Flexible Pay and Labor Supply: Evidence from Uber's Instant Pay

M. Keith Chen, Katherine Feinerman, and Kareem Haggag*

March 14, 2024

Abstract

Modern tech platforms provide workers real-time control over when they work, and increasingly, flexible pay: the option to be paid immediately after work. We investigate the labor supply effects of pay flexibility and the implications of present-biased preferences among gig-economy workers. Using granular data from a nationwide randomized controlled trial at Uber, we estimate the effects of switching from a fixed weekly pay schedule to *Instant Pay*, a system that allows on-demand, within-day withdrawals. We find that flexible pay substantially increased both drivers' work time and earnings (ITT: 2%; TOT: 18-37%). Furthermore, the response is significantly higher when drivers are further away from the end of their counterfactual weekly pay cycle, aligning with predictions of hyperbolic discounting models. We discuss welfare and broader implications in contexts in which workers have the ability to flexibly supply labor.

*Chen: Anderson School of Management, University of California at Los Angeles, keith.chen@anderson.ucla.edu, **Feinerman**: Anderson School of Management, University of California at Los Angeles, katherine.feinerman@anderson.ucla.edu, **Haggag**: Anderson School of Management, University of California at Los Angeles, kareem.haggag@anderson.ucla.edu. We thank Linda Babcock, Hal Hershfield, Sam Hirshman, Devin Pope, and seminar participants at UCLA and the Society for Judgment and Decision Making conference for helpful comments and suggestions.

Before COVID-19, roughly 30 percent of the US labor force reported some flexibility over when their workdays started and stopped, a figure that has accelerated with the growth of remote work (Mas and Pallais, 2020). Gig work platforms extend this control further, allowing workers real-time control over their schedules. While this flexibility has obvious benefits for workers, it also introduces the potential for labor supply decisions to reflect many of the same biases and intertemporal patterns commonly found in consumption and savings decisions. That is, when workers can choose their labor supply, the typical structure of upfront costs (effort) and delayed benefits (wages) may result in procrastination and the under-supply of work. As a result, variation in the timing of pay over relatively short horizons may have surprisingly large effects on labor supply.

We use a large-scale natural field experiment at Uber to study the labor supply effects of a key component of the human resources toolkit: the timing/flexibility of pay. Prior to the experiment, Uber – like roughly a quarter of employers in the US (BLS, 2023) – paid workers on a weekly cycle.[1] Across most markets, drivers could take any number of rides from Monday 4am to the following Monday at 4am and would be paid their accumulated earnings on the subsequent Wednesday. Thus, rides driven on Sunday would be paid roughly 3 days later, while those on Monday (after 4am) would be paid roughly 9 days later. In March 2016, Uber began to experimentally roll out *Instant Pay*, a feature that allowed drivers to access their accumulated earnings on-demand. This program thus allowed drivers in the treatment group to reduce the lag between effort and pay by between 3 to 9 days. While standard, time-consistent preferences suggest this shift should have little-to-no effect on behavior (Parsons and Van Wesep, 2013), present bias suggests otherwise.

We find that the *offer* of Instant Pay (intent-to-treat) increases daily labor supply measures by 2.4% (work minutes), 2.4% (earnings in dollars), and 1.5% (number of distinct driving sessions), all statistically significant at the 1% level. This translates to a treatment-on-the-treated (TOT) increase in work time by 17.6% or 36.9%, earnings by 17.6% or 37.2%, and number of sessions by 11.2% or 22.9% depending on the measure of take-up utilized.[2] The labor supply response significantly increases with the counterfactual pay delay (i.e., the response is largest on Monday and smallest on Sunday), aligning with predictions of hyperbolic discounting models. By contrast, distance from the usual payday (Wednesday) has less predictive power, suggesting that the labor supply response is not just driven by

[1]This is based on February 2023 estimates from the BLS Current Employment Statistics which surveys roughly 122,000 business and government agencies (a third of all nonfarm payroll employees in the US). Among this sample, 27% have a weekly pay period, 43% biweekly, 19.8% semimonthly, and 10.3% monthly.

[2]In terms of levels, these figures are 24.6 or 52.3 minutes, \$9.09 or \$19.31, and 0.07 or 0.15 sessions.

sharply-binding liquidity constraints. We also find relatively consistent responses across various heterogeneity cuts, including the median income of the driver's home Census Block Group. Finally, benchmarking against past experimental estimates of Uber drivers' wage elasticities (Chen et al., 2019), the TOT of Instant Pay is roughly equivalent to the labor supply response to a 19% increase in wages.

Broadly, our paper contributes to the literature documenting evidence consistent with present bias in the field. Present bias captures the idea that individuals prefer a larger, delayed reward to a smaller, immediate reward when that choice is in the future, but reverse this preference when the choice is in the present; a phenomenon reflected in hyperbolic and quasi-hyperbolic discounting models.[3] Researchers have implicated present bias to explain empirical patterns observed in contexts ranging across consumption (Shapiro, 2005), household finance (Angeletos et al., 2001; Meier and Sprenger, 2010; Kuchler and Pagel, 2021; Goda et al., 2020), health decisions (DellaVigna and Malmendier, 2006; Fang and Wang, 2015), among others (DellaVigna, 2009).[4] We extend this literature by considering the implications of present bias for a key part of job contract design (pay timing/flexibility).

More specifically, our work relates to a smaller literature on present bias in personnel and labor economics (Cadena and Keys, 2022).[5] For example, DellaVigna and Paserman (2005) find that survey-based measures of short-run impatience predict measures of job search and unemployment duration. In perhaps the closest work to our own, Kaur et al. (2015) run a field

[3]See Ericson and Laibson (2019) for a broad overview of models described under an umbrella of "present focus" intended to remove the prejudgment that the behavior is a mistake.

[4]See also research on demand for commitment, an implication of *sophisticated* hyperbolic discounting (Ashraf et al., 2006; Beshears et al., 2020; John, 2020). Most closely related to our own work is a pair of studies on demand for commitment in the form of less frequent pay. Casaburi and Macchiavello (2019) find in multiple experiments in the Kenyan dairy market that farmers are willing to pay sizable costs to receive less frequent pay as a commitment device. Similarly, Brune et al. (2021) show sizeable demand for deferred payments (into a soft-commitment savings account) in a field experiment among agricultural workers in Malawi, and that such deferred payments increase labor supply (as they serve as a savings device with a higher effective interest rate due to the lower transaction costs, temptation, or kin taxes relative to other available savings options). Opting into Uber's Instant Pay can be thought of as unwinding a commitment device, and so it's not obvious from this literature whether this should increase or decrease labor supply. However, unlike Brune et al. (2021), the "commitment" horizon implied by the usual pay schedule is just one week (as opposed to three months), workers have existing bank accounts (whereas the Malawi farmer sample is almost fully unbanked), and the contexts are quite different. On the latter point, it's worth noting that Berk et al. (2023) finds very little take-up (0.7%) by workers for an employer-based payroll deduction savings account in the US.

[5]Relatedly, a few experiments use real-effort tasks to isolate consumption trade-offs (e.g., Augenblick et al. (2015); Augenblick and Rabin (2019)), addressing the critique that many lab studies study the receipt of monetary payments rather than the utility flows underlying models of intertemporal choice. These real-effort studies typically find more evidence for present bias than choices involving monetary trade-offs (Imai et al., 2021).

experiment with data entry workers and find that they demonstrate strong payday effects (i.e., output increases as the weekly payday grows closer) and demand dominated contracts to solve the self-control problem in effort provision – both results consistent with present bias. Our work extends on these by studying an arguably more natural way of addressing the self-control problem – rather than introducing new upfront (conditional) costs to align short-term and long-term preferences, pay flexibility instead brings delayed benefits into the present. This is especially important in light of evidence that workers demand flexibility in their jobs, whether that be over their schedules (Chen et al., 2019) or remote/in-office modality (Hansen et al., 2023). Frakes and Wasserman (2020) show that US patent examiners procrastinate on their tasks (which may account for up to 1/6th of the patent backlog), and that this procrastination increases with remote work. As remote work becomes more prevalent – Aksoy et al. (2022) find that job listings in the US allowing remote work increased three-fold from 2019 to 2023 – and the external controls of the workplace are relaxed, the role of present bias in labor supply may become more pronounced. Pay flexibility may thus play an increasingly important role in labor supply decisions beyond the gig economy.

Finally, our results relate to a long line of empirical studies of consumption responses to expected and unexpected household income shocks. Mostly aimed at testing the life-cycle permanent income hypothesis, this line starts with Hall (1978) and is summarized in Jappelli and Pistaferri (2010). Recent research using high-quality account-level data on consumption and borrowing finds surprisingly large ("excess") sensitivity of spending to predictable income flows such as the receipt of a welfare transfer (e.g., Gelman et al. (2014); Baker (2018); Kueng (2018); Zhang (2022); Gelman (2022)). Olafsson and Pagel (2018), for example, show excess spending responses to the receipt of a paycheck across the entire income distribution, even among the most liquid consumers. Much of this work argues that the patterns are best explained by time-inconsistent preferences. More recently, researchers have used financial aggregators to examine the relationship between pay frequency and consumption, across both marketing (De La Rosa and Tully, 2022) and finance (Baugh and Correia, 2022; Baugh and Wang, 2022), finding mixed evidence.[6] We extend this by looking at labor supply rather than

[6] Baugh and Wang (2022) find a higher likelihood of financial shortfalls (reflected in bank overdrafts, bounced checks, and online payday lending usage) among social security benefit recipients with longer pay cycles (5 vs. 4 weeks) but similar due dates for recurring bills – they argue these results are best explained by reliance on simplistic budgeting heuristics. By contrast, Baugh and Correia (2022) examine cross-sectional variation in pay frequency in an account aggregator – they find that those paid at a higher frequency have more financial shortfalls, despite lowering credit card borrowing and consumption. They argue this result can be explained by extending a model of optimal pay timing for present-biased workers (Parsons and Van Wesep (2013)) with a costless borrowing option and an illiquid savings vehicle – theoretically, the higher liquidity of frequent pay should increase the likelihood of investing in the illiquid savings vehicle which would help

spending/borrowing and by looking at pay *flexibility* rather than pay *frequency*. That is, we study a policy in which workers can opt into on-demand pay timing rather than the effects of different lengths of fixed schedules set by the employer. As more companies begin to offer pay flexibility, the labor supply implications will also need to be considered.[7]

1 Background and Experimental Design

1.1 Study Setting

This study took place in partnership with Uber, a popular ride-sharing platform. Uber drivers use their own or leased vehicles to provide rides to customers at their discretion. The platform imposes minimal restrictions on working hours, allowing for considerable flexibility in labor supply and scheduling.[8] The majority of Uber drivers work part-time – for example, among drivers studied in Chen et al. (2019), the majority work fewer than 12 hours per week and a substantial fraction who are active in one week are not in the next week. As noted by Hall and Krueger (2018), for many drivers, driving is complementary to other activities such as caregiving, employment, and school attendance. Over an 8-month period close to our own study period, Chen et al. (2023) notes that over one million individuals worked on UberX, Uber's peer-to-peer service in the United States that is our focus.

1.2 Instant Pay

Prior to the experiment, Uber paid drivers on a weekly cadence, with accumulated earnings from Monday (4am) to Monday (4am) deposited on the following Wednesday. Drivers were

explain the other three empirical results. Finally, De La Rosa and Tully (2022) find that higher pay frequency is correlated with increased spending among clients of a financial services provider, and in lab studies, find evidence that higher pay frequency increases subjective wealth perceptions (but that such effects depend on the timing of expenses).

[7] As further discussed in the conclusion, several fintech companies have emerged in recent years that facilitate pay flexibility at companies such as Walmart, including DailyPay, Even, EarnIn, PayActiv, among others.

[8] As described in Chen et al. (2019), as ride requests emerge, Uber's algorithm assigns these to drivers based on proximity. At the time of our study, riders were charged a base fare, along with additional costs determined by the distance and duration of the trip. Fares were standardized within each city and subject to dynamic pricing during periods of high demand relative to driver availability in specific areas. Both drivers and riders were informed of any applicable surge multiplier ahead of the trip, with the total fare and corresponding driver earnings adjusting in tandem during these peak times. Excluding various taxes, fees, and promotional adjustments, drivers received a portion of the total fare less Uber's service fee (roughly 20 to 30% during the study period of 2016). Thus, a driver's income is influenced by their choices regarding work hours and location, as well as the fluctuating demand and availability of other drivers and riders.

required to link a checking account to their Uber profile and would receive their payment via a direct deposit to this account. Starting in March 2016, Uber launched a new feature called "Instant Pay" to a randomized selection of drivers across several markets (cities) before its national release to all drivers on May 3. This pay flexibility product allowed drivers to "cash out" their accumulated earnings prior to the usual payday.[9] Specifically, drivers could press a button in the Uber app that would initiate a transfer of their earnings that would arrive minutes later in a separate debit card. Drivers could complete up to 5 of these cash outs per day.

During the experimental roll-out, drivers in the treatment group who signed up for Instant Pay would be provided with an Uber Debit Card from their implementation partner, GoBank.[10] The debit card had no withdrawal fees (from their network of 42,000 ATMs), no overdraft/NSF/penalty fees, and account maintenance fees (typically $8.95/month) were waived for 6 months every time the driver received a direct deposit or Instant Pay from their Uber earnings. The product continued to pick up usage after the national launch (after our experimental study period). By 2017, roughly a year after the launch, it was reported that "hundreds of thousands" of drivers were using Instant Pay and that $1.3 billion dollars were cashed out in this way (Etherington, 2017). By 2019, the head of Uber Money reported that 70% of driver payments were made using Instant Pay (Son, 2019).

1.3 Experimental Design & Implementation

Enrollment into the study was done on a rolling basis both within and between cities. Each city had a specific start date on which it began contacting drivers, and a city enrolled roughly 80% of its drivers into the study on this common start date. Upon enrollment, drivers were assigned to treatment or control at roughly an even split. In all cases, treatment entailed access to Instant Pay as well as an email informing the drivers of the program; an example of which can be seen in Appendix Figure A.1.[11] The email noted that it is a pilot program, the

[9] Lyft, Uber's primary rival ride-sharing platform, previously piloted a similar pay flexibility product called "Express Pay" in 2015 in conjunction with Stripe, the large financial services provider (Yeung, 2015).

[10] Uber later allowed drivers to link Instant Pay to their own personal debit cards (with most Visa, MasterCard, and Discover cards accepted), but currently charges a $0.85 fee for cash-outs to a personal card, whereas there is still no fee for cash-outs to an Uber Debit Card.

[11] The example email from Southern California addressed the driver by their first name and included the following text, "*Today we're introducing Instant Pay, a product that gives you the flexibility to access your earnings whenever you need them. With Instant Pay, there is no credit check, withdrawal fee, or minimum amount required to cash out. Apply below, and if you're approved, you'll receive your Uber Debit Card from GoBank in 5-8 days!*," followed by a link to apply, a link to a frequently asked questions page, and some graphics and text describing the program (as shown in Appendix Figure A.2).

result of their valued feedback, would be launched to more drivers in the future, and thanked drivers "for all they do." The second page of the email included additional details including a forecast for time to receive the card if approved ("7-10 days"). While all cities contacted drivers in the treatment group upon study enrollment, cities had discretion over whether to send an email to drivers in the control group.[12] Cities also had discretion over whether to send treated drivers additional emails and SMS messages encouraging program participation, including testimonials from fellow drivers and limited-time bonuses for registering. These messages were generally sent to a random portion of a city's treated drivers within the first two weeks of its common start date, so the volume of communications about *Instant Pay* varied both across and within cities.[13]

Because drivers would only be eligible for the study if they had worked a specific number of hours in the previous two weeks, enrollment was rolling within a city after that date. Specifically, each city would execute a uniform SQL query on all drivers who were not yet enrolled in the study. If these drivers newly satisfied an eligibility criterion (based on the number of rides in the previous two weeks exceeding a threshold), they would be "enrolled" into the study, randomized to treatment or control, and contacted accordingly. The randomization is thus essentially stratified by date within each city, and so in our analysis we include cohort fixed effects, where cohort is defined as the date of enrollment interacted with the driver's city, as noted in Section 3.

The experiment ran for two months, beginning in Minneaplis-St.Paul, MN on March 2, 2016 and gradually expanding across markets. While the initial launch plan included plans to randomize across 168 markets in total, Uber pulled the data on April 16 (after

[12] Appendix Figures A.3 and A.4 show an example of the email sent to the control group in the Bay Area market, which was also done in most other markets. The email is quite similar to the treatment group email, but the tagline notes "Coming Soon," the body of the text notes it will be available "in the coming months," and the button at the bottom links to "Keep Me Updated" rather than to the application.

[13] In some markets, drivers in the treatment group were randomized to receive this launch email alone or both the email and SMS. The content of the SMS was, "UBER: You're now able to get paid faster! We're excited to launch Instant Pay due to your valued feedback. Apply at http://t.uber.com/instantpay". Moreover, half of the drivers in the treatment group who didn't sign up within the first seven days were randomized to receive a follow-up email with a driver testimonial and a heading of "Why Other Drivers Love Instant Pay." There were also two other optional emails sent at the discretion of the individual cities. At three days from the launch email, drivers in the treatment group in some cities received an email and/or SMS with the offer to answer questions (the SMS content was as follows: "UBER: We are excited to help you get paid faster than ever before! Come visit us at psc, we'll help you sign up and answer your final questions related to our new Instant Pay product. Don't have questions? Sign up now at http://t.uber.com/instantpay". Finally, there was another optional randomized email that was sent in some cities to drivers who had not signed up within the first 10 days – 30% of these drivers received an incentive (up to $20, but the amount was also at the discretion of the city) if they signed up within the 48 hours following that email.

just 46 markets had begun implementation) and deemed the program an operational success (i.e., minimal/manageable levels of fraud) and launched nationally to all drivers earlier than planned on May 3, 2016. As discussed further in Section 2, our dataset is composed of the two months pulled by Uber in 2016, spanning February 17 to April 16, and thus the last implementing city in our sample is Philadelphia, PA which launched on April 12, 2016.

Appendix Table A.5 shows the experimental roll-out by cities in our sample. The table includes the total number of drivers, modal start date, the proportion of drivers who were enrolled on that date, the proportion assigned to treatment, and the average number of days drivers in that city are observed after their enrollment date (or after the enrollment date plus three weeks thereafter). The last point highlights an issue of implementation discussed further in Section 2.3 – since drivers would only be able to easily use Instant Pay after being approved by GoBank and receiving their debit card (which they were told could take 7 to 10 days), we expect treatment effects to emerge with a lag from enrollment into the study. Using three weeks as a data-driven cutoff for when we might expect to see treatment effects implies that only cities implementing at least three weeks prior to April 16 identify our estimates. Thus, as the penultimate column of the table shows, this effectively reduces our identifying variation to the 13 markets that enrolled before this cutoff, beginning with Minneapolis-St.Paul on March 2 and ending with Los Angeles, CA on March 18 (the other markets are Madison, San Diego, Seattle on March 8, Boston on March 10, and San Francisco, Washington DC, Columbus, Austin, Chicago, Nashville, Dallas on March 14).

2 Data

Our data comes directly from Uber records on drivers enrolled in the randomized trial over February to April 2016. We directly observe a driver's experimental condition and study enrollment date and supplement it with basic characteristics such as age, sex, and Uber driving tenure. We then combine information on trips, Uber Debit sign-ups, and *Instant Pay* cash outs over this time period to construct our main measures of labor supply and program take-up.

2.1 Data Construction and Key Variables

For the majority of drivers, we lack information on the enrolling city and must impute one in order to include cohort (i.e., city-by-enrollment-date) fixed effects. We pull possible city candidates from one file on driver demographics and another detailing the cities in which drivers were working at the time.[14] By comparing each candidate city's modal experiment start date with a driver's enrollment date, we assign a value based on "consistency" and duration between start dates. A candidate city is "consistent' if its common experiment start date is at least as early as that driver's start date – if two cities satisfy consistency, we select the option with the start date closer to the driver's. When we cannot assign a city based on our criteria, we place drivers in a catch-all "Synthetic City" to avoid dropping drivers. There are also a handful of unnamed cities in our data that cannot be matched to cities in the launch plan. We do not modify assignments for drivers that are initially linked with these unknown cities, and fewer than 1% of drivers in our final sample are assigned either the synthetic city or an unnamed city.[15]

Our first labor supply measure is minutes worked per driving session aggregated over a calendar day, where driving sessions are blocks of time in which drivers are completing trips. Sessions may include breaks in which we do not observe an active trip, but dormant periods exceeding 2 hours are used to delineate driving sessions. Minutes worked per session is then defined as the length of the entire session (i.e., the difference between the drop-off time of the last ride in a session and the pick-up time of the first ride of the session). We also define a session's earnings by summing up net earnings from all corresponding trips. A driver's net earnings for a trip is the ride fare less Uber's commission (or a fixed $3.00 per Uber Eats trip). We analogously aggregate session earnings within a specific date for our daily earnings measure. Note that sessions are assigned to the day in which they began, such that a driver who nets $50 from 8am to 10am on March 3 and $100 from 10pm into 2am of the following date will have 2 driving sessions, 360 minutes worked, and $150 assigned to March 3 as both sessions began on that date.

As discussed by Chen et al. (2019), there are various ways to define labor supply for Uber drivers and each carries its own drawbacks. We believe that our method strikes a reasonable balance to estimate the time spent working that is not captured within the trips data, such as waiting for ride requests. Similarly, defining working time at the session level avoids overestimating labor supplied when drivers can work multiple trips simultaneously. For example, drivers may generate several Uber Pool trip records with nested or overlapping start and end times, and simply summing up trip length would double-count minutes worked in these cases.

[14] While Uber drivers can only be actively registered with one region at any point in time, drivers can work anywhere within the same state. Similarly, drivers can change which local market they are registered with.

[15] Dropping drivers in the Synthetic City from our analysis sample has no meaningful effect on estimates.

For treated drivers, we observe the dates they reached various stages of the Uber Debit funnel: submitting an application to GoBank, and if approved, registering for an account, activating the account, and performing a "cash-out". Not all drivers interested in taking advantage of the newly offered payment flexibility were able to activate a card. Around 4.1% of treated drivers who applied were declined, and GoBank surveys reported a mix of reasons why people fell off at different stages, such as technical issues completing the sign-up process, confusion around fees, and diminished interest. As such, we create time-varying measures of take-up for the earliest date drivers reach each of these phases.

We also count the daily number of *Instant Pay* "cash outs" initiated and total amount withdrawn for each driver. These cash outs are transfers out of Uber earnings as *Instant Pay* deposits to their GoBank accounts. Drivers could access their Uber earnings before they received their Uber debit cards if they initiated a second transfer from GoBank to another bank account. Such transactions would be subject to standard processing wait times and limits on the number of bank transfers within a given period, and thus offered limited payment flexibility compared to the full *Instant Pay* product. Still around 5.9% of drivers who ever cash out earnings did so before they activated their cards.

2.2 Balance Analysis and Summary Statistics

Our final dataset is a balanced panel of 224,987 drivers spanning February 17 to April 16, 2016. Appendix Table A.1 provides driver-level baseline (pre-experiment) summary statistics and balance by treatment condition. T-tests of the four variables taken from Uber's demographics table (sex, age, median household income of driver's home Census Block Group, driver's previous Uber experience) fail to reject the null hypotheses of equality between the treatment and control groups.[16] The same is true for the measures of labor supply in the two weeks prior to the experiment's first enrollment date (i.e., February 17 to March 1). Finally, we report an F-test for joint significance, which also fails to reject the null (F=1.32). This specification includes cohort fixed effects to match our regression analysis, but omitting the fixed effects also produces a similar F-statistic (1.22).

The table shows that 16% of drivers are female, the average age is 43, and the typical driver lives in Census Block Group with a median income of sixty-five thousand dollars (using the 2016 American Community Survey 5-year). Drivers in our sample average one session of

[16]Driver experience in months is computed by taking the difference between April 16, 2016 and the date on which a driver was onboarded to Uber (and thus may span long breaks). We are missing coverage on demographics for about 6% of drivers.

work every two days, for 111 minutes and $38 per day. This aggregates to average totals of roughly 8 shifts, roughly 1,660 minutes, 6 days worked, and $574 in this baseline period.

2.3 Take-Up and *Post-Treatment*

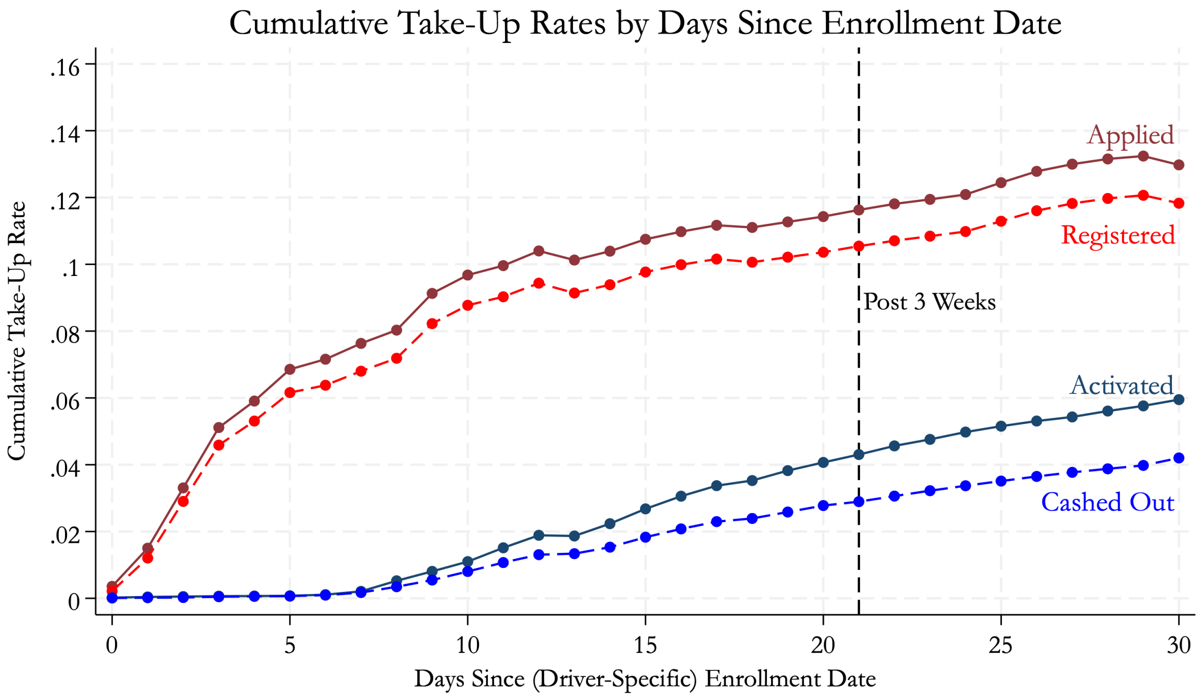
Since converting an emailed *offer* of Instant Pay into actual *access* to the service requires a number of steps, a key issue is how to define when the *post-treatment* period starts. If we define *post* too early (e.g., on the date of the enrollment/email) this may attenuate treatment effects by identifying off days in which the driver won't yet have the ability to use the product. This is because there are access delays due to at least three margins: (a) The driver needs to first apply (requiring first noticing/reading the email, and then filling out the application), (b) GoBank has to review the application to decide if the driver will be approved, (c) After approval, there may are delays in mailing the debit card (as the email noted "*you'll receive your card in 7-10 days*"). On the other hand, defining the *post* period too late risks inflating standard errors as there will be fewer days identifying the treatment effect. We thus face a bias-variance trade-off in this decision.[17] Because our post-implementation observation window is short (roughly 33 days on average) this issue is of first-order importance.[18]

To take a data-driven approach to defining the post period, in Figure 1 we look at cumulative take-up rates by the number of days since a driver is enrolled in the experiment. This figure suggests that take-up, as proxied by application, appears to level off somewhere between 10 and 25 days after enrollment, while activation instead appears to continue on a somewhat linear pathway from 7 out to 30 days. There is no clear cut-off point across all four measures of take-up, but 21 days appears to somewhat balance the concerns, and so we make this our primary measure of *Post*. In Section 4.3.1, we show robustness to varying this definition from 0 to 30 days.

[17]The coefficient may also vary if there are heterogeneous treatment effects by "time since gaining access." For example, if there is a novelty effect, then we may expect larger effects in the earlier days of the window (which would counteract the attenuation discussed earlier when defining the *Post* window earlier). On the other hand, if drivers gradually ramp up their usage of the card, this would suggest larger effects in the other direction (producing larger coefficients the further out the *Post* window date is defined).

[18]Limiting to the 57,229 treated drivers who were enrolled at least 22 days prior to April 16, we observe them roughly 33 days after enrollment on average, with the 10th percentile being 29 days and the 90th being 37 days. As noted below, we focus on treatment effects 21 days after enrollment, so our treatment effects are identified off 13 known cities with experiment start dates at least three weeks before the April 16 cutoff and a handful of drivers (498) in the catch-all synthetic city. This group of early cities consists of 113,892 drivers with 57,229 assigned to treatment and 56,663 to control.

Figure 1: Take-Up Rates by Days Since Enrollment (Launch Email Date)



Notes: This figure's sample is composed of drivers in the treatment group and includes one observation per driver-day. Each point corresponds to the mean on the day since the (driver-specific) enrollment date. For example, if one driver was enrolled into the study on March 2 and another started on March 9, Day 1 would correspond to March 3 and 10 respectively for those drivers. Note that since this is a cumulative measure, and the post-enrollment data windows vary by cohort/cities, the denominator can change throughout (e.g., explaining the downward dip between days 12 and 13).

Finally, Panel A of Appendix Table A.2 shows take-up rates by treatment group among the 113,879 drivers who have at least one observation in the post-treatment period (i.e., at least one day observed at least 21 days after their experiment enrollment). We limit to this sample as the take-up rates in this group are relevant for our primary estimation sample (i.e., for comparing the intent-to-treat and treatment-on-the-treated estimates). We see that in this sample take-up measured by application is 13%, by registration is 12%, by activation is 6%, and by ever making a cash-out before the study ending is 4%.[19] The table also shows that there is only one-sided non-compliance, as the take-up rates are zero in the control

[19]If we expand to the full sample of 224,987 drivers (including the drivers who were enrolled too late to observe observations 21 days after their enrollment), these treatment group take-up rates are 11%, 11%, 4%, and 3% respectively. These figures are relevant for calculating the TOT if we define "post-treatment" as beginning on the date of enrollment, as shown in Figure 2 discussed in Section 4.3.1.

group.

Panel B of Appendix Table A.2 telegraphs our regression estimates by showing simple t-tests of differences in the outcomes using just post-treatment observations. These differences capture the intent-to-treat (ITT), i.e. the effects of *offer* of treatment. Using this driver-day level sample, we see that work time (total minutes driven) increases by 3 minutes (roughly 2.2% of the control group mean), earnings by $1.13 (or 2.2%), and session count by 0.01 (1.8%). In Section 4 we estimate these treatment effects more precisely using the full panel structure.

3 Empirical Strategy

We first examine the impact of the randomized *offer* of *Instant Pay* on driver outcomes, i.e. the intent-to-treat (ITT) of *Instant Pay*. We use the full two months of our balanced panel (February 17 to April 16), and thus the number of pre- vs. post-implementation observations varies within and between cities (e.g., implementation started in Houston on March 30, and we thus have 6 weeks of pre-enrollment data and roughly 2 weeks of post-enrollment data for that city, though for some Houston drivers, enrollment begins after March 30).[20] As discussed in Section 1, since both cities and drivers became eligible at different points in time, we implement a specification that includes fixed effects for the randomization stratification unit of the cohort (i.e., city-by-enrollment-date) interacted with date fixed effects to account for temporal dynamics. Finally, as noted in Section 2.3, we define the post-*treatment* period as 21 days after enrollment to account for delayed sign-up and mailing of debit cards (see Section 4.3 for robustness to this choice).

Specifically, we estimate the following driver-day level regression specification:

$$y_{izj} = \beta_1 * Treatment_{iz} * Post_{izj} + \eta_{zj} + \epsilon_{izj}, \tag{1}$$

where y_{izj} is an outcome (work time, earnings, or the number of sessions) for driver i working in city z on date j, regressed on an indicator for whether the driver is in the treatment group ($Treatment_{iz}$) interacted with a time-varying indicator variable that takes on the value 1 three weeks after the driver's study enrollment date ($Post_{izj}$), as well as a set of cohort-by-date fixed effects (η_{zj}). We cluster standard errors at the unit of randomization (driver). The key coefficient from Equation 1, the reduced-form parameter β_1, captures the effect of being

[20] By enrollment, we mean the date on which drivers were randomized to their treatment condition and received their corresponding emails.

offered Instant Pay.

We examine outcomes in terms of levels, as well as with specifications that allow interpretation in percentages (e.g., to facilitate comparisons to other labor supply elasticities). We show results with the common log(1+Y) transformation that behaves like log(Y) at large values but is defined at zero. However, recent work has shown that these transformations are sensitive to the units of the outcome (Chen and Roth, 2023). As such, our preferred estimates use Poisson regression, a consistent estimator of the ATE in percentages (Chen and Roth, 2023; Wooldridge, 2010).[21]

As discussed in Section 2.3, take-up is somewhat mechanically depressed due to the short observation window and logistical/implementation challenges. Since we're especially interested in the response of those with access to the card, and would like to forecast the treatment effects in the present (where take-up may be as high as 70%), we also present treatment-on-the-treated (TOT) estimates, i.e., we scale up the ITT to account for the fact that not all drivers took up the treatment before the end of our observation window. That is, we estimate instrumental variable specifications where the second stage is of the form:

$$y_{izj} = \beta_1 * Takeup_{iz} * Post_{izj} + \eta_{zj} + \epsilon_{izj}, \tag{2}$$

where $Takeup_{iz}$ is an indicator that takes value 1 if a driver *ever* took up the *Instant Pay* at any point prior to the end of the observation window (April 16). Of course, takeup is not random and may be correlated with unobservables, and so we instrument $Takeup_{iz} * Post_{izj}$ with the variable $Treatment_{iz} * Post_{izj}$ which is uncorrelated with the error due to randomization. As with the ITT, $Post_{izj}$ is defined as 21 days after the driver-specific enrollment date, allowing a simple comparison of the ITT and TOT (i.e. the latter point estimate is equal to the ITT divided by the take-up rate).[22]

We show TOT estimates in which *Takeup* is defined as either: (1) Applying for the card, or (2) Activating the card after receiving it in the mail. Each of these definitions has trade-offs which we discuss further in Section 4.2. The key identification assumption is that the assignment of treatment only affects labor supply through the channel of takeup (i.e.,

[21] This is implemented in Stata using *ppmlhdfe* to facilitate the inclusion of high-dimensional fixed effects.

[22] An alternative is to instead allow for time-varying measure of takeup that turns on when the driver does so (and stays on thereafter) and to instrument this with the binary treatment indicator. We show in the Appendix that this specification delivers qualitatively similar results to our primary specification. We prefer the version shown in equation 2 because it enables a more straightforward scaling of the ITT just by the proportion of drivers who eventually take up in the cross-section, rather than mixing variation driven by the endogenous *timing* of takeup.

through application in case 1, or the receipt and activation of the card in case 2). Again, our preferred estimate uses Poisson regression; however, because the inclusion of fixed effects precludes us from using the *ivpoisson* Stata command for estimation (Angrist, 2001; Chen and Roth, 2023), we use a control function approach and the delta method to calculate standard errors as suggested by Lin and Wooldridge (2019). That is, we estimate the first stage via ordinary least squares regression (OLS) and then add the residuals from the first stage as a control in the second-stage Poisson regression.[23]

4 Main Results

This section describes the impacts of Instant Pay on driver labor supply. We start by reporting the intent-to-treat (ITT) estimates, i.e. the effects of the *offer* of Instant Pay on measures of labor supply. Then we turn to the treatment-on-the-treated (TOT), i.e. the effects of Instant Pay among those who take it up. For the TOT, we show results using two possible measures of take-up: ever applying for Instant Pay and ever activating an Instant Pay account (before the April 16 end of our observation window). We show all results using our preferred approach (Poisson regression) as well as OLS using both transformed (log(1+Y)) and untransformed outcomes.

4.1 Intent-to-Treat (ITT)

Table 1, Panel A reports the intent-to-treat (ITT), i.e., estimates of equation 1. Columns 1-3 report our preferred estimates using Poisson regression, while columns 4-6 use OLS after transforming the outcomes to log(1+Y), and finally columns 7-9 report OLS specifications of the untransformed outcomes. Our first measure of labor supply is the number of minutes worked over a day. Column 1 shows that the offer of Instant Pay increased daily labor supply by 2.4% in terms of minutes worked, while Column 2 shows that this is reflected in earnings similarly increasing by 2.4%.[24]. Finally, Column 3 shows that the number of distinct work sessions increases by 1.5%, suggesting that the total increase in work time comes both from longer shifts and more shifts. All results are statistically significant at the 1% level.

[23] In the Appendix we also estimate using *ivpoisson* and no fixed effects (using just the post-treatment observations, since panel estimates would be biased in absence of the fixed effects) – we obtain qualitatively similar estimates using this approach as well.

[24] It's not obvious that work time and earnings should increase in lockstep. For example, it could have been that drivers already worked at times of day with the highest earnings opportunities, and thus that the marginal product of labor is diminishing.

Columns 4 to 6 report results with a log-like transformation estimated by OLS. Work time increases by 0.037 log points, earnings by 0.031 log points, and sessions by 0.005 log points. While it's common to translate these into percentage effects (i.e., 3.7%, 3.1%, and 0.5%) recent work suggests this may not be warranted, as these log-like transformations are sensitive to the scale of the units, unlike percentages (Chen and Roth, 2023).

Finally, Columns 7 to 9 report results OLS regressions in levels. Minutes increases by 3.33 minutes, earnings by $1.23, and sessions by 0.01. Dividing these point estimates by the pre-period mean of the dependent variable in the control group suggests these outcomes increased by 2.8%, 2.9%, and 1.7% respectively. These results are all quite similar to the Poisson regression results, suggesting the ITT estimates are robust to the estimation method.

Table 1: Main Results

	Poisson			Log(1+Y)			Levels		
	Minutes	Dollars	Sessions	Minutes	Dollars	Sessions	Minutes	Dollars	Sessions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Intent-to-Treat (ITT)									
Treat X Post	0.024***	0.024***	0.015***	0.037***	0.031***	0.005***	3.327***	1.229***	0.010***
	(0.006)	(0.006)	(0.005)	(0.011)	(0.009)	(0.002)	(0.844)	(0.316)	(0.003)
N	13,403,588	13,403,370	13,403,588	13,498,680	13,498,680	13,498,680	13,498,680	13,498,680	13,498,680
DepVarMean	121.524	41.999	0.580	2.281	1.845	0.355	120.586	41.674	0.575
Panel B: IV (TOT, Take-Up = Applied)									
Applied X Post	0.176***	0.176***	0.112***	0.276***	0.229***	0.040***	24.587***	9.085***	0.071***
	(0.045)	(0.045)	(0.037)	(0.083)	(0.069)	(0.013)	(6.243)	(2.336)	(0.023)
N	13,403,588	13,403,370	13,403,588	13,498,680	13,498,680	13,498,680	13,498,680	13,498,680	13,498,680
DepVarMean	121.524	41.999	0.580	2.281	1.845	0.355	120.586	41.674	0.575
Panel C: IV (TOT, Take-Up = Activated)									
Activated X Post	0.369***	0.372***	0.229***	0.588***	0.486***	0.085***	52.263***	19.312***	0.152***
	(0.095)	(0.096)	(0.079)	(0.177)	(0.146)	(0.028)	(13.273)	(4.967)	(0.049)
N	13,403,588	13,403,370	13,403,588	13,498,680	13,498,680	13,498,680	13,498,680	13,498,680	13,498,680
DepVarMean	121.524	41.999	0.580	2.281	1.845	0.355	120.586	41.674	0.575

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Panel A reports ITT estimates from regressions with fixed effects for the interactions of randomization cohort (city-by-enrollment-date) and date. The "Post" period is defined as beginning 3 weeks (21 days) after the (driver-specific) enrollment date. Columns 1-3 report Poisson regressions, while columns 4-6 report OLS specifications in which the outcome variable is the logarithm of the outcome value plus one. Columns 7-10 report OLS of the untransformed outcome. Panel B reports instrumental variable estimates using the randomized assignment of treatment as an instrument for take-up (defined as the driving have ever applied for a card prior to April 16; mean = 13%). Panel C is similar but uses having ever activated a card as the definition of take-up (Mean = 6%). Columns 1 to 3 of Panels B and C use a control function approach where the first stage is estimated via OLS (*reghdfe*), residuals are included as a control in the second-stage Poisson regression (*ppmlhdfe*), and the delta method is used to compute standard errors (using *margins, eydx*). DepVarMean is the mean in the control group in the pre-period. Standard errors are clustered by driver.

4.2 Treatment-on-the-Treated (TOT)

Panels B and C report the TOT estimates, i.e., IV estimates of equation 2 in which the randomized offer of Treatment instruments for Take-Up, both interacted with a post-treatment indicator. Panel B focuses on the loosest measure of take-up, which is whether the driver submitted an application. As shown in Appendix Table A.2, roughly 13% of drivers in the sample ever applied. In Panel C, the measure of take-up is whether the driver ever activated their Uber debit card (which would require first applying for the card and registering for GoBank online), something that roughly 6% of drivers completed before April 16. Among drivers who ever activated a card, it took roughly 6.9 days on average for drivers to apply for the card after receiving the first enrollment/launch email, and roughly 18.1 days to activate their card.

It's not obvious which TOT measure is most relevant *a priori*. On the one hand, cashing out seems a natural candidate as it reflects the most direct measure of actually utilizing the product to change pay frequency. After all, drivers who do not cash out have not had an actual change in their pay frequency. However, it's plausible that drivers shift their labor supply after activating their card – as it's the presence of the card that changes the *ability to access* additional earnings in the event of a cash need that may not yet have arisen before the end of our observation window (Among drivers who ever activated a card, we observe them for 15.1 days on average after activating their card and 15 days on average after their first cash out). For this reason, we prefer *activation* over *cash-out* as a measure of access. However, even *activation*, comes well after drivers have taken a formidable step in the take-up process. As such after application, drivers may adjust their behavior in anticipation before receiving the card. We present both estimates, and only give preference to the latter (i.e., take-up = application) as a more conservative measure.

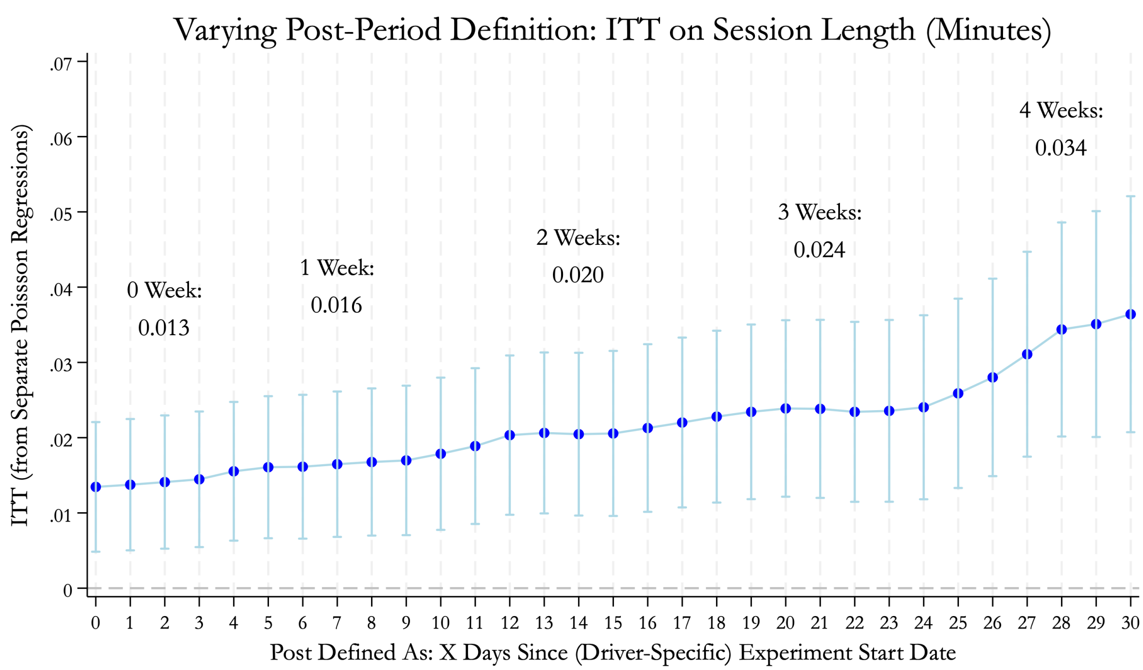
Columns 1 to 3 of Panel B, Table 1 – using *Applied* as a measure of Takeup – finds TOT effects of daily work time increasing by 17.6%, earnings by 17.6%, and the number of sessions by 11.2%. In terms of levels estimated via OLS (Columns 7 to 9), and percentages relative to the control group mean, these are increases in worktime of 24.6 minutes (20%), earnings of $9.09 (22%), and the number of sessions of 0.07 (12%). Using card *Activation* as the measure of take-up, Panel C shows work time increasing by 36.9%, earnings by 37.2%, and sessions by 22.9%. The level increases are 52.3 minutes (43%), $19.31 (46%), and 0.15 sessions (26%).

4.3 Robustness

4.3.1 *Post* Window Definition

A natural question is how to define the *Post* period. As noted in Section 3, if we define *Post* too early for drivers to receive their cards, then we potentially attenuate the treatment effect. By contrast, if we define *Post* too late, then we both reduce power by limiting the observation window and may mischaracterize the total treatment effect if there is dynamic heterogeneity (e.g., novelty effects). In our primary analysis, we settled on 21 days after enrollment to define *Post*; however, since this choice was not preregistered, one may naturally be concerned about specification search. Figure 2 repeats the specification shown in Table 1, Panel A, Column 1 for 30 different possible definitions of *Post*, ranging from 0 to 30 days. We see that the effect remains statistically significant across all of these, and that our preferred estimate (2.4% at 21 days) is roughly in the middle between 0 days (a 1.3%) and 28 days (3.4%).

Figure 2: Robustness to *Post* Window Definition



Notes: Each dot corresponds to a coefficient on the interaction term (Treat X Post) from a separate Poisson regression for each "Post" definition, ranging from 0 to 30 days. The bars correspond to the 95% confidence interval (+/- 1.96*SE) on that point estimate. The Poisson regressions include fixed effects for the interactions of randomization cohort and date. Standard errors are clustered by driver.

4.3.2 TOT Estimation

Another degree of freedom is how to estimate the treatment-on-the-treated given the staggered roll-out and endogenous timing of take-up by drivers. Our primary approach is to define "ever take-up" as a time-invariant property of a driver and to instrument this with the similarly time-invariant treatment indicator (with both terms being interacted with a "post" variable capturing the exogenous enrollment timing, plus 21 days). This has the desirable property of preserving the usual cross-sectional logic that the TOT is simply the ITT divided by the take-up rate. By contrast, another approach found in the literature (e.g., Atkin et al. (2017)) is to instead construct a time-varying measure of take-up that turns on when the driver *chooses* to apply or activate their card (and stays on thereafter), and to then simply instrument this with the binary treatment indicator. Because this scaling mixes both the take-up rate and the endogenous timing of take-up, it does not facilitate as straightforward a comparison to the ITT. Nonetheless, for completeness, we estimate this specification as well in Appendix Table A.3. Comparing the primary Poisson IV specifications to Table 1 we see that this approach produces somewhat larger coefficients on minutes (22.2% vs. 17.6%) and dollars (21.4% vs. 17.6%) that remain significant at the 5% level, but a slightly smaller effect on session count (9.5% vs. 11.2%) that crosses the significance threshold. When take-up is defined as activation, the differences are even larger across minutes (82.1% vs. 36.9%), earnings (78.2% vs. 37.2%), and sessions (35.4% vs. 22.9%), again with only the first two remaining significant. We also see that the log(1+Y) specifications lose significance, though the estimates in levels remain robust.

A related question is whether to use the baseline (pre-enrollment) observations at all. We do so to increase the precision of our estimates; however, this requires including cohort fixed effects to ensure the staggered roll-out does not introduce bias. The complication this poses is that estimating the TOT in our preferred Poisson specification requires a control function approach and using the delta method to compute standard errors. By contrast, if we only used the post-treatment observations, we could more safely drop the fixed effects and use *ivpoisson* in Stata (using a two-step generalized method of moments estimator) which will calculate the correct clustered standard errors. In Appendix Table A.4, we reproduce Table 1 using just the post-treatment (i.e., more than 21 days after the driver-specific enrollment date) observations and omitting any fixed effects. Panel A shows that this produces quite similar ITT point estimates (e.g. 2.1% vs. 2.4% for minutes in Column 1), with all remaining significant at the 1% level. Columns 1 to 3 of Panels B and C are our main interest, where we now estimate using *ivpoisson gmm*, and we again see quite similar point estimates (e.g., 14.9% vs. 17.6% for minutes in Panel B and 29.3% vs. 36.9% in Panel C) and standard errors.

4.4 Magnitude

One potential complexity in interpreting the magnitude of our results is that we observe only on-platform labor supply, and most Uber drivers also do non-Uber work. This may complicate the interpretation of our measured elasticities – is Instant Pay making working on Uber more attractive in absolute terms, or only relative to other work? Perhaps most importantly, many Uber drivers also work on non-Uber gig platforms, and substitution away from these may represent relatively modest changes in the absolute attractiveness of work on Uber.

To provide context for our results which helps address this complexity, we first consider how the observed labor supply responses compare to previously estimated wage elasticities among Uber drivers (Chen et al., 2019). Labor responses to experimentally randomized increases in Uber wages share the same interpretational issues we discuss above, and allow us both to ask if responses to Uber wages seem non-normatively high, and to translate Instant-Pay labor-supply responses to equivalent increases in Uber wage rates. Chen et al. (2019) find a median driver-wage labor supply elasticity of 1.92 (25th percentile: 1.81; 75th percentile: 2.01). As noted there, these elasticities are somewhat high relative to estimates from the labor economics literature, but they note that the bulk of Uber drivers in their sample work roughly 10 hours per week and thus have more scope for adjustment than other classes of workers.

To translate our estimates to a wage elasticity, it's not *a priori* obvious whether to use the ITT or TOT. While wage shifts have perfect compliance (i.e., they apply to all drivers), the programmatic costs of Instant Pay (e.g., subsidizing transaction fees) are only borne by the drivers who take it up. This pushes in the direction of benchmarking relative to the TOT for considering cost-effectiveness. Similarly, for thinking about the behavioral mechanism, the TOT may be closer to the right individual-level estimate, as the lower ITT is mechanically driven to zero by the low take-up during the short pilot period (itself an artifact of the gradual diffusion of take-up that may have reached as high as 70% by 2019). On the other hand, the TOT may not extrapolate to the full sample, as those with the highest valuation for Instant Pay (and thus, possibly the largest LATEs) may be the earlier adopters.[25] As a

[25]Because our measure of take-up and utilization is relatively short, it's also plausible that early adopters are the most attentive (to emails or notifications), and the correlation between attentiveness and valuation

result, both benchmarks may be useful.

If we use 1.92 as the wage elasticity, this implies that the ITT estimate of a 2.4% (work minutes) labor supply response to the offer of *Instant Pay* is roughly equivalent to raising wages for all drivers by 1.24% (2.4/1.92). Similarly, the TOT using application (Panel B of Table 1) of 17.6% would imply is roughly equivalent to raising wages by 9.2%, while the TOT using activation of the card (Panel C of Table 1) of 36.9% is equivalent to raising wages by 19.2%. All of these suggest that Instant Pay makes work on Uber considerably more attractive to Uber drivers.

5 Mechanism and Heterogeneity

What mechanism can explain the observed labor supply responses to flexible pay? As previously noted, Instant Pay made the effort costs and financial benefits of work roughly coincident, reducing the lag between them by several days. As such, a leading explanation is that the response has something to do with time preferences. However, the fact that we observe such a large response is difficult to reconcile with exponential discounting. This leads us to hypothesize that the response is either driven by sharply-binding liquidity constraints (combined with pressing financial needs) and/or by hyperbolic discounting. While we cannot directly observe liquidity constraints or time preferences, we provide some indirect evidence suggesting the response may be driven by the latter. It's worth noting that even if we did observe liquid assets, it would still be challenging to disentangle preferences from liquidity, as preferences can drive observed liquidity levels, a point well made by Gelman (2022).[26]

To make some headway in understanding mechanisms, we take three tracks. First, as noted in Section 4.4, the TOT response is roughly equivalent to a 19% increase in wages. Since the product is effectively reducing the average delay between effort and pay by 6 days (weighting each day of the week equally), this is a 6-day discount factor of 0.84, and so under an exponential discounting model it would imply an annual discount factor of 0.00002268 ($= 0.84^{365/6}$). Such a discount factor is implausible as all drivers have a savings vehicle (the linked checking accounts required of all Uber drivers to receive the usual direct deposit) and have access to short-term credit such as payday loans at much lower rates.

of Instant Pay is unclear (and thus the TOT may underestimate the TOT for individuals who eventually adopt).

[26]Gelman (2022) uses linked consumption and liquid asset data paired with a buffer stock model of consumption to implement an Euler equation test. He concludes that excess sensitivity of consumption to income (paychecks) is ultimately driven by quasi-hyperbolic discounting (though liquidity constraints, themselves driven by preferences, are a proximate cause).

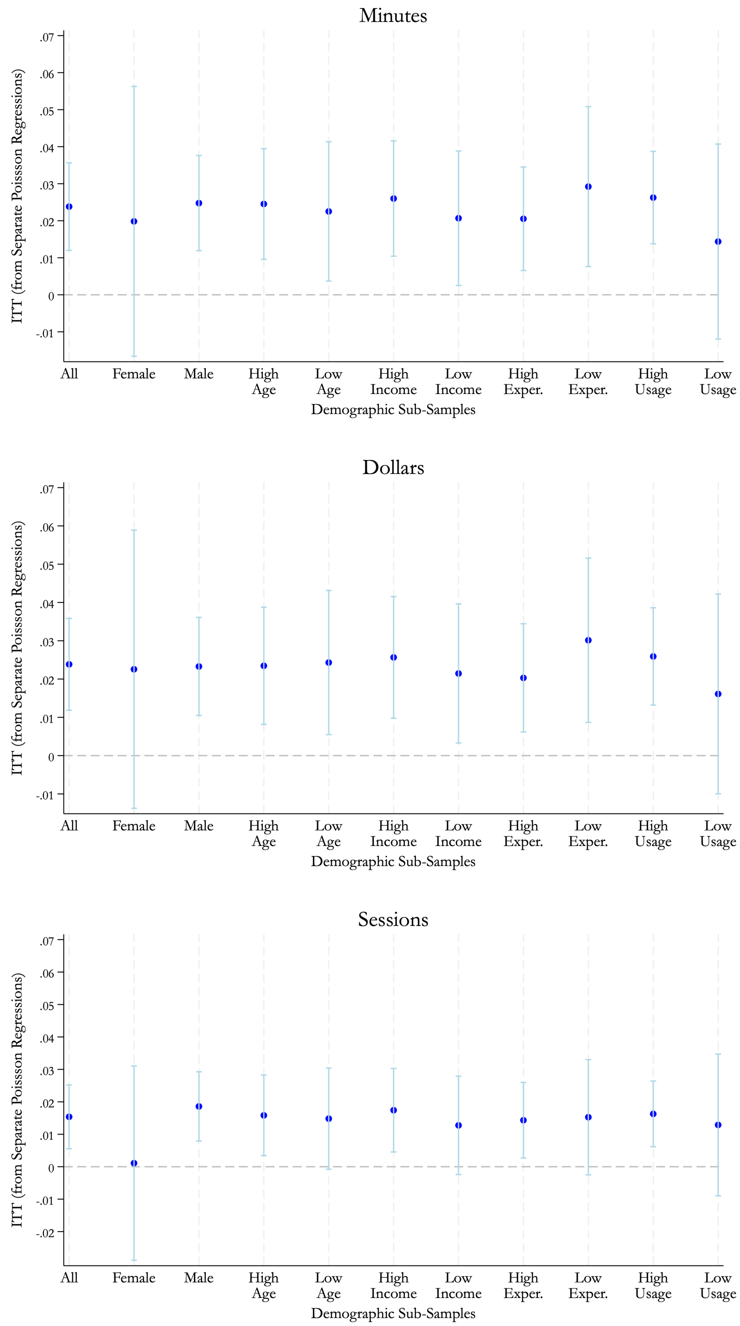
By contrast, it could be reconciled by a quasi-hyperbolic discounting with a short-term discount factor of $\beta = 0.84$ and a long-term discount factor close to 1. While this would still imply a large present bias relative to the literature, it's close to estimates using real-effort tasks, e.g. Augenblick and Rabin (2019) who find a $\beta = 0.83$. While this simple calculation mostly dismisses the exponential discounting explanation, in conjunction with the institutional details (accessible credit at lower rates and the existence of a savings account), it also sheds light on the plausibility of liquidity constraints as the ultimate cause.

Second, we examine heterogeneity by drivers' background characteristics at baseline. Figure 3 repeats the specification reported in Column 1, Panel A of Table 1, i.e., the ITT specification estimated using Poisson regression, first for the full sample and then for various sub-samples. In particular, we find little evidence for heterogeneity across all three outcomes by the sample splits. These include sex (male/female) and median splits on age (42), proxied income ($58,750), experience (16 months), and usage (0.4 sessions per day). Of particular note is the lack of heterogeneity by the median income of the drivers' home block group. Notwithstanding the critique noted in (Gelman, 2022), we would expect a larger response among lower-income drivers if liquidity constraints drove the labor supply response, but we do not see this result directionally or statistically. While the measurement error implied by using the income proxy could drive this difference to zero, it's at least suggestive against the liquidity constraints explanation.

Finally, we examine more model-driven patterns in heterogeneous treatment effects implied by present bias and liquidity constraints respectively. In particular, drivers in the control group had to wait 9 days to receive pay for work done on Monday, 8 days for Tuesday, and so on until a gap of 3 days for work done on Sunday. Instant Pay shrunk this gap to zero across all days of the week. Certain formulations of present bias would predict a behavioral response that shrinks as the counterfactual pay delay gets smaller (in particular, the 1/N hyperbolic discounting model, but not the quasi-hyperbolic discounting model). Figure 4 shows this pattern when we separately estimate by each day of the week, with a significant 3.1% ITT on Monday mostly decreasing out to an insignificant 0.8% ITT on Sunday. By contrast, liquidity constraints and pressing financial needs either predicts no systematic pattern (i.e., labor supply responses will match idiosyncratic financial shocks) or that the pattern will match background variation in liquidity. One liquidity event that we know with certainty in the control group is the usual paycheck direct deposits on Wednesday. If liquidity constraints are operative, we expect the smallest response on Wednesday when control group drivers receive their usual paycheck and for the response to increase out to Tuesday. While

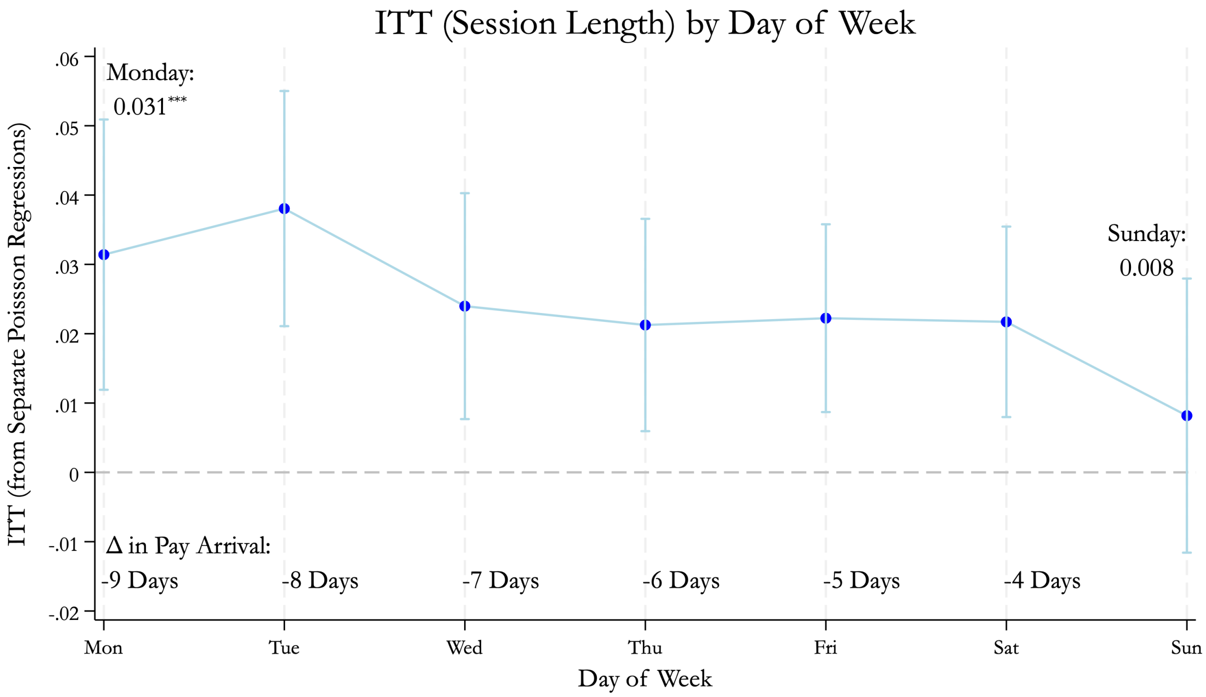
Figure 4 doesn't show this pattern of increasing effects from Wednesday to Tuesday, we run a horserace through regressions in Table 2. Specifically, we interact $Treatment_{iz} * Post_{izj}$ with categorical variables counting the number of days from Monday (0 to 6) or from Wednesday (0 to 6). Column 2 suggests this linear specification with just the Days-from-Monday interaction parallels the sub-sample analysis fairly well, showing an ITT of 3.5% on Monday and this effect decreasing by 0.3pp each day out. By contrast, Column 3 shows that the interaction with Days-from-Wednesday is insignificant (though in the direction predicted by liquidity constraints). The coefficient however shrinks closer to zero in Column 4 when both interactions are estimated, with only the Days-from-Monday interaction remaining significant in Panel A. Together, the three exercises suggest that present bias better explains the observed pattern than liquidity constraints on their own.

Figure 3: Heterogeneity by Demographic Sub-Samples (ITT)



Notes: Each dot corresponds to a coefficient on the interaction term (Treat X Post) from a separate poisson regression for each sub-group subsample. The bars correspond to the 95% confidence interval (+/- 1.96*SE) on that point estimate. The Poisson regressions include fixed effects for the interactions of randomization cohort and calendar date. Standard errors are clustered by driver.

Figure 4: ITT by Day of Week



Notes: Each dot corresponds to a coefficient on the interaction term (Treat X Post) from a separate Poisson regression for each "Day of the Week" sub-sample. The bars correspond to the 95% confidence interval (+/- 1.96*SE) on that point estimate. The Poisson regressions include fixed effects for the interactions of randomization cohort and date. Standard errors are clustered by driver.

Table 2: Days-from-Monday vs. Days-from-Wednesday

	(1)	(2)	(3)	(4)
Panel A: Minutes				
Treat X Post	0.024***	0.035***	0.020***	0.032***
	(0.006)	(0.009)	(0.007)	(0.010)
Treat X Post X Days Since Monday		-0.003**		-0.003*
		(0.002)		(0.002)
Treated X Post X Days from Wednesday			0.002	0.001
			(0.001)	(0.001)
N	13,403,588	13,403,588	13,403,588	13,403,588
DepVarMean	121.524	121.524	121.524	121.524
Panel B: Dollars				
Treat X Post	0.024***	0.032***	0.020***	0.029***
	(0.006)	(0.009)	(0.007)	(0.010)
Treat X Post X Days Since Monday		-0.003		-0.002
		(0.002)		(0.002)
Treated X Post X Days from Wednesday			0.002	0.001
			(0.001)	(0.001)
N	13,403,370	13,403,370	13,403,370	13,403,370
DepVarMean	41.999	41.999	41.999	41.999
Panel C: Sessions				
Treat X Post	0.015***	0.021***	0.017***	0.025***
	(0.005)	(0.007)	(0.005)	(0.008)
Treat X Post X Days Since Monday		-0.002		-0.002
		(0.001)		(0.001)
Treated X Post X Days from Wednesday			-0.001	-0.001
			(0.001)	(0.001)
N	13,403,588	13,403,588	13,403,588	13,403,588
DepVarMean	0.580	0.580	0.580	0.580

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Column 1 repeats the specification used in Table 1 to estimate the ITT on minutes worked (Panel A), earnings (Panel B), and session count (Panel C). Column 2 modifies this by including an interaction with Days-from-Monday which goes from 0 (Monday) to 6 (Sunday), while Column 3 instead does so for Days-from-Wednesday from 0 (Wednesday) to 6 (Tuesday). Column 4 reports a specification with both interactions included. Specifically, all specifications are from Poisson regressions with fixed effects for the interactions of randomization cohort (city-by-enrollment-date) and date. DepVarMean is the mean in the control group in the pre-period. Standard errors are clustered by driver.

6 Conclusion

Using a nationwide experiment at Uber, this paper shows that allowing relatively instant access to accumulated earnings can increase the labor supply of workers. While previous work has documented how variation in pay frequency can affect consumption in contexts with limited work flexibility (Baugh and Correia, 2022; Zhang, 2022), this paper instead focuses on labor supply responses to introducing flexible *pay* in a context with flexible *work*. Although standard models suggest that changing pay frequency from weekly to on-demand should have little effect, we find that the labor supply response is on par with increasing wages by 19%. Moreover, the increase in work hours is concentrated in days further away from the previous/control payday, in line with previous research documenting payday effects in labor supply (Kaur et al., 2015).

Estimating the welfare effects of *Instant Pay* is beyond the scope of this paper; however, we discuss some potential channels through which both consumer and driver welfare may be affected. On the consumer side, by increasing the supply of drivers at any given moment, the policy may lower wait times and slightly depress fares via the surge multiplier – both of which should increase consumer welfare.[27] Second, among liquidity-constrained drivers, increasing pay frequency may improve ride quality by reducing the psychological effects of financial concerns (e.g., improving drivers' attention and focus as in Kaur et al. (2022)).[28]

The effects of Instant Pay on driver welfare depend on whether drivers are sophisticated or naïve hyperbolic discounters. If drivers do not suffer from present bias, then Instant Pay may increase welfare by relaxing very tightly binding liquidity constraints or by allowing them to withdraw pay at a cadence better matching their expenses (e.g., withdrawing on Mondays if lumpy expenses fall on Mondays rather than Wednesdays, which may matter if drivers misforecast their expenses). If drivers are sophisticated hyperbolic discounters (i.e.,

[27]It's worth noting that we do not see evidence of statistically significant price changes. On the driver side, even if prices slightly decrease, if the increase in consumer demand lowers slack time between rides, then the hourly wage rate may still increase. Assuming that time between rides is costlier for drivers than the additional miles driven, this could increase driver welfare.

[28]Kaur et al. (2022) test the effects of varying the timing of payment while holding total pay constant – specifically, they employ Indian factory workers for a two-week contract and randomize some to be paid part of the total amount 4 days early. They find that these treated workers exhibit a 7% increase in output in the final 4 days and find evidence this is through fewer costly, unintentional mistakes by alleviating the psychological effects of financial strain. However, whether these financial strain effects would be relevant among Uber drivers in the US (who have relatively stable employment prospects) is unclear. As they note, "*the relative liquidity boost from being paid daily or monthly would be different in long-run employment—indeed, having such stable employment would limit the likelihood that a worker faces large financial strain in the first place (e.g., Morduch and Schneider, 2017).*"

they correctly forecast their future periods' labor-leisure trade-offs) then they will only enroll in Instant Pay if it is welfare-improving. This is because enrollment has a hassle cost and only facilitates early access to payment after getting the card in a future period (typically two to three weeks later). Thus, if infrequent (weekly) pay serves as a commitment device to save or limit temptation spending, a sophisticate would only undo it by opting into Instant Pay if they correctly anticipate that the additional incentive to work would fund the increase in consumption in a way that at least weakly increases present discounted expected utility.[29] There is, however, ambiguity on welfare in the case of full or partial naivete.[30] A naïve driver may enroll in the program to alleviate an anticipated future mid-week expense but fail to realize that the additional liquidity will create a permanent increase in temptation. On the other hand, they may also underestimate how this future temptation to spend will increase their incentive to work. This second biased belief about work incentives could be enough to outweigh the distortion created by the self-control problem over consumption.[31] Finally, while we've focused on present bias and beliefs about present bias (naivete), it's possible that workers also misforecast expenses (e.g., suffer from "budget neglect") and that faster access to pay exacerbates overspending. While this concern may be somewhat muted in considering the switch from weekly to daily pay, it could be relevant for larger shifts in pay frequency, and may motivate pairing such shifts with planning aids (Augenblick et al., 2022).

This paper shows both that (a) there is demand for increased pay frequency even among weekly-paid workers, and that (b) increased pay frequency can benefit the firm by increasing labor supply through a relatively inexpensive lever.[32] As such, firms are likely to increasingly

[29] A recent literature studies worker preferences over pay timing. Parsons and Van Wesep (2013) develop a model of optimal pay by firms with present-biased workers with the result that a sophisticated worker would be willing to accept a lower average wage in exchange for a contract that better matches pay timing with regular consumption needs (because her future self would not make the savings decisions that her current self prefers). See also Casaburi and Macchiavello (2019) and Brune et al. (2021) for RCTs on less frequent or deferred pay.

[30] Kaur et al. (2022) make a similar point, "*Suppose that a worker is paid monthly and also has rent due monthly. If that worker receives a weekly payment, self-control problems may lead them to save too little and at the end of the month they may not be able to make rent payments. Weekly payment may – when combined with lumpy consumption and imperfect self-control – create more financial strain.*"

[31] The exact sign on this will depend on the shape of the effort cost function and slack in the work hours constraint. For example, for naïve present-biased drivers who cannot work any additional hours, *Instant Pay* increases the temptation to spend without being able to adjust the work margin. Whether the additional benefits of *Instant Pay* (alleviating some financial strain, allowing the worker choice over the payday) outweigh this cost is unclear in this case. Future research pairing this variation with detailed consumption and liquid asset data – paired with a structural model – would allow a more complete treatment.

[32] Our paper also relates to recent studies on the effects of allowing consumers to pay for Uber rides by cash instead of credit in Mexico and Panama (Alvarez and Argente, 2022a,b). This policy also increased driver labor supply (and entry); however, because it largely affected consumer demand, it's unclear whether

offer such pay flexibility, particularly in contexts in which the firm benefits from increased labor supply. This shift may already be observed in the marketplace. For example, the largest US employer, Walmart, allows workers to receive up to 50% of their pay early once per week via the similarly-named *Instapay*, facilitated by the financial application *Even* (Corkery, 2017). This service was utilized by over 200,000 workers just 8 months after the 2017 launch (Crosman, 2018).[33] Similar services are also being offered and utilized by employees of McDonalds, Outback Steakhouse, DoorDash, GrubHub, among others (Gee, 2017). While our paper provides causal evidence on the labor supply margin, future work may benefit from pairing an RCT with administrative data on labor supply as well as spending and borrowing (e.g., provided by an application like *Even*) to better understand the full financial impact of these policies, both in contexts in which labor may be flexibly supplied as well as those with more rigid schedules. Moreover, given previous work documenting demand for commitment in the form of deferred pay in developing country contexts (Casaburi and Macchiavello, 2019; Brune et al., 2021), there may be value to offering workers the option to commit to pay deferral as part of the menu of pay contracts.

any part of the labor supply response is due to the greater liquidity afforded by cash payments (a point unexplored by these studies). Our paper suggests that the labor supply response should be an important policy consideration in the choice of whether to allow cash payments (in markets without Instant Pay).

[33]Though the service costs $6/month to access, Walmart waived this fee as part of their COVID response (Walmart, 2023). *Even* also offers financial planning and savings tools which may help address some of the consumption misforecasting risks of higher pay frequency, as noted above, though more research is needed. As they note, "*Even gets access to Walmart employees' bank account and payroll data, as well as their work schedules. It knows the hours they're scheduled for, how much they're going to get paid and what bills they have coming up. "We built machine learning models that figure out what your bills are and they predict how much it's OK for you to spend,*" Schlossberg said."(Crosman, 2018)

References

Aksoy, C. G., J. M. Barrero, N. Bloom, S. Davis, M. Dolls, and P. Zarate (2022): "Working from Home Around the World," Tech. Rep. w30446, National Bureau of Economic Research, Cambridge, MA.

Alvarez, F. and D. Argente (2022a): "Consumer Surplus of Alternative Payment Methods," *Working Paper*.

——— (2022b): "On the Effects of the Availability of Means of Payments: The Case of Uber," *The Quarterly Journal of Economics*, 137, 1737–1789.

Angeletos, G.-M., D. Laibson, A. Repetto, J. Tobacman, and S. Weinberg (2001): "The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation," *Journal of Economic Perspectives*, 15, 47–68.

Angrist, J. D. (2001): "Estimation of Limited Dependent Variable Models With Dummy Endogenous Regressors: Simple Strategies for Empirical Practice," *Journal of Business & Economic Statistics*, 19, 2–28.

Ashraf, N., D. Karlan, and W. Yin (2006): "Tying Odysseus to the Mast: Evidence From a Commitment Savings Product in the Philippines," *The Quarterly Journal of Economics*, 121, 635–672.

Atkin, D., A. K. Khandelwal, and A. Osman (2017): "Exporting and Firm Performance: Evidence from a Randomized Experiment*," *The Quarterly Journal of Economics*, 132, 551–615.

Augenblick, N., K. Jack, S. Kaur, F. Masiye, and N. Swanson (2022): "Budget Neglect in Consumption Smoothing: A Field Experiment on Seasonal Hunger," Tech. rep., Working Paper.

Augenblick, N., M. Niederle, and C. Sprenger (2015): "Working over Time: Dynamic Inconsistency in Real Effort Tasks," *The Quarterly Journal of Economics*, 130, 1067–1115.

Augenblick, N. and M. Rabin (2019): "An Experiment on Time Preference and Misprediction in Unpleasant Tasks," *The Review of Economic Studies*, 86, 941–975.

Baker, S. R. (2018): "Debt and the Response to Household Income Shocks: Validation and Application of Linked Financial Account Data," *Journal of Political Economy*, 126, 1504–1557.

Baugh, B. and F. Correia (2022): "Does Paycheck Frequency Matter? Evidence from Micro Data," *Journal of Financial Economics*, 143, 1026–1042.

Baugh, B. and J. Wang (2022): "When Is It Hard to Make Ends Meet?" *Working Paper*.

Berk, S. H., J. Beshears, J. Garg, J. J. Choi, and D. Laibson (2023): "Employer-Based Short-Term Savings Accounts," *Working Paper*.

Beshears, J., J. J. Choi, C. Harris, D. Laibson, B. C. Madrian, and J. Sakong (2020): "Which early withdrawal penalty attracts the most deposits to a commitment savings account?" *Journal of Public Economics*, 183, 104144.

BLS (2023): "Length of pay periods in the Current Employment Statistics survey," .

Brune, L., E. Chyn, and J. Kerwin (2021): "Pay Me Later: Savings Constraints and the Demand for Deferred Payments," *American Economic Review*, 111, 2179–2212.

Cadena, B. C. and B. J. Keys (2022): "The labor market consequences of impatience," *IZA World of Labor*.

Casaburi, L. and R. Macchiavello (2019): "Demand and Supply of Infrequent Payments as a Commitment Device: Evidence from Kenya," *American Economic Review*, 109, 523–555.

Chen, J. and J. Roth (2023): "Logs with Zeros? Some Problems and Solutions," *The Quarterly Journal of Economics*, qjad054.

Chen, M. K., J. A. Chevalier, P. E. Rossi, and L. Currier (2023): "Suppliers and Demanders of Flexibility: The Demographics of Gig Work," *Working Paper*.

Chen, M. K., J. A. Chevalier, P. E. Rossi, and E. Oehlsen (2019): "The Value of Flexible Work: Evidence from Uber Drivers," *Journal of Political Economy*, 127.

Corkery, M. (2017): "Walmart Will Let Its 1.4 Million Workers Take Their Pay Before Payday," *The New York Times*.

Crosman, P. (2018): "Walmart's pay-advance app Even used by 200,000 employees," *American Banker*.

De La Rosa, W. and S. M. Tully (2022): "The Impact of Payment Frequency on Consumer Spending and Subjective Wealth Perceptions," *Journal of Consumer Research*, 48, 991–1009.

DellaVigna, S. (2009): "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 47, 315–372.

DellaVigna, S. and U. Malmendier (2006): "Paying Not to Go to the Gym," *American Economic Review*, 96, 694–719.

DellaVigna, S. and M. D. Paserman (2005): "Job Search and Impatience," *Journal of Labor Economics*, 23.

Ericson, K. M. and D. Laibson (2019): "Intertemporal choice," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 2, 1–67.

Etherington, D. (2017): "Uber's Instant Pay has cashed out $1.3B to drivers in just one year," *TechCrunch*.

Fang, H. and Y. Wang (2015): "Estimating Dynamic Discrete Choice Models with Hyperbolic Discounting, With an Application to Mammography Decisions," *International Economic Review*, 56, 565–594.

Frakes, M. D. and M. F. Wasserman (2020): "Procrastination at the Patent Office?" *Journal of Public Economics*, 183, 104140.

Gee, K. (2017): "Workers Get Faster Access to Wages With These New Apps - WSJ," *The Wall Street Journal*.

Gelman, M. (2022): "The Self-Constrained Hand-to-Mouth," *The Review of Economics and Statistics*, 104, 1096–1109.

Gelman, M., S. Kariv, M. D. Shapiro, D. Silverman, and S. Tadelis (2014): "Harnessing naturally occurring data to measure the response of spending to income," *Science*, 345, 212–215.

Goda, G. S., M. R. Levy, C. F. Manchester, A. Sojourner, and J. Tasoff (2020): "Who is a passive saver under opt-in and auto-enrollment?" *Journal of Economic Behavior & Organization*, 173, 301–321.

Hall, J. V. and A. B. Krueger (2018): "An Analysis of the Labor Market for Uber's Driver-Partners in the United States," *ILR Review*, 71, 705–732.

Hall, R. E. (1978): "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86, 971–987.

Hansen, S., P. J. Lambert, N. Bloom, S. Davis, R. Sadun, and B. Taska (2023): "Remote Work across Jobs, Companies, and Space," Tech. Rep. w31007, National Bureau of Economic Research, Cambridge, MA.

Imai, T., T. A. Rutter, and C. F. Camerer (2021): "Meta-Analysis of Present-Bias Estimation using Convex Time Budgets," *The Economic Journal*, 131, 1788–1814.

Jappelli, T. and L. Pistaferri (2010): "The Consumption Response to Income Changes," *Annual Review of Economics*, 2, 479–506.

John, A. (2020): "When Commitment Fails: Evidence from a Field Experiment," *Management Science*, 66, 503–529.

Kaur, S., M. Kremer, and S. Mullainathan (2015): "Self-Control at Work," *Journal of Political Economy*, 123, 1227–1277.

Kaur, S., S. Mullainathan, S. Oh, and F. Schilbach (2022): "Do Financial Concerns Make Workers Less Productive?" *Working Paper*.

Kuchler, T. and M. Pagel (2021): "Sticking to Your Plan: The Role of Present Bias for Credit Card Paydown," *Journal of Financial Economics*, 139, 359–388.

Kueng, L. (2018): "Excess Sensitivity of High-Income Consumers," *The Quarterly Journal of Economics*, 133, 1693–1751.

Lin, W. and J. M. Wooldridge (2019): "Testing and Correcting for Endogeneity in Nonlinear Unobserved Effects Models," in *Panel Data Econometrics*, Elsevier, 21–43.

Mas, A. and A. Pallais (2020): "Alternative Work Arrangements," *Annual Review of Economics*, 12, 631–58.

MEIER, S. AND C. SPRENGER (2010): "Present-Biased Preferences and Credit Card Borrowing," *American Economic Journal: Applied Economics*, 2, 193–210.

OLAFSSON, A. AND M. PAGEL (2018): "The Liquid Hand-to-Mouth: Evidence from Personal Finance Management Software," *The Review of Financial Studies*, 31, 4398–4446.

PARSONS, C. A. AND E. D. VAN WESEP (2013): "The Timing of Pay," *Journal of Financial Economics*, 109, 373–397.

SHAPIRO, J. M. (2005): "Is there a daily discount rate? Evidence from the food stamp nutrition cycle," *Journal of Public Economics*, 89, 303–325.

SON, H. (2019): "Uber announces deeper push into financial services with Uber Money," *CNBC Finance*.

WALMART (2023): "Even FAQ for Walmart Associates," .

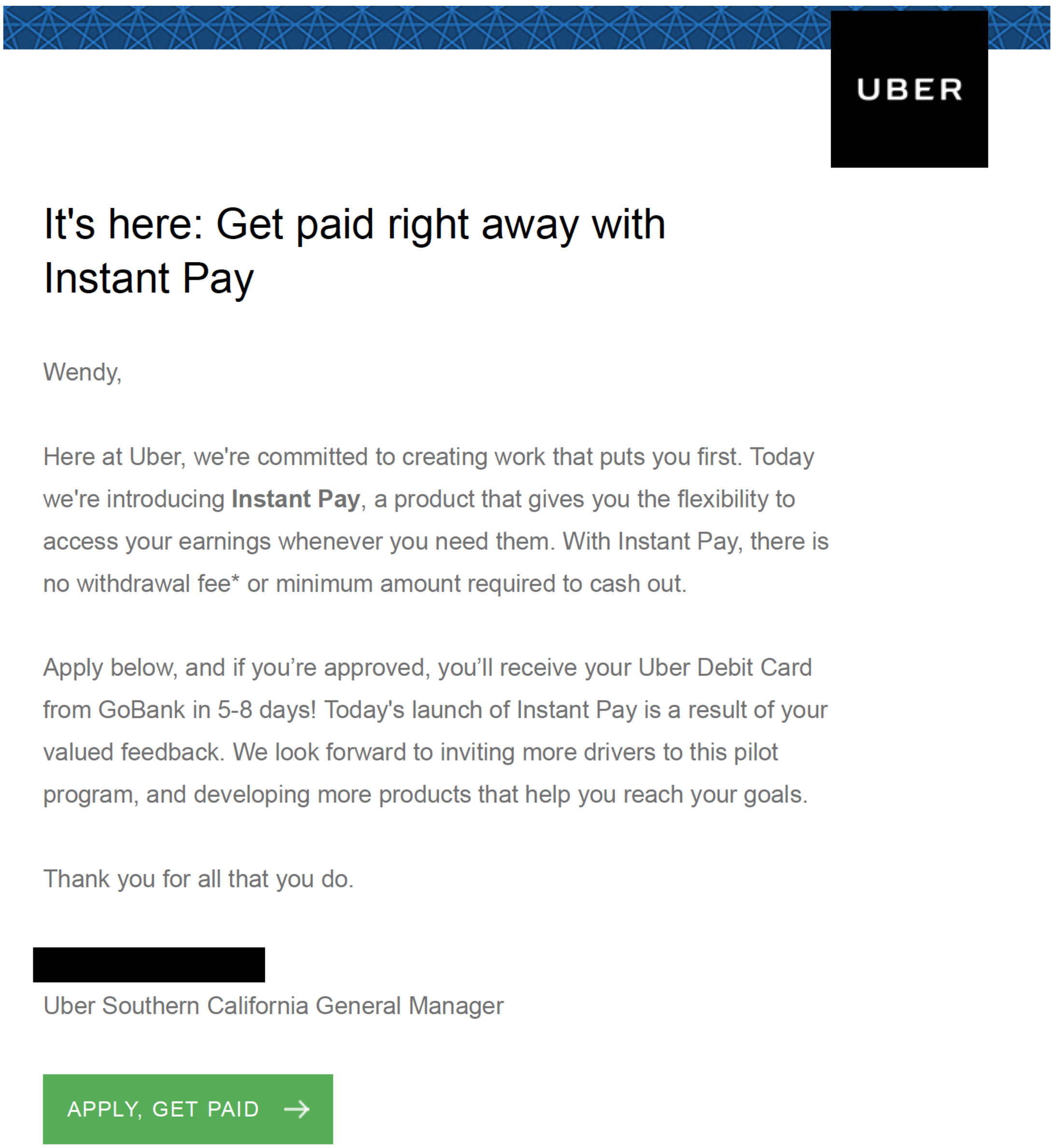
WOOLDRIDGE, J. (2010): *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

YEUNG, K. (2015): "Lyft partners with Hertz and Shell, taps Stripe to pay its 100k drivers faster," *VentureBeat*.

ZHANG, C. Y. (2022): "Consumption Responses to Pay Frequency: Evidence from 'Extra' Paychecks," *Working Paper*.

Appendix A: Figures and Tables

Figure A.1: Launch Date Email (Treatment)



UBER

It's here: Get paid right away with Instant Pay

Wendy,

Here at Uber, we're committed to creating work that puts you first. Today we're introducing **Instant Pay**, a product that gives you the flexibility to access your earnings whenever you need them. With Instant Pay, there is no withdrawal fee* or minimum amount required to cash out.

Apply below, and if you're approved, you'll receive your Uber Debit Card from GoBank in 5-8 days! Today's launch of Instant Pay is a result of your valued feedback. We look forward to inviting more drivers to this pilot program, and developing more products that help you reach your goals.

Thank you for all that you do.

[redacted]

Uber Southern California General Manager

APPLY, GET PAID →

Figure A.2: Launch Date Email (Treatment, Details)

GET THE INSTANT PAY FREQUENTLY ASKED QUESTIONS ›

Apply for your Uber Debit Card from GoBank in minutes. If approved, you'll receive your card in 7-10 business days.

Cash out your earnings instantly and easily at any time, with no minimum deposit amount. GoBank's $8.95 monthly membership fee is waived for 6 months each time you directly deposit your Uber earnings.

Use your Uber Debit Card anywhere Visa is accepted or withdraw cash at more than 42,000 in-network ATMs nationwide — for free.

Figure A.3: Launch Date Email (Optional for Control)

UBER

Coming soon: Get paid right away with Instant Pay

,

Here at Uber, we're committed to creating work that puts you first. And today, we're announcing Instant Pay, a pilot program that will allow drivers to get paid instantly, for free, with no withdrawal fees or minimum amount required to cash out.

This feature will be available to San Francisco drivers in the coming months, so please let us know if you're interested in learning more by submitting the form below.

Today's announcement of Instant Pay is a result of your valued feedback. We look forward to inviting you to this pilot program, and developing more features that help you reach your goals.

Thank you for all that you do.

[redacted]

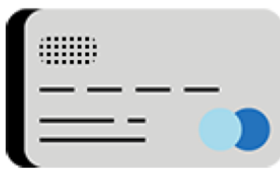
Uber Bay Area General Manager

KEEP ME UPDATED →

Figure A.4: Launch Date Email (Optional Control, Details)

Instant Pay in 3 Easy Steps

GET THE INSTANT PAY FAQS ›



We'll reach out and invite you to apply for your Uber Debit Card. If approved, you'll receive your card in 7-10 days.

Cash out your earnings instantly and easily at any time, with no minimum deposits or transaction fees.

Use your Uber Debit Card in store or withdraw cash at more than 42,000 ATMs nationwide — for free.

Table A.1: Balance Table (Pre-Experiment Driver Characteristics)

	(1)		(2)		T-test
	Treatment		Control		P-value
Variable	N	Mean/SE	N	Mean/SE	(1)-(2)
Female	107290	0.16	106623	0.16	0.39
		(0.00)		(0.00)	
Age (Years)	107237	42.99	106571	42.92	0.11
		(0.04)		(0.04)	
Median HH Income (Home Block Group)	105621	64,584.94	104872	64,671.43	0.64
		(96.42)		(98.15)	
Uber Experience (Months)	107290	20.36	106623	20.26	0.12
		(0.05)		(0.05)	
Avg Number of Shifts (Daily)	112835	0.53	112152	0.53	0.60
		(0.00)		(0.00)	
Total Number of Shifts (Baseline)	112835	7.93	112152	7.92	0.60
		(0.02)		(0.02)	
Avg Minutes Worked (Daily)	112835	111.03	112152	110.53	0.17
		(0.37)		(0.37)	
Total Minutes Worked (Baseline)	112835	1,665.39	112152	1,657.93	0.17
		(5.53)		(5.51)	
Avg Earnings (Daily)	112835	38.35	112152	38.18	0.23
		(0.13)		(0.13)	
Total Earnings (Baseline)	112835	575.29	112152	572.70	0.23
		(1.94)		(1.94)	
Total Days Worked (Baseline)	112835	5.91	112152	5.92	0.91
		(0.01)		(0.01)	
F-test of joint significance (F-stat)					1.32
F-test, number of observations					210342

Notes: The value displayed for t-tests are p-values. The value displayed for F-tests are the F-statistics. Fixed effects using variable city_x_cohort are included in all estimation regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table A.2: Summary Statistics (Outcomes; Post-(Enrollment+21))

	(1) Treatment		(2) Control		T-test Difference
Variable	N	Mean/SE	N	Mean/SE	(1)-(2)
Panel A: Take-Up (Driver-Level)					
Ever Applied	57222	0.13	56657	0.00	0.13***
		(0.00)		(0.00)	
Ever Registered	57222	0.12	56657	0.00	0.12***
		(0.00)		(0.00)	
Ever Activated Card	57222	0.06	56657	0.00	0.06***
		(0.00)		(0.00)	
Ever Withdrew Cash	57222	0.04	56657	0.00	0.04***
		(0.00)		(0.00)	
Panel B: Outcomes (Driver-Day Level)					
Total Minutes Driven (Daily)	671947	139.54	660875	136.53	3.01***
		(0.25)		(0.25)	
Total Earnings (Daily)	671947	52.13	660875	51.00	1.13***
		(0.10)		(0.10)	
Number of Shifts (Daily)	671947	0.56	660875	0.55	0.01***
		(0.00)		(0.00)	

Notes: The value displayed for t-tests are the differences in the means across the groups. The value displayed for F-tests are the F-statistics. City X Cohort fixed effects are included in all estimation regressions. Observations are at the driver-day level and include all days between the (driver-specific) experiment start date until April 16. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels.

Table A.3: Robustness: Time-Varying Take-Up TOT

	Poisson			Log(1+Y)			Levels		
	Minutes	Dollars	Sessions	Minutes	Dollars	Sessions	Minutes	Dollars	Sessions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: IV (TOT, Take-Up = Applied)									
Applied	0.222**	0.214**	0.095	0.233	0.192	0.033	27.984**	9.314**	0.061
	(0.101)	(0.103)	(0.079)	(0.158)	(0.130)	(0.026)	(12.491)	(4.411)	(0.046)
N	13,403,588	13,403,370	13,403,588	13,498,680	13,498,680	13,498,680	13,498,680	13,498,680	13,498,680
DepVarMean	121.524	41.999	0.580	2.281	1.845	0.355	120.586	41.674	0.575
Panel B: IV (TOT, Take-Up = Activated)									
Activated	0.821**	0.782**	0.354	0.874	0.722	0.124	105.175**	35.007**	0.230
	(0.381)	(0.387)	(0.297)	(0.594)	(0.488)	(0.096)	(46.953)	(16.581)	(0.173)
N	13,403,588	13,403,370	13,403,588	13,498,680	13,498,680	13,498,680	13,498,680	13,498,680	13,498,680
DepVarMean	121.524	41.999	0.580	2.281	1.845	0.355	120.586	41.674	0.575

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table repeats Table 1 Panels B and C, but instead of the endogenous variable being the interaction between a binary "ever take-up" variable and a binary "post" (equal to 1 for dates 21 days after the driver-specific enrollment date) variable, it is instead a single driver-specific take-up variable that equals 0 until the driver takes up and equals 1 after (and always equals 0 for those in the control group or those who never take-up in the treatment group). Columns 1 to 3 use a control function approach where the first stage is estimated via OLS (*reghdfe*), residuals are included as a control in the second-stage Poisson regression (*ppmlhdfe*), and the delta method is used to compute standard errors (using *margins, eydx*). DepVarMean is the mean in the control group in the pre-period. Standard errors are clustered by driver.

Table A.4: Robustness: Just post-treatment observations, no fixed effects, & Poisson TOT with *ivpoisson*

	Poisson			Log(1+Y)			Levels		
	Minutes	Dollars	Sessions	Minutes	Dollars	Sessions	Minutes	Dollars	Sessions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Intent-to-Treat (ITT)									
Treatment	0.021***	0.022***	0.015***	0.034***	0.028***	0.005***	3.000***	1.118***	0.009***
	(0.006)	(0.006)	(0.005)	(0.011)	(0.009)	(0.002)	(0.856)	(0.323)	(0.003)
N	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050
DepVarMean	138.057	50.949	0.623	2.494	2.053	0.383	138.057	50.949	0.623
Panel B: IV (TOT, Take-Up = Applied)									
Ever Applied	0.149***	0.150***	0.105***	0.254***	0.210***	0.038***	22.122***	8.245***	0.069***
	(0.040)	(0.041)	(0.034)	(0.084)	(0.070)	(0.013)	(6.315)	(2.381)	(0.023)
N	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050
DepVarMean	138.057	50.949	0.623	2.494	2.053	0.383	138.057	50.949	0.623
Panel C: IV (TOT, Take-Up = Activated)									
Ever Activated Card	0.293***	0.295***	0.211***	0.540***	0.445***	0.081***	46.964***	17.505***	0.146***
	(0.073)	(0.075)	(0.065)	(0.178)	(0.148)	(0.028)	(13.407)	(5.056)	(0.050)
N	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050	1,333,050
DepVarMean	138.057	50.949	0.623	2.494	2.053	0.383	138.057	50.949	0.623

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes:This table repeats Table 1 but limits the sample to "post-treatment" observations (i.e. 21 days afterrr the driver-specific enrollment date) and omits fixed effects from all specifications. Columns 1-3 of Panels B and C are estimated using the *ivpoisson gmm* command, i.e. a two-step generalized method of moments estimator.

Table A.5: Experimental Roll-Out and Summary Statistics by City

City	Total Drivers	Modal Start Date	Pr(Started on Mode)	Pr(Treated)	Avg Days Post (3 weeks)	Avg Days Post (0 weeks)
Minneapolis - St. Paul, MN	2,825	02mar2016	0.68	0.57	17.61	36.94
Madison, WI	631	08mar2016	0.72	0.51	13.30	32.58
San Diego, CA	5,845	08mar2016	0.70	0.50	13.14	32.38
Seattle, WA	5,044	08mar2016	0.75	0.50	13.88	33.43
Boston, MA	13,474	10mar2016	0.77	0.51	12.34	31.92
San Francisco, CA	20,286	14mar2016	0.75	0.50	8.80	28.13
Washington, D.C.	17,754	14mar2016	0.75	0.50	8.83	28.25
Columbus, OH	2,264	14mar2016	0.72	0.52	8.44	27.31
Austin, TX	5,334	14mar2016	0.67	0.49	8.25	27.16
Chicago, IL	22,938	14mar2016	0.76	0.50	8.75	28.02
Nashville, TN	3,218	14mar2016	0.70	0.50	8.31	27.17
Dallas, TX	8,296	14mar2016	0.66	0.49	7.93	26.45
Los Angeles, CA	26,883	18mar2016	0.76	0.50	5.60	24.85
Orlando, FL	4,826	30mar2016	0.83	0.49	0.00	14.58
Denver, CO	5,524	30mar2016	0.82	0.49	0.00	14.53
Houston, TX	5,140	30mar2016	0.80	0.50	0.00	14.43
Phoenix, AZ	5,735	30mar2016	0.83	0.51	0.00	14.52
Orange County, CA	4,305	30mar2016	0.82	0.48	0.00	14.46
Charlotte, NC	2,399	04apr2016	0.86	0.50	0.00	10.07
Portland, (OR)	2,113	04apr2016	0.89	0.50	0.00	10.33
Tampa Bay, FP	3,988	04apr2016	0.86	0.50	0.00	10.09
New Orleans, LA	2,693	04apr2016	0.87	0.50	0.00	10.16
New Jersey	10,053	04apr2016	0.88	0.51	0.00	10.23
Baltimore, MD	2,783	04apr2016	0.86	0.48	0.00	10.16
Las Vegas, NV	4,349	04apr2016	0.86	0.50	0.00	10.16
Raleigh-Durham, NC	2,146	04apr2016	0.86	0.51	0.00	10.12
Connecticut	2,903	04apr2016	0.87	0.50	0.00	10.17
Pittsburgh, PA	2,139	04apr2016	0.88	0.51	0.00	10.03
Detroit, MI	2,528	11apr2016	0.94	0.49	0.00	3.83
Hampton Roads	1,408	11apr2016	0.93	0.52	0.00	3.78
Rhode Island	1,079	11apr2016	0.95	0.50	0.00	3.84
Richmond, VA	1,057	11apr2016	0.94	0.50	0.00	3.82
Milwaukee, WI	1,306	11apr2016	0.95	0.48	0.00	3.84
Kansas City, (?)	1,260	11apr2016	0.94	0.50	0.00	3.84
St Louis, MO	1,141	11apr2016	0.91	0.46	0.00	3.71
Charleston, SC	972	11apr2016	0.95	0.51	0.00	3.84
Indianapolis, IN	1,852	11apr2016	0.92	0.51	0.00	3.75
Cincinnati, OH	1,376	11apr2016	0.95	0.51	0.00	3.85
Palm Springs, CA	1,079	11apr2016	0.95	0.50	0.00	3.84
San Antonio, TX	804	11apr2016	0.92	0.52	0.00	3.75
Sacramento, CA	1,109	11apr2016	0.94	0.54	0.00	3.82
Salt Lake City, UT	1,149	11apr2016	0.93	0.49	0.00	3.80
Honolulu, HI	873	11apr2016	0.94	0.53	0.00	3.83
Cleveland, OH	1,592	11apr2016	0.95	0.49	0.00	3.85
Tucson, AZ	848	11apr2016	0.95	0.46	0.00	3.83
Philadelphia, PA	5,855	12apr2016	0.95	0.50	0.00	2.87

Notes: Observations are at the driver-level. Average Days Post (0 Weeks) is computed by first calculating the number of days between the last day observed in our data (April 16) and the driver's entry into the experiment (minus 1), and then taking the driver-weighted average for this value within each city. Average Days Post (3 Weeks) does the same, but subtracts 21 and censors at zero before taking the average. Note that this table omits the roughly 1% drivers for whom a city could not be identified.