independent draws from the Dirichlet-multinomial distribution (the multi-category generalization of the beta-binomial distribution), with joint density proportional to

$$\prod_i \left(\frac{\Gamma(\sum_x \alpha_x)}{\Gamma(\sum_x n_{x,i} + \alpha_x)} \prod_x \frac{\Gamma(n_{x,i} + \alpha_x)}{\Gamma(\alpha_x)} \right) \tag{30}$$

(Mosimann, 1962). From (30), $\boldsymbol{\alpha}_X$ can be estimated by maximum likelihood. We then take the maximum a posteriori estimate of each vector $(p_{i,x})_{x \in X}$ based on the data $(n_{i,x})_{x \in X}$ and the prior $\mathrm{Dirichlet}(\hat{\boldsymbol{\alpha}}_X)$. The resulting estimates are

$$\hat{p}_{x,i} = \frac{n_{x,i} + \hat{\alpha}_x}{\sum_{x'} n_{x',i} + \hat{\alpha}_{x'}}, \tag{31}$$

from which we see that the empirical Bayes approach amounts to shrinking the basic estimates $n_{x,i} / \sum_{x'} n_{x',i}$ toward $\hat{\alpha}_x / \sum_{x'} \hat{\alpha}_{x'}$, where estimates for users with fewer pre-experiment observations are shrunk more.[69]

Tables I.1 through I.5 reproduce heterogeneity analyses for whether downtown, whether at an airport, day and time of week, and distance to public transit using the inverses of the empirical Bayes propensity estimates as observation weights. The non-downtown vs. downtown results do not show a higher VOT in downtown sessions, suggesting that the observed higher VOT in the unweighted data may be attributable to user selection. Results for non-airport vs. airport sessions are similar, though the airport estimates remain very imprecise due to the small sample size. In contrast, the results for day and time of week are similar to those found in the unweighted data. We find that the VOT is highest in the morning commute, suggesting that this result is not primarily due to user selection. Results using inverse propensity weights based on distance to transit bins are also similar to those observed in the unweighted data.

Table I.1
2SLS regression results by whether session is downtown, using empirical Bayes inverse propensity weights.

	Non-downtown	Downtown
ln(ETA)	-0.0397^{***}	-0.0233^{***}
	(0.0066)	(0.0033)
ln(1 + PT)	-0.3874^{***}	-0.3425^{***}
	(0.0666)	(0.0199)
ETA Elasticity	-0.0740^{***}	-0.0360^{***}
	(0.0125)	(0.0049)
PT Elasticity	-0.7219^{***}	-0.5288^{***}
	(0.1259)	(0.0300)
VOT	29.63^{***}	24.17^{***}
	(6.34)	(3.47)
Control Avg. ETA	4.14	2.34
Control Avg. Price	19.96	13.83
Control Req. Rate	0.533	0.674
Controls	x	x
N	1986393	3190965
R^2	0.059	0.057

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\text{ETA})$ and $\ln(1 + \text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

We compare these inverse-propensity-weighted estimates to those obtained from the Heckman (1979) two-step correction procedure. For each context x in the range of contexts X, we estimate

[69] For more details on empirical Bayes Dirichlet-multinomial estimation, see Maritz and Lwin (1989), section 4.5.

Table I.2
2SLS regression results by whether session is at an airport, using empirical Bayes inverse propensity weights.

	Non-airport	Airport
ln(ETA)	-0.0266^{***}	-0.0181
	(0.0018)	(0.0159)
$\ln(1+\mathrm{PT})$	-0.3673^{***}	-0.6019^{*}
	(0.0132)	(0.3244)
ETA Elasticity	-0.0431^{***}	-0.0301
	(0.0029)	(0.0264)
PT Elasticity	-0.5956^{***}	-1.0021^{*}
	(0.0214)	(0.5449)
VOT	18.85^{***}	18.07
	(1.35)	(21.06)
Control Avg. ETA	3.09	3.08
Control Avg. Price	13.39	30.86
Control Req. Rate	0.620	0.606
Controls	x	x
N	5048270	129088
R^2	0.074	0.079

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table I.3
2SLS regression results by day of week, using empirical Bayes inverse propensity weights.

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
ln(ETA)	-0.0273^{***}	-0.0291^{***}	-0.0305^{***}	-0.0259^{***}	-0.0255^{***}	-0.0248^{***}	-0.0228^{***}
	(0.0032)	(0.0032)	(0.0032)	(0.0031)	(0.0029)	(0.0028)	(0.0029)
$\ln(1+\mathrm{PT})$	-0.3762^{***}	-0.4145^{***}	-0.4155^{***}	-0.3555^{***}	-0.3050^{***}	-0.3844^{***}	-0.3878^{***}
	(0.0283)	(0.0318)	(0.0314)	(0.0221)	(0.0150)	(0.0204)	(0.0189)
ETA Elasticity	-0.0458^{***}	-0.0474^{***}	-0.0493^{***}	-0.0428^{***}	-0.0409^{***}	-0.0393^{***}	-0.0381^{***}
	(0.0054)	(0.0052)	(0.0052)	(0.0050)	(0.0046)	(0.0044)	(0.0048)
PT Elasticity	-0.6305^{***}	-0.6742^{***}	-0.6720^{***}	-0.5872^{***}	-0.4898^{***}	-0.6095^{***}	-0.6475^{***}
	(0.0473)	(0.0515)	(0.0505)	(0.0364)	(0.0241)	(0.0324)	(0.0316)
VOT	19.25^{***}	18.76^{***}	20.46^{***}	20.13^{***}	21.38^{***}	18.12^{***}	15.90^{***}
	(2.54)	(2.38)	(2.48)	(2.50)	(2.44)	(2.09)	(2.07)
Control Avg. ETA	3.12	3.03	2.94	3.07	3.38	2.87	3.11
Control Avg. Price	13.79	13.49	13.67	14.13	14.40	13.45	14.00
Control Req. Rate	0.606	0.627	0.630	0.610	0.625	0.634	0.604
Controls	x	x	x	x	x	x	x
N	621203	632247	623344	722306	913007	902130	763121
R^2	0.078	0.073	0.073	0.069	0.064	0.076	0.086

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\mathrm{ETA})$ and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

a probit regression

$$\mathbb{P}\{\text{session } ij \text{ is in context } x | W_{ij}\} = \Phi(W'_{ij}\beta_x) \tag{32}$$

where the vector W_{ij} consists of a binary indicator of whether user i has any pre-experiment sessions; the number of user i's pre-experiment sessions; and the fraction of user i's pre-experiment sessions which occurred in context j (set to 0 for users with no pre-experiment sessions). In the 2SLS regression on the subsample of sessions in context x, we then include the inverse Mills ratio (IMR) based on the probit (32) as a regressor in the first and second stages.

The results of this procedure are shown in Tables I.6 through I.10. In almost all regressions, the estimated coefficient on the IMR is statistically significant, but the estimated elasticities and VOTs are little changed versus the uncorrected results.

Table I.4
2SLS regression results by time of week, using empirical Bayes inverse propensity weights.

	Weekdays				
	6–10 AM	10 AM–4 PM	4 PM–7 PM	7 PM–11 PM	11 PM–6 AM
ln(ETA)	-0.0471^{***}	-0.0267^{***}	-0.0270^{***}	-0.0151^{***}	-0.0246^{***}
	(0.0043)	(0.0034)	(0.0033)	(0.0030)	(0.0038)
$\ln(1+\mathrm{PT})$	-0.3907^{***}	-0.3962^{***}	-0.3302^{***}	-0.3449^{***}	-0.3165^{***}
	(0.0265)	(0.0346)	(0.0223)	(0.0274)	(0.0200)
ETA Elasticity	-0.0767^{***}	-0.0470^{***}	-0.0450^{***}	-0.0244^{***}	-0.0405^{***}
	(0.0067)	(0.0059)	(0.0055)	(0.0048)	(0.0062)
PT Elasticity	-0.6361^{***}	-0.6969^{***}	-0.5519^{***}	-0.5558^{***}	-0.5202^{***}
	(0.0420)	(0.0603)	(0.0370)	(0.0439)	(0.0327)
VOT	29.23^{***}	19.62^{***}	20.78^{***}	13.08^{***}	19.93^{***}
	(3.02)	(2.86)	(2.72)	(2.63)	(3.07)
Control Avg. ETA	3.60	2.92	3.18	2.65	3.83
Control Avg. Price	14.53	14.15	13.48	13.14	16.34
Control Req. Rate	0.662	0.587	0.614	0.631	0.619
Controls	x	x	x	x	x
N	545313	859760	659754	888423	558857
R^2	0.066	0.063	0.077	0.081	0.060

	Weekends				
	6–10 AM	10 AM–4 PM	4 PM–7 PM	7 PM–11 PM	11 PM–6 AM
ln(ETA)	-0.0283^{***}	-0.0310^{***}	-0.0323^{***}	-0.0229^{***}	-0.0145^{***}
	(0.0074)	(0.0041)	(0.0045)	(0.0042)	(0.0040)
$\ln(1+\mathrm{PT})$	-0.3670^{***}	-0.4117^{***}	-0.3665^{***}	-0.4302^{***}	-0.3397^{***}
	(0.0489)	(0.0289)	(0.0262)	(0.0423)	(0.0229)
ETA Elasticity	-0.0514^{***}	-0.0546^{***}	-0.0542^{***}	-0.0365^{***}	-0.0216^{***}
	(0.0135)	(0.0072)	(0.0076)	(0.0067)	(0.0059)
PT Elasticity	-0.6672^{***}	-0.7258^{***}	-0.6145^{***}	-0.6855^{***}	-0.5076^{***}
	(0.0893)	(0.0509)	(0.0441)	(0.0674)	(0.0339)
VOT	17.27^{***}	19.52^{***}	22.83^{***}	16.01^{***}	13.28^{***}
	(4.82)	(2.75)	(3.33)	(3.11)	(3.63)
Control Avg. ETA	4.01	3.08	3.12	2.58	2.78
Control Avg. Price	14.99	13.33	13.45	12.93	14.47
Control Req. Rate	0.557	0.574	0.600	0.633	0.683
Controls	x	x	x	x	x
N	135429	412620	280795	386487	449920
R^2	0.076	0.082	0.093	0.092	0.063

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\mathrm{ETA})$ and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table I.5
2SLS regression results by distance to nearest public transit, using empirical Bayes inverse propensity weights.

	Under 50m	50 to 200m	200 to 800m	Over 800m
$\ln(\text{ETA})$	-0.0275^{***}	-0.0246^{***}	-0.0285^{***}	-0.0282^{**}
	(0.0032)	(0.0022)	(0.0037)	(0.0140)
$\ln(1 + \text{PT})$	-0.3299^{***}	-0.3445^{***}	-0.4143^{***}	-0.4957^{***}
	(0.0207)	(0.0161)	(0.0319)	(0.1585)
ETA Elasticity	-0.0412^{***}	-0.0391^{***}	-0.0499^{***}	-0.0577^{**}
	(0.0047)	(0.0035)	(0.0067)	(0.0291)
PT Elasticity	-0.4940^{***}	-0.5476^{***}	-0.7261^{***}	-1.0129^{***}
	(0.0306)	(0.0255)	(0.0569)	(0.3290)
VOT	25.03^{***}	20.07^{***}	19.87^{***}	14.42^{*}
	(3.02)	(1.92)	(2.80)	(7.94)
Control Avg. ETA	2.58	2.85	3.43	5.24
Control Avg. Price	12.90	13.35	16.54	22.11
Control Req. Rate	0.686	0.638	0.560	0.487
Controls	x	x	x	x
N	1054335	2554328	1269849	298846
R^2	0.062	0.067	0.061	0.101

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\text{ETA})$ and $\ln(1 + \text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table I.6
2SLS regression results by whether session is downtown, using Heckman's two-step correction.

	Non-downtown	Downtown
ln(ETA)	-0.0317^{***}	-0.0231^{***}
	(0.0028)	(0.0021)
$\ln(1 + \mathrm{PT})$	-0.4025^{***}	-0.3520^{***}
	(0.0303)	(0.0128)
IMR	-0.0285^{***}	-0.0079^{***}
	(0.0014)	(0.0014)
ETA Elasticity	-0.0599^{***}	-0.0344^{***}
	(0.0054)	(0.0031)
PT Elasticity	-0.7612^{***}	-0.5241^{***}
	(0.0574)	(0.0191)
VOT	17.94^{***}	21.65^{***}
	(1.97)	(1.99)
Control Avg. ETA	4.36	2.28
Control Avg. Price	16.57	12.54
Control Req. Rate	0.533	0.674
Controls	x	x
N	1986393	3190965
R^2	0.049	0.059

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1 + \mathrm{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table I.7
2SLS regression results by whether session is at an airport, using Heckman's two-step correction.

	Non-airport	Airport
$\ln(\text{ETA})$	-0.0266^{***}	-0.0135
	(0.0018)	(0.0094)
$\ln(1+\text{PT})$	-0.3673^{***}	-0.3336^{***}
	(0.0132)	(0.0924)
IMR	0.0077	-0.0800^{***}
	(0.0103)	(0.0274)
ETA Elasticity	-0.0431^{***}	-0.0223
	(0.0029)	(0.0155)
PT Elasticity	-0.5951^{***}	-0.5514^{***}
	(0.0214)	(0.1525)
VOT	18.84^{***}	26.15
	(1.35)	(18.24)
Control Avg. ETA	3.09	2.92
Control Avg. Price	13.37	31.40
Control Req. Rate	0.620	0.606
Controls	x	x
N	5048270	129088
R^2	0.074	0.082

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\text{ETA})$ and $\ln(1+\text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table I.8

2SLS regression results by day of week, using Heckman's two-step correction.

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
ln(ETA)	-0.0288^{***}	-0.0285^{***}	-0.0282^{***}	-0.0251^{***}	-0.0245^{***}	-0.0249^{***}	-0.0244^{***}
	(0.0031)	(0.0031)	(0.0031)	(0.0030)	(0.0028)	(0.0026)	(0.0028)
$\ln(1 + \mathrm{PT})$	-0.3800^{***}	-0.4235^{***}	-0.4050^{***}	-0.3559^{***}	-0.2948^{***}	-0.3763^{***}	-0.3834^{***}
	(0.0277)	(0.0305)	(0.0300)	(0.0214)	(0.0145)	(0.0193)	(0.0180)
IMR	-0.2625^{***}	-0.2370^{***}	-0.3020^{***}	-0.0990^{***}	0.3593^{***}	-0.0511^{***}	-0.1208^{***}
	(0.0172)	(0.0151)	(0.0169)	(0.0281)	(0.0262)	(0.0066)	(0.0092)
ETA Elasticity	-0.0478^{***}	-0.0457^{***}	-0.0451^{***}	-0.0413^{***}	-0.0393^{***}	-0.0395^{***}	-0.0407^{***}
	(0.0052)	(0.0050)	(0.0050)	(0.0049)	(0.0045)	(0.0041)	(0.0046)
PT Elasticity	-0.6315^{***}	-0.6806^{***}	-0.6473^{***}	-0.5850^{***}	-0.4728^{***}	-0.5957^{***}	-0.6395^{***}
	(0.0461)	(0.0491)	(0.0480)	(0.0352)	(0.0233)	(0.0306)	(0.0300)
VOT	19.87^{***}	17.82^{***}	19.31^{***}	19.43^{***}	21.23^{***}	18.73^{***}	17.25^{***}
	(2.45)	(2.21)	(2.40)	(2.41)	(2.45)	(2.04)	(2.01)
Control Avg. ETA	3.13	3.03	2.93	3.07	3.38	2.86	3.12
Control Avg. Price	13.70	13.38	13.55	14.08	14.41	13.50	14.07
Control Req. Rate	0.606	0.627	0.630	0.610	0.625	0.634	0.604
Controls	x	x	x	x	x	x	x
N	621203	632247	623344	722306	913007	902130	763121
R^2	0.082	0.078	0.078	0.071	0.065	0.077	0.086

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\mathrm{ETA})$ and $\ln(1 + \mathrm{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table I.9
2SLS regression results by time of week, using Heckman's two-step correction.

	Weekdays				
	6–10 AM	10 AM–4 PM	4 PM–7 PM	7 PM–11 PM	11 PM–6 AM
ln(ETA)	-0.0447^{***}	-0.0266^{***}	-0.0281^{***}	-0.0152^{***}	-0.0246^{***}
	(0.0037)	(0.0032)	(0.0031)	(0.0029)	(0.0035)
$\ln(1+\text{PT})$	-0.3963^{***}	-0.4128^{***}	-0.3171^{***}	-0.3400^{***}	-0.3363^{***}
	(0.0220)	(0.0348)	(0.0206)	(0.0255)	(0.0192)
IMR	-0.1783^{***}	-0.2269^{***}	-0.2383^{***}	-0.1498^{***}	-0.1351^{***}
	(0.0034)	(0.0045)	(0.0072)	(0.0087)	(0.0056)
ETA Elasticity	-0.0680^{***}	-0.0457^{***}	-0.0461^{***}	-0.0241^{***}	-0.0398^{***}
	(0.0056)	(0.0055)	(0.0052)	(0.0045)	(0.0056)
PT Elasticity	-0.6033^{***}	-0.7094^{***}	-0.5198^{***}	-0.5401^{***}	-0.5434^{***}
	(0.0336)	(0.0598)	(0.0338)	(0.0405)	(0.0311)
VOT	26.23^{***}	18.27^{***}	22.42^{***}	13.42^{***}	18.55^{***}
	(2.39)	(2.56)	(2.70)	(2.58)	(2.61)
Control Avg. ETA	3.54	2.95	3.16	2.60	3.81
Control Avg. Price	13.72	13.93	13.33	13.04	16.07
Control Req. Rate	0.662	0.587	0.614	0.631	0.619
Controls	x	x	x	x	x
N	545313	859760	659754	888423	558857
R^2	0.094	0.075	0.086	0.086	0.064
	Weekends				
	6–10 AM	10 AM–4 PM	4 PM–7 PM	7 PM–11 PM	11 PM–6 AM
ln(ETA)	-0.0304^{***}	-0.0304^{***}	-0.0285^{***}	-0.0211^{***}	-0.0163^{***}
	(0.0060)	(0.0037)	(0.0040)	(0.0038)	(0.0034)
$\ln(1+\text{PT})$	-0.3652^{***}	-0.4145^{***}	-0.3660^{***}	-0.4220^{***}	-0.3431^{***}
	(0.0444)	(0.0264)	(0.0234)	(0.0389)	(0.0194)
IMR	-0.2326^{***}	-0.2172^{***}	-0.2204^{***}	-0.1647^{***}	-0.0913^{***}
	(0.0130)	(0.0131)	(0.0174)	(0.0114)	(0.0039)
ETA Elasticity	-0.0551^{***}	-0.0534^{***}	-0.0478^{***}	-0.0335^{***}	-0.0239^{***}
	(0.0109)	(0.0066)	(0.0067)	(0.0061)	(0.0050)
PT Elasticity	-0.6614^{***}	-0.7286^{***}	-0.6150^{***}	-0.6697^{***}	-0.5025^{***}
	(0.0805)	(0.0465)	(0.0393)	(0.0617)	(0.0284)
VOT	17.91^{***}	18.95^{***}	20.18^{***}	15.07^{***}	14.99^{***}
	(3.95)	(2.49)	(2.92)	(2.87)	(3.10)
Control Avg. ETA	4.13	3.09	3.14	2.59	2.76
Control Avg. Price	14.79	13.32	13.59	13.01	14.54
Control Req. Rate	0.557	0.574	0.600	0.633	0.683
Controls	x	x	x	x	x
N	135429	412620	280795	386487	449920
R^2	0.081	0.083	0.092	0.092	0.064

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1+\text{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

Table I.10
2SLS regression results by distance to nearest public transit, using Heckman's two-step correction.

	Under 50m	50 to 200m	200 to 800m	Over 800m
ln(ETA)	-0.0270^{***}	-0.0250^{***}	-0.0282^{***}	-0.0278^{***}
	(0.0028)	(0.0022)	(0.0032)	(0.0066)
$\ln(1+\mathrm{PT})$	-0.3382^{***}	-0.3359^{***}	-0.4456^{***}	-0.5687^{***}
	(0.0174)	(0.0154)	(0.0274)	(0.0777)
IMR	0.0148^{***}	0.0242^{***}	0.0162^{***}	-0.0161^{***}
	(0.0029)	(0.0023)	(0.0021)	(0.0024)
ETA Elasticity	-0.0395^{***}	-0.0393^{***}	-0.0506^{***}	-0.0575^{***}
	(0.0041)	(0.0034)	(0.0057)	(0.0136)
PT Elasticity	-0.4953^{***}	-0.5286^{***}	-0.8015^{***}	-1.1733^{***}
	(0.0256)	(0.0242)	(0.0493)	(0.1603)
VOT	23.91^{***}	20.70^{***}	16.95^{***}	10.76^{***}
	(2.56)	(1.91)	(2.03)	(2.78)
Control Avg. ETA	2.47	2.82	3.60	5.29
Control Avg. Price	12.34	13.07	16.08	19.39
Control Req. Rate	0.686	0.638	0.560	0.487
Controls	x	x	x	x
N	1054335	2554328	1269849	298846
R^2	0.063	0.071	0.058	0.060

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) and $\ln(1+\mathrm{PT})$ instrumented by experimental group indicators. Value of time expressed in 2015 US dollars per hour. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.

J Reweighting Elasticity Estimates

In inspecting our results on heterogeneity of time elasticities, one may note that the control average ETA often varies considerably between subsamples. In light of the evidence from the second experiment suggesting that ETA elasticities vary over the demand curve, this observation calls into question whether our heterogeneity results are due to differences in the demand curve between subsamples, or whether they are due to us estimating elasticities at different points in different samples.

To see the issue formally, consider a simplified version of our main model, in which demand D_i in session i is a function of the (logged) waiting time $T_i \geq 0$, session and user characteristics X_i, and an unobserved shock ε_i:

$$D_i = D_i(T_i, X_i, \varepsilon_i). \tag{33}$$

Suppose further that z_i is a binary instrument for T_i, satisfying the assumptions of independence (z_i is independent of $(T_i(X_i, 1), T_i(X_i, 0))$, conditional on X_i), exclusion ($D_i(T, X, z) = D_i(T, X, z')$ for all z, z'), and strict monotonicity ($T_i(X_i, 1) > T_i(X_i, 0)$). Now, consider restricting the sample to observations i for which $X_i = x$ and running an IV regression of D_i on T_i using z_i as an instrument. The resulting coefficient on T_i will be an estimate of

$$\beta_x = \frac{\mathbb{E}[D_i(T_i(1), X_i, 1) - D_i(T_i(0), X_i, 0)|X_i = x]}{\mathbb{E}[T_i(1) - T_i(0)|X_i = x]} \tag{34}$$

$$= \int_0^\infty \mathbb{E}\left[\frac{\partial D_i}{\partial T}(T)|T_i(0) < T < T_i(1), X_i = x\right] \omega(T|x)\, dT \tag{35}$$

where

$$\omega(T|x) \propto \mathbb{P}[T_i(0) < T < T_i(1)|X_i = x],$$

and $\omega(T|x)$ is normalized to integrate to 1. The proof is the same as that of Theorem 1 in Angrist et al. (2000).

The above result shows that for two values of X, say x and x', separate IV regressions for the subsamples of observations with $X_i = x$ and $X_i = x'$ will estimate quantities that may differ for two reasons:

1. $\mathbb{E}\left[\frac{\partial D_i}{\partial T}(T)|T_i(0) < T < T_i(1), X_i = x\right]$ may differ from $\mathbb{E}\left[\frac{\partial D_i}{\partial T}(T)|T_i(0) < T < T_i(1), X_i = x'\right]$ for some fixed T; or

2. $\omega(\cdot, x)$ may differ from $\omega(\cdot, x')$.

Differences of the first kind are economically interesting, as they indicate that the average slope of the demand curve at some fixed point differs between the two samples. Differences of the second kind are nuisances, as they are caused by our semi-elasticity estimates for the two subsamples being measured at different points.

Note that $\omega(\cdot|x)$ is determined by the distribution $(T_i(0), T_i(1))|X_i = x$. Hence, if $(T_i(0), T_i(1))|X_i = x \sim (T_i(0), T_i(1))|X_i = x'$, we would have $\omega(\cdot|x) = \omega(\cdot|x')$, and any differences between β_x and $\beta_{x'}$ would be of the first kind.

We may attempt to set $(T_i(0), T_i(1))|X_i = x \sim (T_i(0), T_i(1))|X_i = x'$ by reweighting each subsample. If such weights can be found, and can be chosen independently of z_i, we may then run weighted IV regressions on each subsample to estimate semi-elasticities which represent weighted average derivatives over the same sections of the demand curve, with the same weights, and can therefore be more readily compared.[70]

We attempt such reweighting for the second experiment, in which we can observe what a session's ETA quote would have been under each treatment level. To find a reweighting scheme that (approximately) equates the joint distributions of potential ETAs, we employ the maximum entropy rebalancing idea of Hainmueller (2012). This entails solving a smooth, convex, constrained optimization problem to find the maximum entropy weight vector w which sets certain moments in the data equal to pre-specified target moments. For our application, we use moments on the following variables:

- the logged counterfactual ETA for each treatment level, and the squares and cross-products of these values;
- indicators of each time category of the week;
- indicators of precipitation type;
- an indicator of whether the session is downtown;
- an indicator of whether the user is a business user;
- indicators for distance to nearest public transit;
- indicators for each region; and
- log of the price multiplier and its square.

For each variable, the target moment is the moment in the full data sample. Matching moments of logged counterfactual ETA for each treatment, as well as squares and cross-products of these, is intended to approximately balance the conditional joint distributions of counterfactual ETAs in each subsample. We also add various moment conditions on control variables to protect against the possibility that reweighting to balance counterfactual ETA distributions causes the samples to

[70] A similar argument shows that the semi-elasticity estimates from 2SLS regressions with multiple binary instruments can also be compared across subsamples so long as the conditional joint distributions of potential ETAs for each value of the instrument vector are the same between the two subsamples: $(T_i(z_0), \ldots, T_i(z_k))|X_i = x \sim (T_i(z_0), \ldots, T_i(z_k))|X_i = x'$.

become unbalanced on other observables.[71] We focusing on a particular heterogeneity dimension, we drop the associated moment condition (e.g., when comparing time elasticities across regions, we do not reweight by region.)

Table J.1 gives the entropy-balance-reweighted 2SLS results by region. The results are quite similar to those found without reweighting: we still find New York City and Washington, D.C. to have the largest ETA elasticities and Miami to have the smallest ETA elasticity, suggesting that these regional variations in ETA elasticities reflect true differences in the underlying demand curves, and not just differences in prevailing ETA levels. Our observed results comparing ETA elasticities at different times of the week are also robust to reweighting (Tables J.2 and J.3).

Table J.1
Reweighted 2SLS results by region.

	San Francisco	New York City	Chicago	D.C.	Miami	New Jersey	Boston	Philadelphia	Atlanta	Los Angeles
ln(ETA)	-0.0247^{***}	-0.0463^{***}	-0.0249^{***}	-0.0306^{***}	-0.0165^{***}	-0.0182^{***}	-0.0271^{***}	-0.0194^{***}	-0.0101^{***}	-0.0219^{***}
	(0.0021)	(0.0014)	(0.0014)	(0.0015)	(0.0026)	(0.0023)	(0.0021)	(0.0023)	(0.0033)	(0.0015)
$\ln(1+\text{PT})$	-0.2379^{***}	-0.3478^{***}	-0.2826^{***}	-0.2875^{***}	-0.1567^{***}	-0.3065^{***}	-0.3392^{***}	-0.2526^{***}	-0.1809^{***}	-0.2434^{***}
	(0.0047)	(0.0028)	(0.0027)	(0.0023)	(0.0081)	(0.0039)	(0.0026)	(0.0041)	(0.0078)	(0.0041)
ETA Elasticity	-0.0361^{***}	-0.0882^{***}	-0.0406^{***}	-0.0491^{***}	-0.0257^{***}	-0.0319^{***}	-0.0440^{***}	-0.0331^{***}	-0.0162^{***}	-0.0345^{***}
	(0.0030)	(0.0027)	(0.0022)	(0.0025)	(0.0040)	(0.0042)	(0.0035)	(0.0039)	(0.0055)	(0.0023)
Control Avg. ETA	3.37	3.41	3.38	3.39	3.38	3.41	3.38	3.40	3.39	3.40
Control Req. Rate	0.690	0.538	0.621	0.631	0.648	0.576	0.622	0.592	0.625	0.640
Controls	x	x	x	x	x	x	x	x	x	x
N	1325977	1261730	1101478	894309	798597	631453	677806	473281	487427	1842581
R^2	0.076	0.068	0.071	0.055	0.049	0.061	0.068	0.064	0.053	0.060

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local hour of week and week of year, session geohash5, business user, and decile of pre-experiment lifetime rides.

Table J.2
Reweighted 2SLS results by time of day (weekdays).

	6–10 AM	10 AM–4 PM	4 PM–7 PM	7 PM–11 PM	11 PM–6 AM
ln(ETA)	-0.0367^{***}	-0.0240^{***}	-0.0288^{***}	-0.0253^{***}	-0.0210^{***}
	(0.0012)	(0.0010)	(0.0013)	(0.0012)	(0.0016)
$\ln(1+\text{PT})$	-0.3441^{***}	-0.2960^{***}	-0.3154^{***}	-0.2783^{***}	-0.2508^{***}
	(0.0019)	(0.0024)	(0.0023)	(0.0027)	(0.0030)
ETA Elasticity	-0.0558^{***}	-0.0400^{***}	-0.0467^{***}	-0.0422^{***}	-0.0338^{***}
	(0.0019)	(0.0017)	(0.0020)	(0.0019)	(0.0025)
Control Avg. ETA	3.39	3.39	3.38	3.39	3.38
Control Req. Rate	0.668	0.606	0.625	0.606	0.628
Controls	x	x	x	x	x
N	1196076	1795586	1207487	1447026	777861
R^2	0.084	0.074	0.077	0.092	0.077

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local hour of week and week of year, session geohash5, business user, and decile of pre-experiment lifetime rides.

[71] It is, of course, still possible that the reweighting causes the samples to become unbalanced on unobservables.

Table J.3
Reweighted 2SLS results by time of day (weekends).

	6–10 AM	10 AM–4 PM	4 PM–7 PM	7 PM–11 PM	11 PM–6 AM
ln(ETA)	-0.0293^{***}	-0.0263^{***}	-0.0258^{***}	-0.0218^{***}	-0.0198^{***}
	(0.0025)	(0.0015)	(0.0019)	(0.0017)	(0.0018)
$\ln(1+\text{PT})$	-0.3135^{***}	-0.3109^{***}	-0.2980^{***}	-0.2557^{***}	-0.2282^{***}
	(0.0040)	(0.0028)	(0.0035)	(0.0038)	(0.0032)
ETA Elasticity	-0.0493^{***}	-0.0448^{***}	-0.0426^{***}	-0.0358^{***}	-0.0305^{***}
	(0.0044)	(0.0025)	(0.0032)	(0.0028)	(0.0028)
Control Avg. ETA	3.40	3.40	3.39	3.39	3.38
Control Req. Rate	0.602	0.594	0.614	0.616	0.656
Controls	x	x	x	x	x
N	314065	886668	498999	678722	692149
R^2	0.073	0.072	0.077	0.085	0.065

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local hour of week and week of year, session geohash5, business user, and decile of pre-experiment lifetime rides.

Table J.4 gives reweighted results by precipitation type. Here, the difference between elasticities is no longer statistically significant.

Table J.5 gives reweighted results by whether the session is downtown. Here, the result reverses relative to the unweighted results, and we find a significantly larger ETA elasticity for downtown sessions. This suggests that the counterintuitive unweighted result of non-downtown elasticities being larger was driven by ETAs being larger in non-downtown sessions, and not by non-downtown sessions being more responsive to changes in waiting time at the same base ETA level. Similarly, Table J.6 shows that the ETA elasticity is greater when passengers are closer to public transit; this result runs counter to the unweighted estimates but is consistent with economic intuition.

The validity of this method depends on the quality of the counterfactual data, and the assumption that treatment does not affect counterfactual outcomes. For some observations, the no or only incomplete counterfactual data is available; for other sessions, the counterfactual data has apparent "errors": the reported "counterfactual" ETA for the treatment a session actually received differs from the actual ETA. Missing and incorrect counterfactuals affect about 2% of sessions, and are attributable to small technical errors in logging. In Table J.8, we show that whether counterfactual data is missing and the degree of the counterfactual are not significantly predicted by treatment status.

In Table J.9, we regress the reported counterfactuals for each treatment group on indicator variables of treatment recieved. Here we find statistically significant (though economically small) evidence that sessions which had their ETAs increased more show lower counterfactual ETAs

Table J.4
Reweighted 2SLS results by precipitation type.

	None	Rain
ln(ETA)	-0.0268^{***}	-0.0244^{***}
	(0.0005)	(0.0016)
$\ln(1+\mathrm{PT})$	-0.2852^{***}	-0.2909^{***}
	(0.0015)	(0.0028)
ETA Elasticity	-0.0435^{***}	-0.0396^{***}
	(0.0008)	(0.0026)
Control Avg. ETA	3.39	3.38
Control Req. Rate	0.624	0.625
Controls	x	x
N	7467167	2027472
R^2	0.074	0.077

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local hour of week and week of year, session geohash5, business user, and decile of pre-experiment lifetime rides.

Table J.5
Reweighted 2SLS results by whether session downtown.

	Non-downtown	Downtown
ln(ETA)	-0.0171^{***}	-0.0377^{***}
	(0.0011)	(0.0011)
$\ln(1+\mathrm{PT})$	-0.2658^{***}	-0.2899^{***}
	(0.0033)	(0.0018)
ETA Elasticity	-0.0288^{***}	-0.0587^{***}
	(0.0019)	(0.0017)
Control Avg. ETA	3.39	3.41
Control Req. Rate	0.598	0.653
Controls	x	x
N	4262346	4918765
R^2	0.065	0.067

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local hour of week and week of year, session geohash5, business user, and decile of pre-experiment lifetime rides.

Table J.6
Reweighted 2SLS results by distance to public transit.

	Under 50m	50 to 200m	200 to 800m	Over 800m
ln(ETA)	-0.0340^{***}	-0.0261^{***}	-0.0221^{***}	-0.0014
	(0.0011)	(0.0006)	(0.0011)	(0.0116)
$\ln(1+\text{PT})$	-0.2902^{***}	-0.2880^{***}	-0.2864^{***}	-0.2467^{***}
	(0.0018)	(0.0011)	(0.0026)	(0.0240)
ETA Elasticity	-0.0532^{***}	-0.0420^{***}	-0.0370^{***}	-0.0025
	(0.0017)	(0.0010)	(0.0019)	(0.0215)
Control Avg. ETA	3.40	3.40	3.39	3.34
Control Req. Rate	0.648	0.629	0.605	0.554
Controls	x	x	x	x
N	1958132	4764641	2244268	527598
R^2	0.069	0.072	0.076	0.095

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local hour of week and week of year, session geohash5, business user, and decile of pre-experiment lifetime rides.

Table J.7
Reweighted 2SLS results by whether business user.

	Non-business	Business
ln(ETA)	-0.0258^{***}	-0.0339^{***}
	(0.0005)	(0.0019)
$\ln(1+\text{PT})$	-0.2882^{***}	-0.2787^{***}
	(0.0011)	(0.0036)
ETA Elasticity	-0.0420^{***}	-0.0526^{***}
	(0.0007)	(0.0028)
Control Avg. ETA	3.39	3.39
Control Req. Rate	0.622	0.655
Controls	x	x
N	8839169	655470
R^2	0.073	0.087

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. ln(ETA) instrumented by experimental treatment indicators. Controls include local hour of week and week of year, session geohash5, business user, and decile of pre-experiment lifetime rides.

Table J.8
Regressions of errors in counterfactual (CF) data on treatment indicators for experiment 2.

	CF Data Missing	CF Error	Squared CF Error
Plus 60+	0.0000	−0.0000	−0.0013
	(0.0001)	(0.0002)	(0.0012)
Plus 150+	−0.0001	−0.0001	−0.0006
	(0.0001)	(0.0002)	(0.0014)
Plus 240+	0.0002	−0.0003	0.0039^{**}
	(0.0001)	(0.0003)	(0.0017)
Controls	x	x	x
N	9668820	9494639	9494639
R^2	0.168	0.002	0.007
F	0.9	0.5	2.8^{**}

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\text{ETA})$ instrumented by experimental treatment indicators. Controls include local hour of week and week of year, session geohash5, business user, and decile of pre-experiment lifetime rides.

then control sessions. Since we know that the experimental treatments were assigned randomly, this effect may be caused by the following mechanism: in location-times subject to higher ETA treatments, fewer requests are made, resulting in more available drivers nearby and a lower reported counterfactual ETA. If this mechanism is occuring, the "counterfactuals" we observe are not truly representative of what ETAs would have been experienced under different treatment regimes.

This mechanism constitutes an interaction effect, but the estimates in Table J.9 suggest that such effects are small in magnitude, shifting ETAs by about four seconds at most.

Table J.9
Regressions of counterfactual ETAs (in minutes) for each treatment level on treatment indicators.

	CF Control	CF Plus 60	CF Plus 150	CF Plus 240
Plus 60+	-0.0303^{***}	-0.0295^{***}	-0.0332^{***}	-0.0357^{***}
	(0.0016)	(0.0018)	(0.0018)	(0.0018)
Plus 150+	-0.0379^{***}	-0.0378^{***}	-0.0430^{***}	-0.0472^{***}
	(0.0019)	(0.0022)	(0.0021)	(0.0021)
Plus 240+	-0.0499^{***}	-0.0523^{***}	-0.0539^{***}	-0.0649^{***}
	(0.0024)	(0.0030)	(0.0028)	(0.0028)
Controls	x	x	x	x
N	9494639	9494639	9494639	9494639
R^2	0.466	0.567	0.557	0.543
F	272.9^{***}	198.8^{***}	255.1^{***}	329.3^{***}

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. $\ln(\text{ETA})$ instrumented by experimental treatment indicators. Controls include local hour of week and week of year, session geohash5, business user, and decile of pre-experiment lifetime rides.
The dependent variable for each column is the counterfactual ETA reported for a particular treatment level, while the independent variables are indicators of which treatment a session actually received.

K Ride Types

As noted in the text, Lyft offers multiple ride types, while our analysis is not ride-type-specific. In particular, our main proxy for demand—the indicator of whether a session had a ride request—does not differentiate between ride types, and the last-in-session ETA used in our regressions is not for a fixed ride type, but for whatever ride type the passenger selected last.

We consider in this section whether the presence of ride type effects may bias our results. We first note that the experimental variation does induce some small but significant effect on a passenger's consideration and choice of ride type. In Table K.1, we regress indicators of whether a session considered the Shared at any point on the experimental treatment indicators. We find small but statistically significant effects on whether a session considers the Shared mode; the individual point estimates suggest that high price treatments increase the probability of looking at Shared, low price treatments decrease the probability, and high ETA treatments leave it unchanged. An F test of the hypotheses that the coefficient on High ETA Normal Price is 0, that the coefficients on High ETA High Price and Normal ETA High Price are equal, and High ETA Low Price and Normal ETA Low price returns a p-value of 0.23.

Since the ETA variation does not significantly affect which ride types are considered, ride type effects should not bias our estimate of the effect of ETA. Such effects may, however, bias our price elasticity estimates: the base price of a ride is a component of the error term in our main model, and the price instruments may be correlated with this component through the effect on ride type. This correlation would invalidate the exogeneity of our price instruments, resulting in

Table K.1
Effect of treatments on ride types considered.

	Considered Shared	Considered XL
High ETA High Price	0.009***	0.001
	(0.003)	(0.001)
High ETA Normal Price	0.004**	0.000
	(0.002)	(0.000)
High ETA Low Price	−0.006***	0.000
	(0.002)	(0.000)
Normal ETA High Price	0.010***	0.000
	(0.002)	(0.001)
Normal ETA Low Price	−0.006***	−0.000
	(0.002)	(0.000)
Controls	x	x
N	5177358	5177358
R^2	0.293	0.012
F	13***	1

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Clustered standard errors in parentheses. F test is of the null that all the coefficients on the instruments equal 0. Controls include local week of year, local hour of week, user geohash5, business user, airport, and decile of pre-experiment lifetime rides.
The dependent variables for the two columns are indicators of whether a session ever selected the Shared or XL ride types, respectively.

an underestimate of the price elasticity and an overestimate of the VOT.

We argue, however, that such bias is likely to be small. For rides completed by control passengers in our sample, the average Standard price is $\$14.65$, the average XL price is 11.06, and the average XL price is 28.22. The three types account for 70.7%, 27.9%, and 1.3% of rides, respectively. From the results in Table K.1, we see that the Normal ETA High Price treatment increase the probability of considering the Shared type by 1.0 percentage points. As not all passengers induced to look at the Shared mode necessarily switch to considering only that mode, this implies at most that Normal ETA High Price passengers have their probability of considering Standard lowered by 1.0 percentage points and their probability of considering Shared raised by 1.0 percentage points. This shift implies that the average price faced by such passengers is lowered from 13.83 (the average price paid by control passengers) to, at worst,

$$(0.707 - 0.01) \times 14.65 + (0.279 + 0.01) \times 11.06 + 0.013 \times 28.22 = 13.77,$$

a 0.43% decrease, due solely to the effect on ride choice. For comparison, the direct effect of the Normal ETA High Price treatment on the price, through the Prime Time multiplier, is around 5%, around twenty times larger.

L Distribution of VOT

Rather than evaluating our VOT estimate at the full-sample average ETAs and prices, we can plug in individual ETA and price estimates for each session to obtain a full distribution of VOTs. By also including interaction terms in our main model, we may also plug in individual ETA and price semi-elasticities of demand.

We estimate the following equation with 2SLS:

$$Request = \beta_0 + \beta_1 \ln(\mathrm{ETA}) + \ln(\mathrm{ETA}) \times D' B_1 + \beta_2 \ln(1 + \mathrm{PT}) + \ln(1 + \mathrm{PT}) \times D' B_2 + (controls) + \varepsilon,$$

where D is a vector of dummy variables indicating dimensions. We include region, time category, precipitation type, bin of distance to nearest public transit, and binary indicators for business users, airports, and downtown. Our first-stage equations interact the experimental treatment indicators with these same dummy variables:

$$\ln(\mathrm{ETA}) = \gamma_0 + \sum_{j=1}^{5} [\gamma_j (\mathrm{T}j) + (\mathrm{T}j) D' \Gamma_j] + (controls) + u$$
$$\ln(1 + \mathrm{PT}) = \gamma_0 + \sum_{j=1}^{5} [\delta_j (\mathrm{T}j) + (\mathrm{T}j) D' \Delta_j] + (controls) + v.$$

For each session, the relevant price and time semi-elasticities of demand are given from the estimates of the above equation; these will vary between sessions according to their different values of D.

To obtain estimates of ETA and $(1 + PT)$ for each session, we first apply the exponential function to the fitted values of the first-stage equation. These will generally be biased estimates. We correct for this bias by multiplying by the correction factor of Goldberger (1968), which may be written as

$$F = {}_0F_1(\mathrm{df}_{\mathrm{resid}}/2, (\mathrm{df}_{\mathrm{resid}}/2)(\hat{\sigma}^2/2 - \widehat{Var}(\widehat{\gamma}_0))).$$

Here $\mathrm{df}_{\mathrm{resid}}$ is the residual degrees of freedom in the logged first-stage equation; $\hat{\sigma}^2$ is the usual estimator of the residual variance; $\widehat{Var}(\widehat{\gamma}_0)$ is the standard estimator of the variance of the estimate of the intercept; and ${}_0F_1$ is the confluent hypergeometric function

$${}_0F_1(v, z) = \sum_{k=0}^{\infty} \frac{z^k}{(v)_k k!}$$

where $(v)_k$ is the Pochammer symbol $(v)_k = \frac{\Gamma(v+n)}{\Gamma(v)}$ and Γ is Euler's gamma function. Our fitted

ETA and (1 + PT) values are then

$$\widehat{ETA} = \exp(\ln(\widehat{\mathrm{ETA}}))F_{\ln(\mathrm{ETA})}$$
$$\widehat{1+PT} = \exp(\ln(\widehat{1+\mathrm{PT}}))F_{\ln(1+\mathrm{PT})},$$

where $\ln(\widehat{\mathrm{ETA}})$ and $\ln(\widehat{1+\mathrm{PT}})$ are the first-stage fitted values and $F_{\ln(\mathrm{ETA})}$ and $F_{\ln(1+\mathrm{PT})}$ are the Goldberger correction factors corresponding to each equation.

The estimated price multiplier is then converted to an estimated price by multiplying it by a prediction of the base price (that is, $Price/(1+PT)$) derived from a linear model using data on completed control rides. We then compute a VOT for each session as the ratio of its estimated time and price semi-elasticities, multiplied by the ratio of its estimated price and ETA.

Figure L.1 shows the resulting distribution of VOT estimates. The distribution has a mean of \$22.13/hour and a median of \$20.76/hour, both of which are close to our full-sample estimate of \$19.39/hour. The distribution has a standard deviation of \$10.19 and an inter-quartile range is \$13.47, indicative of the large variability in VOTs between different individuals and contexts.

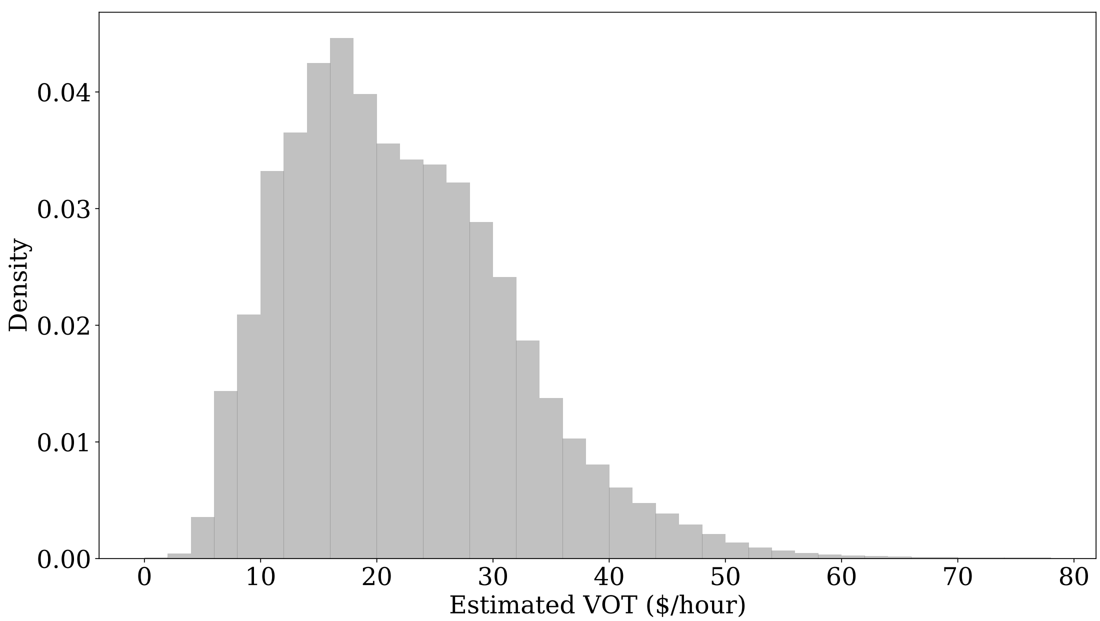


Figure L.1
Distribution of VOT estimates.

M Causality with Interaction Effects

In this section, we argue that our IV estimates maintain an Angrist et al. (2000)-type interpretation as causal weighted average derivatives, despite the possibility of cross-session interaction due to marketplace effects.

Let $\mathbf{Z}$ be the N-dimensional random vector giving the treatment assignment of every session in the experiment, with distribution function $F_{\mathbf{Z}}$. Let Z_i denote the ith entry of $\mathbf{Z}$ and $\mathbf{Z}_{-i}$ the vector $\mathbf{Z}$ with the ith entry deleted. Let $T_i(\mathbf{Z}) = T_i(Z_i, \mathbf{Z}_{-i})$ be the waiting time of the ith session; in general, this quantity depends on all of $\mathbf{Z}$ due to interaction effects. Also let $D_i(T, \mathbf{Z})$ be the ith session's demand (conversion rate) when the waiting time is T and the treatment vector is $\mathbf{Z}$.

The following properties hold:

1. *Independence*: $\mathbf{Z}$ is independent of $T_i(\mathbf{z})$ and $D_i(\mathbf{z}) := D_i(T_i(\mathbf{z}), \mathbf{z})$ for all i and $\mathbf{z}$. This property follows because treatment assignments are fully randomized in the experiment.

   A consequence is that Z_i is independent of $T_i(\mathbf{z}) = T_i(z_i, \mathbf{z}_{-i})$ and $D_i(\mathbf{z}) = D_i(z_i, \mathbf{z}_{-i})$ for all i and $\mathbf{z}$. It follows that $T(z) := E[T_i(z, \mathbf{Z}_{-i})] = E[T_i(z, \mathbf{Z}_{-i})|Z_i = z] = E[T_i|Z_i = z]$ and similarly $D(z) := E[D_i(z, \mathbf{Z}_{-i})] = E[D_i|Z_i = 1]$.

   This result says that we can compute the average waiting time and demand for a session assigned treatment z by taking averages over all sessions assigned z. The only difference in our case vs. the no-interaction case is that the average waiting time and demand for a session assigned treatment z is an average not only over the population of sessions, but also over the possible treatment realizations for all other sessions. Hence, this average may in general depend on the distribution of $\mathbf{Z}$; for example, a distribution that assigns more sessions to a high ETA treatment $z = 1$ would result in a larger decrease in overall market demand, leaving more supply available and pushing down the average control session ETA $T(0)$. To emphasize this dependence on the distribution of $\mathbf{Z}$, we might more accurately write the average waiting time and demand for a session assigned treatment z by $T(z, F_{\mathbf{Z}})$ and $D(z, F_{\mathbf{Z}})$, but suppress this detail in our notation for simplicity.

2. *Exclusion*: $D_i(T, \mathbf{z}) = D_i(T, \mathbf{z}')$ for all i, T, $\mathbf{z}$, and $\mathbf{z}'$. This property follows from the design of our experiment: the ETA is the only aspect of the rider's decision to request that is affected by the experiment, and so holding the ETA fixed, the vector of treatment assignments for all sessions does not affect the rider's decision.

3. *Nonzero effect of treatment on ETA*: $T(z)$ is a nontrivial function of z. This property is validated empirically by noting that sessions with different ETA treatments have different average ETAs.

4. *Monotonicity*: For all z and z', either $\Pr\{T(z) \leq T(z')\} = 1$ or $\Pr\{T(z') \leq T(z)\} = 1$. If 1 denotes the high ETA treatment, we have $\Pr\{T(0) \leq T(1)\} = 1$ by design of the experi-

ment: since the high ETA treatment removes nearby drivers from being eligible to pick up a session, it cannot decrease a session's ETA.

These properties are exactly Assumptions 1–4 of Angrist et al. (2000), from which we conclude that the IV estimand from a regression of D_i on T_i using Z_i as an instrument,

$$\frac{E[D_i|Z_i = 1] - E[D_i|Z_i = 0]}{E[T_i|Z_i = 1] - E[T_i|Z_i = 0]},$$

can be written as

$$\int_0^\infty E\left[\frac{\partial D_i}{\partial T}(T)\middle| T_i(0) \leq T \leq T_i(1)\right] \omega(T)\, dT \tag{36}$$

with

$$\omega(T) = \frac{\Pr\{T_i(0) \leq T \leq T_i(1)\}}{\int_0^\infty \Pr\{T_i(0) \leq r \leq T_i(1)\}\, dr}. \tag{37}$$

This result illustrates that, even without strict SUTVA, our IV estimates can be interpreted as a weighted average of LATE derivatives of demand with respect to ETA. Since $T(0)$ and $T(1)$ depend on $F_{\mathbf{Z}}$, the presence of interaction effects modify the notion of "local" and the weighting of the derivatives at different points. More concretely, we conclude that under any distribution $F_{\mathbf{Z}}$ of the treatment assignments satisfying the properties 1–4 above, our IV estimates would have causal interpretations as weighted average derivatives, but the particular choice of $F_{\mathbf{Z}}$ would affect where on the demand curve these derivatives are evaluated, and how different derivatives at different points on the demand curve are weighted in the final estimate.

References

Angrist, J. D., K. Graddy, and G. W. Imbens (2000). The interpretation of instrumental variables estimators in simultaneous equations models with an application to the demand for fish. *The Review of Economic Studies 67*(3), 499–527.

Federal Highway Administration (2017). National Household Travel Survey.

Goldberger, A. S. (1968). The interpretation and estimation of Cobb-Douglas functions. *Econometrica 36*(3/4), 464–472.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis 20*(1), 25–46.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica 47*(1), 153–161.

Maritz, J. S. and T. Lwin (1989). *Empirical bayes methods*. Routledge.

Mosimann, J. E. (1962). On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions. *Biometrika 49*(1/2), 65–82.

U.S. Bureau of Labor Statistics (2016). May 2016 metropolitan and nonmetropolitan area occupation employment and wages estimates.

U.S. Bureau of Labor Statistics (2020). Consumer price index for all urban consumers: All items in u.s. city average [cpiaucsl].