NBER WORKING PAPER SERIES

FLEXIBLE PAY AND LABOR SUPPLY: EVIDENCE FROM UBER'S INSTANT PAY

M. Keith Chen
Katherine Feinerman
Kareem Haggag

Working Paper 33177
http://www.nber.org/papers/w33177

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2024

We thank Linda Babcock, Hal Hershfield, Sam Hirshman, Devin Pope, Dmitry Taubinsky, Oleg Urminsky, seminar participants at Carnegie Mellon University, UCLA, Johns Hopkins University, and audiences at the Society for Judgment and Decision Making conference, the Behavioral Decision Research in Management conference, the Los Angeles Experiments Workshop, and the Stanford Institute for Theoretical Economics Experimental Economics session for helpful comments and suggestions. Chen is a former employee of Uber, but no longer has a material financial relationship with the company. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2024 by M. Keith Chen, Katherine Feinerman, and Kareem Haggag. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Flexible Pay and Labor Supply: Evidence from Uber's Instant Pay
M. Keith Chen, Katherine Feinerman, and Kareem Haggag
NBER Working Paper No. 33177
November 2024
JEL No. J0, J20, J3, R4

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 drivers' work time. Furthermore, consistent with present bias, the response is significantly higher when drivers are further away from the end of their counterfactual weekly pay cycle. We discuss welfare and broader implications in contexts in which workers have the ability to flexibly supply labor.

M. Keith Chen
Anderson School of Management
University of California, Los Angeles
Entrepreneurs Hall
110 Westwood Plaza
Los Angeles, CA 90095
keith.chen@anderson.ucla.edu

Katherine Feinerman
Anderson School of Management
University of California, Los Angeles
Entrepreneurs Hall
110 Westwood Plaza
Los Angeles, CA 90095
katherine.feinerman.phd@anderson.ucla.edu

Kareem Haggag
Anderson School of Management
University of California, Los Angeles
Entrepreneurs Hall
110 Westwood Plaza
Los Angeles, CA 90095-1481
and NBER
kareem.haggag@anderson.ucla.edu

Prior to 2020, 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. 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*, an earned wage access (EWA) product that allowed drivers to access their accumulated earnings on-demand. This feature 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 in the absence of sharply-binding liquidity constraints (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 1.4% (work minutes) and 1.5% (earnings in dollars), both statistically significant at the 5% level. Depending on the measure of take-up used, these results translate to treatment-on-the-treated increases in work time of 10.1% (or 21.0%) and in earnings of 10.8% (or 22.8%). Taking the largest of these estimates, we benchmark against a labor supply elasticity estimated by Chen et al. (2019) using randomized wage variation in roughly the same time period and set of drivers. This benchmarking suggests that the effect of Instant Pay among those who activated their cards is roughly equivalent to an 11% increase in wages. Several mechanisms could explain this overall rise in labor supply, including sharply binding liquidity constraints and consumption smoothing motives, motivational effects of tying feedback to immediate rewards, and certain forms of present-focused preferences. While each of these mechanisms could be operative for different drivers, we next turn to a set of patterns that implicate the last of these as particularly important.

We present four pieces of evidence pointing toward present bias as an important mechanism underlying the estimated labor supply responses. First, while there is ample evidence that many rideshare drivers are liquidity-constrained (Koustas, 2018), we do not find significantly larger treatment effects among drivers who reside in less affluent Census block groups. Moreover, the magnitude

of the effects when benchmarked against wage elasticities would imply that, for liquidity constraints alone to account for this effect, drivers would need to act as though they do not have access to credit at less than an 11% weekly interest rate – in conjunction with the fact that all drivers must have a bank account, this explanation seems unlikely for many drivers. Second, we develop a proxy for present bias in the spirit of Kaur et al. (2015) (i.e., whether drivers exhibit a payday effect in labor supply during the pre-experiment period) and show this indicator strongly predicts take-up – having above-average payday effects in the baseline predicts a 19% increase in take-up (3pp increase), roughly 2.5 times the size of the relationship with being below the median proxied household income. Third, we examine heterogeneous treatment effects by day-of-the-week and find larger effects on days when the counterfactual pay delay would be larger, with a 1.9% ITT on Monday (9 days from payday) and an insignificant -0.4% ITT on Sunday (3 days from payday). We use data from another experiment to argue this pattern does not simply reflect variation by day-of-the-week in labor supply elasticities more generally. Finally, we examine day-of-the-week patterns in utilization of Instant Pay (i.e., "cash outs") and find that it does not mimic the one found for labor supply responses. This final result casts doubt on both consumption smoothing motives *and* standard accounts of present bias over *consumption*. Instead, these results can be reconciled by present bias over *access to cash*, consistent with other work showing discounting over cash itself (e.g., Andreoni et al. (2018)).

Broadly, our paper contributes to the literature documenting evidence of 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 they reverse this preference when the choice is in the present; a phenomenon reflected in hyperbolic and quasi-hyperbolic discounting models.[1] 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). 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). 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 experiment with data entry workers and find that they demonstrate strong payday effects and demand dominated (commitment) contracts to solve the self-control problem in effort provision – both results consistent with present bias. Our work extends on these by studying an increasingly popular policy that presents another way of addressing the self-control problem – rather than introducing new upfront

[1] 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.

(conditional) costs to align short-term and long-term preferences, pay flexibility instead brings delayed benefits into the present. This inquiry 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.

We also directly contribute to an emerging literature on earned and early wage access (EWA) products that allow workers to receive their earnings ahead of payday, usually with a fee. Earned wage access products are offered in partnership with employers and often integrated with existing payroll systems, while early wage access products are offered directly to consumers (workers). The market for these products has grown considerably in recent years. For example, the number of transactions processed by earned wage access products grew over 90% from 2021 to 2022 (Consumer Finance Protection Bureau, 2024). Early analysis by Baker and Kumar (2018) studying employee turnover across six U.S. firms partnered with the EWA product PayActiv estimated a 19% reduction in turnover among employees who were active EWA users compared with employees who were enrolled but inactive. This result is consistent with work by Murillo et al. (2023) finding that EWA users at two large Mexican firms were 9-12% less likely to leave their firms by the next pay cycle compared to non-users. On the consumption side, De La Rosa and Tully (2022) found that lab participants given the ability to access their pay daily in a life simulation task both chose to get paid more frequently and increased their overall spending relative to participants constrained to a weekly pay schedule. Our main contribution is to study the causal effects of the flexible pay provided through an EWA on labor outcomes in a field setting with flexible work (i.e., where total hours and earnings are adjustable). By benchmarking our treatment effects against findings by Chen et al. (2019) on labor supply responses to randomized wages, we are also able to price our treatment effect estimates. Moreover, our analysis of Instant Pay utilization patterns suggests that drivers respond to the option value of accessing their pay on-demand, showing how workers may benefit from EWA even if they do not withdraw their earnings early or often.

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);

Olafsson and Pagel (2018); Gelman (2022)). 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.[2] We extend this line by looking at labor supply rather than spending and 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 the EWA market continues to grow and more companies begin to offer pay flexibility, the labor supply implications will also need to be considered.

1 Background and Experimental Design

1.1 Study Setting

This study took place in partnership with Uber, a popular peer-to-peer 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.[3] The majority of 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), driving is often complementary to other activities such as caregiving, employment, and school attendance. Over an 8-month period close to our study period, Chen et al. (2023) notes that over one million individuals worked on UberX, the company's flagship service.

[2] 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 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).

[3] 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. 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.

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 required to link a checking account to their Uber profile to receive their weekly payments via direct deposit. 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 earned wage access product allowed drivers to flexibly "cash out" their accumulated earnings prior to the usual payday. Specifically, drivers could press a button in the Uber app to initiate a transfer of the entirety of their current earnings balance that would arrive minutes later in a separate debit card. Drivers could complete up to 5 of these cash-outs per day. Importantly, drivers could not cash out a partial amount of their accumulated earnings, and if a driver did not cash out on or prior to Sunday, those accumulated earnings would be sent for direct deposit (and thus inaccessible until Wednesday). This Sunday deadline is important for interpreting patterns in cash out behavior.

During the experimental roll-out, drivers in the treatment group who signed up for Instant Pay would receive an Uber Debit Card from the implementation partner GoBank. The debit card had no fees for "cash-outs" (transfers from the Uber app to the GoBank account), no withdrawal fees from their network of 42,000 ATMs, and no overdraft/non-sufficient-funds/penalty fees. 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 underlying GoBank account had limited features compared to other complete checking account products, but it did allow for ACH transfers to accounts at other U.S. financial institutions. Meanwhile, Uber's primary rival ride-sharing platform Lyft previously piloted a similar EWA called "Express Pay" in 2015 in conjunction with Stripe. Lyft drivers could use their existing personal debit cards with Express Pay but had to earn at least $50 before cashing out and pay a $0.50 fee per cash-out (Lyft, 2016a).

Uber's EWA continued to pick up usage after the national launch in May, and in August drivers were able to use most Visa, Mastercard, or Discover debit cards with Instant Pay for a $0.50 fee per cash out (Koren, 2016). 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). Starting in 2023, GoBank transitioned all existing Uber Debit cards accounts to GoBank debit cards accounts with an $0.85 fee for Instant Pay cash-outs and removed the ability to use ACH transfers (GoBank, 2022). Drivers using Visa, Mastercard, and Discover debit cards with Instant Pay are now subject to a $1.25 fee per cash out, but some high performing drivers can take advantage of free instant cash-outs after each trip using the Uber Pro debit Mastercard and checking account powered by Branch (Uber, n.d.).

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 roughly 80% of drivers were enrolled into the study on their city's 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. Appendix Figure A.1 displays one such example noting that it is a pilot program and providing a link to the application. The second page of the email included additional details, including a forecast for the time to receive the card if approved ("7-10 days"). While all cities contacted drivers in the treatment group, cities had discretion over whether to email drivers in the control group. Appendix Figures A.3 and A.4 show an example of the email sent to the control group in the Bay Area market, with the tagline note "Coming soon," and a "Keep Me Updated" button instead of an application button. 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.

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. 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", randomized to treatment or control, and contacted accordingly. The randomization is thus essentially stratified by date within each city, so in our analysis we include cohort fixed effects, where cohort is defined as the date of enrollment interacted with the driver's city.

The experiment ran for two months, beginning in Minneapolis-St.Paul, MN on March 2, 2016 and gradually expanding across markets. While the initial launch plan intended to randomize across 168 markets in total, Uber pulled the data on April 16 after just 46 markets had begun implementation. Deeming the program an operational success, Instant Pay 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. Implementation issues in Minneapolis-St.Paul lead to imbalances in baseline driving activity across treatment and control groups, so we exclude this market's drivers out of caution.[4] Thus, the first and last implementing

[4] Appendix Table A.6 reproduces our primary regression results including drivers assigned to Minneapolis-St.Paul, MN and shows an increase in the magnitudes and statistical significance of our treatment effects across all outcomes. The failed random assignment in this city lead to meaningful baseline differences across the two groups. Over the two weeks prior to the experiment, the average daily driving minutes was 44 minutes higher for treatment than control drivers, and the difference in average daily earnings was $15.46. After careful consideration, we chose to drop this city

cities in our sample are Madison, WI and Philadelphia, PA which launched on March 8 and April 12, 2016, respectively.

Appendix Table A.8 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 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 and three weeks thereafter. The last point highlights an issue discussed further in Section 2.3 – how to define the post-period. 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 and effectively reduces our identifying variation to the 12 markets that enrolled before this cutoff, beginning with Madison, WI on March 8 and ending with Los Angeles, CA on March 18; thus, when we change our definition of "post" we expand or contract the set of identifying cities.

2 Data

Our data comes from Uber records on drivers enrolled in the study over February to April 2016. We observe a driver's experimental condition and study enrollment date and supplement this information with basic demographic characteristics, such as age, sex, and Uber driving tenure. We also include the median household income of their Census Block Group using the 2016 American Community Survey 5-year estimates. We then combine information on active minutes worked and payouts associated with trips, Uber Debit sign-ups, and Instant Pay cash-outs to construct our main measures of labor supply and program take-up.

2.1 Data Construction and Key Variables

Given the staggered roll-out of the experiment across city markets, we include fixed effects for the interaction of cohort (city-by-enrollment-date) and calendar date. To do so, we impute the enrolling city for a majority of drivers. We pull possible city candidates from one file on driver demographics and another detailing the cities in which drivers were working at the time. While Uber drivers can only be actively registered with one market, drivers can work anywhere within the same state and change which local market they are registered with. By comparing each candidate city's modal experiment start date with a driver's enrollment date, we assign a city to a driver if a candidate city's common experiment start date is prior to and the closest to the driver's start date. When we cannot assign a city based on this criteria, we place drivers in a catch-all "Synthetic City" to avoid dropping drivers. There are also handful of unnamed cities that cannot be matched to cities in the launch plan and we do not modify assignments for drivers linked with these unknown cities. Fewer

and its drivers to avoid biasing our treatment effects and potentially overstating the causal effects of the program.

than 1% of drivers in our final sample are assigned either the synthetic city or an unnamed city, and dropping these drivers has no meaningful effect on our estimates.

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 actively completing UberX, Uber Pool, and Uber Eats trips. Sessions may include breaks in which we do not observe an active trip, and we use dormant periods exceeding 2 hours 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). Defining session earnings is more straightforward and calculated by summing net earnings over the corresponding trips where net earnings is the trip's ride fare less Uber's commission or a fixed $3.00 per Uber Eats trip.

To construct our daily-level outcomes, we analogously aggregate session minutes and session earnings within a specific date for our daily session minutes and earnings outcomes. The daily session count is the number of sessions on a given calendar date. 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 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 by the trip timestamps, 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. Note that the pulled trips data has coverage issues after April 16: there are no trips recorded between April 17 and 25, and the data remains sparse through the rest of the month. Due to the missing dates, we cut off the sample on April 16. However, in Appendix B we discuss a way to extend the sample for the final two weeks of the experiment by using data from Chen et al. (2019).

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 by pushing a button in the driver app to transfer all accumulated earnings to the GoBank account. 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. To examine take-up over time, we create time-varying measures of take-up for the earliest date drivers reach each of these phases; however, for our take-up analysis and for estimating the

treatment-on-the-treated, we use time-invariant measure of take-up reflecting whether a driver ever reaches that stage on or before April 16.

We also count the daily number of Instant Pay cash-outs 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 an ACH transfer from GoBank to another financial institution. 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.

2.2 Balance Analysis and Summary Statistics

Our final dataset is a balanced panel of 222,162 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 demographic variables (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.[5] The same is true for the measures of labor supply in the two weeks prior to the experiment's first planned 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=1.02). This specification includes cohort fixed effects to match our regression analysis, but omitting the fixed effects produces a similar F-statistic (1.08).

The table shows that 16% of drivers are female, the average age is 43, and that the typical driver lives in a Census Block Group with a median income of $65,000 dollars. Drivers in our sample average one session of work every two days, for 111 minutes and $38 per day. This activity aggregates to average totals of roughly 8 shifts, 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 earnings requires a number of steps, a key issue is how to define when the *post-treatment* period starts. Defining *post* too early may attenuate treatment effects by identifying off days in which the driver lacks the ability to use the service. This is because there are access delays due to at least three margins: (a) The driver needs to apply (requiring first reading the email), (b) GoBank has to review the application to decide if the driver will be approved, (c) After approval, there are delays in receiving the debit card. On the other hand, defining the *post* period too late risks inflating standard errors as there will be

[5] 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.

fewer days identifying the treatment effect. We thus face a bias-variance trade-off in this decision.[6] Because our post-implementation observation window is short (roughly 32 days on average) this issue is of first-order importance.[7]

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 and activation, appears to steadily increase between 10 and 28 days after enrollment. There is no clear cut-off point across all four measures of take-up, but 21 days appears to somewhat balance the concerns, 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.

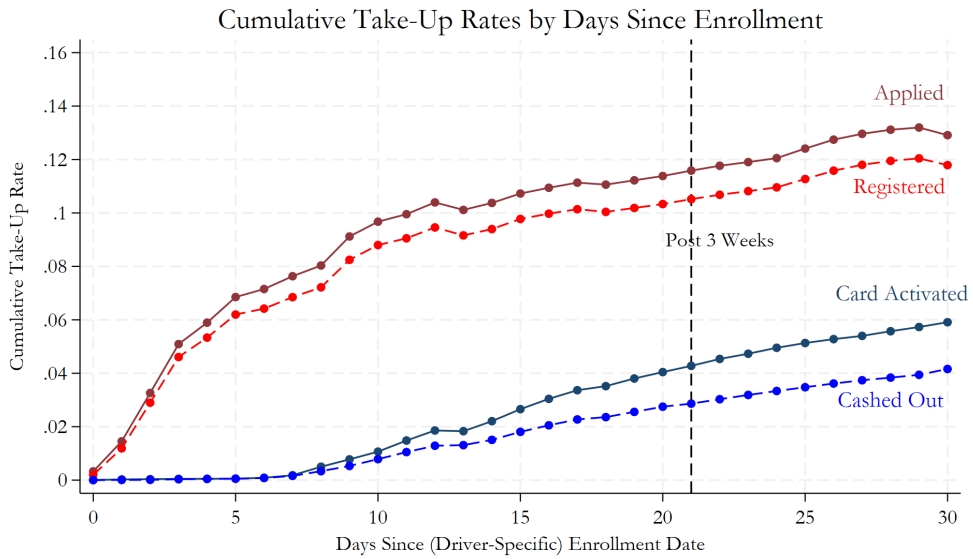
Panel A of Appendix Table A.2 shows take-up rates by treatment group among the 55,841 drivers who have at least one observation 21 days post after their experiment enrollment. The take-up rates in this group are relevant for our primary estimation sample and for comparing the intent-to-treat and treatment-on-the-treated estimates. 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%. The table also shows that there is only one-sided non-compliance, as the take-up rates are zero in the control group.

Finally, 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, capturing the intent-to-treat (ITT) effects. Using this driver-day level sample, we see that work time increases by 1.89 minutes (roughly 1.4% of the control group mean), earnings by \$0.81 (or 1.6%), and session count by less than 0.01 (less than 1.6%). In Section 4 we estimate these treatment effects using the full panel structure.

[6] 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, counteracting the possible attenuation from defining the *Post* window earlier. Conversely, if drivers gradually ramp up their usage of the card, this would producing larger coefficients the further out the *Post* window date is defined.

[7] Limiting to the 55,841 treated drivers who were enrolled at least 22 days prior to April 16, we observe them roughly 32 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 12 known cities with experiment start dates at least three weeks before the April 16 cutoff and a handful of drivers in the catch-all synthetic city (498). This group of early cities consists of 111,503 drivers with 55,841 assigned to treatment and 55,662 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 8 and another started on March 10, Day 1 would correspond to March 9 and 11 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).

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 panel, and thus since implementation is staggered, the number of pre- vs. post-enrollment observations varies within and between cities. For example, implementation began in Houston on March 30, so we have 6 weeks of pre-enrollment data and roughly 2 weeks of post-enrollment data for that city. However, 20% of Houston drivers were enrolled after that date and so have fewer days post-enrollment. As discussed in Section 1, we include fixed effects for the cohort of enrollment (i.e., each distinct randomization) interacted with date fixed effects to isolate treatment versus control differences in the post-treatment period, while accounting for dynamics common to both groups within a cohort over time. Pre-period observations are included for their potential to increase precision and not for causal identification (as treatment is randomized), but we show in Section 4.3.3 that our estimates are quite similar if we limit to post-treatment observations and omit fixed effects. We define the post-*treatment* period as 21 days after enrollment to account for delayed sign-up and mailing of debit cards and show robustness to this choice in Section 4.3.1.

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

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

where y_{izj} is a measure of daily labor supply (e.g., work minutes or earnings) 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 *offered* Instant Pay.

We examine outcomes in terms of levels, as well as with specifications that allow interpretation in percentages in order to facilitate comparisons to other labor supply elasticities. For the latter, since there are many potential zero-valued outcome observations we use Poisson regression due to its favorable econometric properties (Chen and Roth, 2023; Wooldridge, 2010).

As discussed in Section 2.3, take-up is somewhat mechanically depressed due to the short observation window and implementation challenges. Since we're interested in the labor response of drivers actually capable of taking advantage of the greater pay flexibility, we also present treatment-on-the-treated (TOT) estimates, scaling 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:

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

where $Takeup_{iz}$ is an indicator that takes value 1 if a driver *ever* took up the *Instant Pay* at any point prior to April 16. Of course, take-up is not random and may be correlated with unobservables, 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).

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 take-up as measured by either the application in case 1 or the activation of the card in case 2. Again, our preferred estimates use Poisson regressions, and we adopt a control function approach and use the delta method to calculate standard errors as suggested by Lin and Wooldridge (2019). We estimate the first stage via ordinary least squares regression (OLS) and add residuals from the first stage as a control in the second-stage Poisson regression. In Appendix Table A.4 we re-estimate this specification using just the the post-treatment observations (and omitting fixed effects) and obtain qualitatively similar estimates.

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. Then we turn to the treatment-on-the-treated (TOT), and we show results using two possible measures of take-up: ever applying for Instant Pay and ever activating an Uber debit card by the end of our observation window. We show all results using Poisson and OLS regressions.

4.1 Intent-to-Treat (ITT)

Table 1, Panel A reports the intent-to-treat (ITT) estimates of Equation 1. Columns 1-4 report our preferred estimates using Poisson regression, and Columns 7-9 report OLS specifications. Our first measure of labor supply is the number of session minutes worked over a day. Column 1 shows that the offer of Instant Pay increased daily labor supply by 1.4% in terms of minutes worked, and Column 2 reflects a similar 1.5% increase increase in earnings. Note that work time and earnings are not constrained to move together due to drivers' flexible choice of when to supply labor and Uber's dynamic pricing algorithm. For example, if drivers already worked at times of day with the highest earnings opportunities, then the marginal product of labor would be diminishing. However, we do see similar increases in both outcomes that are statistically significant at the 5% level, suggesting that drivers may have increased work time without reducing their average productivity. In Columns 3 and 4 we disaggregate the increase in work time into two margins: whether the driver works at all on a given day (Column 3) and the total time worked on a day conditional on working at all (Column 4). Of course, since the decision to work is downstream of treatment, Column 4 which conditions on this variable should be interpreted with caution as it may suffer from post-treatment bias (Montgomery et al., 2018). Column 3 shows that the probability of working any sessions increases by 0.6%, and Column 4 shows that work time conditional on working increases by 0.8% with only the latter being significant (at the 5% level).

Columns 5 to 8 report results of OLS regressions in levels. Minutes worked increased by 1.91 minutes, and earnings increased by $0.78. The daily probability of working increased by 0.3pp, and work minutes conditional on working increased by 2.38 minutes. Dividing these point estimates by the pre-period mean of the dependent variable in the control group implies these outcomes increased by 1.6%, 1.8%, 0.7%, and 0.8%, respectively. These results are all quite similar to the Poisson regression results and demonstrate that the ITT estimates are robust to the estimation method. Appendix Table A.7 reports estimates using a log(1+Y) transformation that produces similar (slightly larger) estimates.

Table 1: Main Results

	Poisson				OLS (Levels)			
	Minutes	Dollars	Work	Minutes | Work	Minutes	Dollars	Work	Minutes | Work
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Intent-to-Treat (ITT)								
Treat X Post	0.014**	0.015**	0.006	0.008**	1.905**	0.757**	0.003	2.375**
	(0.006)	(0.006)	(0.004)	(0.004)	(0.858)	(0.323)	(0.002)	(1.084)
N	13,242,030	13,241,824	13,242,030	5,795,535	13,329,180	13,329,180	13,329,180	5,795,535
DepVarMean	121.752	42.082	0.435	280.039	120.885	41.781	0.432	280.039
Panel B: IV (TOT, Take-Up = Applied)								
Applied X Post	0.101**	0.108**	0.040	0.060**	14.159**	5.622**	0.019	16.893**
	(0.045)	(0.046)	(0.032)	(0.027)	(6.379)	(2.400)	(0.015)	(7.715)
N	13,242,030	13,241,824	13,242,030	5,795,535	13,329,180	13,329,180	13,329,180	5,795,535
DepVarMean	121.752	42.082	0.435	1.335	120.885	41.781	0.432	280.039
Panel C: IV (TOT, Take-Up = Activated)								
Activated X Post	0.210**	0.228**	0.078	0.130**	30.254**	12.013**	0.042	32.947**
	(0.097)	(0.098)	(0.067)	(0.057)	(13.627)	(5.127)	(0.032)	(15.059)
N	13,242,030	13,241,824	13,242,030	5,795,535	13,329,180	13,329,180	13,329,180	5,795,535
DepVarMean	121.752	42.082	0.435	280.039	120.885	41.781	0.432	280.039

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

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 1-4 report Poisson regressions, while columns 5-8 report OLS of the untransformed outcomes. Minutes and Dollars are the daily totals for minutes worked and dollars earned. Work is an indicator for having driven any sessions on a particular day. 'Minutes | Work" represents regressions of daily minutes worked restricted to days when drivers worked any shifts. 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 4 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 of Equation 2, where the randomized offer of Treatment instruments for Take-Up. Panel B focuses on the loosest measure of take-up, which is whether the driver ever submitted an application. In Panel C, the measure of take-up is whether the driver ever activated their Uber debit card, which required first applying for the card and registering for GoBank online. As shown in Appendix Table A.2, roughly 13% of drivers in the sample ever applied and roughly 6% of drivers activated an Uber debit card before April 16. Among drivers who ever activated a card, it took roughly 7.1 days on average for drivers to apply for the card after receiving the first enrollment email and roughly 18.2 days to activate their card.

It is not obvious which TOT measure is most relevant *a priori*. On one hand, cashing out seems a natural candidate as it reflects the most direct measure of using the EWA. After all, drivers who do not cash out have not had an actual change in their pay flexibility. However, it is plausible

that drivers shift their labor supply after activating their card – as it is 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. 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 to apply for the program, and drivers may adjust their behavior in anticipation before receiving the card.

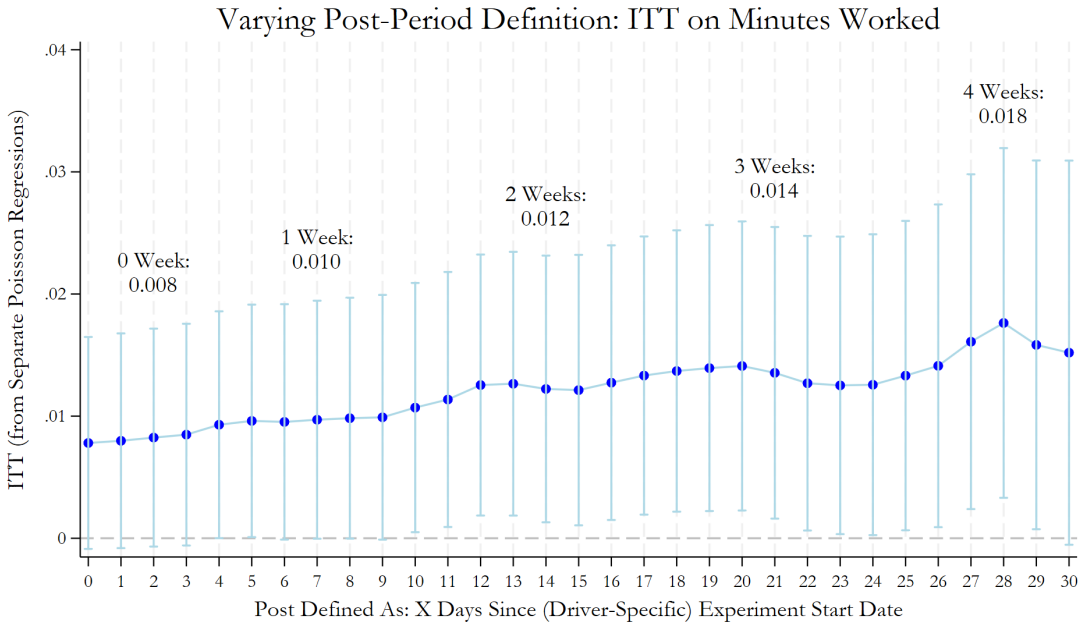
Columns 1 to 4 of Panel B, Table 1 using *Applied* as a measure of take-up – finds TOT effects of daily work time increasing by 10.1%, earnings by 10.8%, the daily probability of working any sessions by 4%, and daily work time conditional on working by 6.0%. Except for the effect on the probability of working, these TOT estimates are statistically significant at the 5% level. In terms of levels estimated via OLS (Columns 5 to 8) and percentages relative to the control group mean, these are increases in worktime of 14.2 minutes (12%), earnings of $5.62 (13%), probability of working of 1.9pp (4%), and worktime on days drivers work of 16.9 minutes (6%). Using card *Activation* as the measure of take-up, Panel C shows work time increasing by 21.0%, earnings by 22.8%, probability of working by 7.8%, and worktime on days drivers work by 13.0%, where only the effect on the probability of working is insignificant. The level increases are 30.3 minutes (25%), $12.01 (29%), 4.2pp (10%), and 32.9 minutes (12%), respectively.

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. In our primary analysis, we settled on 21 days after enrollment to define *Post*; however, since this choice was not preregistered, one may be concerned about specification search. Figure 2 repeats the specification shown in Table 1, Panel A, Column 1 for 31 possible definitions of *Post*, ranging from 0 to 30 days. We see that the effect remains relatively stable and significant at the 5% level across most of the range of values (and significant at the 10% level across the entire range).

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 TOT given the staggered roll-out and endogenous timing of take-up by drivers. We choose to define "ever take-up" as a time-invariant property of a driver and to instrument it with the similarly time-invariant treatment indicator. This approach 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 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 in Appendix Table A.3. Comparing the primary Poisson IV specifications to Table 1 we see that this approach produces larger coefficients on minutes (13.0% vs. 10.1%), dollars (13.3% vs. 10.8%), and minutes conditional on working (11.8% vs. 6%) but cross the significance threshold, while the effect on the probability of work is smaller and remains insignificant (1.8% vs. 4.0%). When take-up is defined as activation, the differences are even larger across minutes (50.7% vs. 21.0%), earnings (48.3% vs. 22.8%), and minutes conditional

on working (44.2% vs. 13%), and slightly smaller (6.4% vs. 7.8%) for the probability of work, but again these suffer from a lack of precision.

4.3.3 Use of Pre-Treatment Observations and Fixed Effects

We next investigate whether results are similar when omitting pre-treatment observations. Our primary specification uses the balanced panel spanning pre- to post-treatment days to potentially increase the precision of our estimates; however, this choice 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 a two-step generalized method of moments estimator to calculate the correct clustered standard errors. In Appendix Table A.4, we reproduce Table 1 using just the observations at least 21 days after driver enrollment and omitting any fixed effects. Panel A shows that this approach produces quite similar ITT point estimates (e.g., 1.3% vs. 1.4% for minutes in Column 1). Columns 1 to 4 of Panels B and C are our main interest, and we again see quite similar point estimates (e.g., 9.6% vs. 10.1% for minutes in Panel B and 19.5% vs. 21.0% in Panel C) and standard errors.

We also examine whether our results are sensitive to omitting date fixed effects from the interaction with cohort fixed effects. Using the full balanced panel, we limit the set of controls to cohort fixed effects and a single binary indicator for the 3-week-post period (to isolate treatment vs. control differences after treatment begins). Appendix Table A.5 shows that this alternative specification produces extremely similar point estimates and standard errors to those in our main table.

4.3.4 Extending the Sample to Examine Dynamics

Finally, we consider how treatment effects evolve over time as a cohort gets further from the enrollment date. If there are novelty effects, learning, or other types of dynamics, then it is possible that the treatment effects estimated over our fairly short window over- or understate the effects that may be observed over a longer horizon. To investigate, we trace out the evolution of treatment effects estimated over each week past enrollment. The main challenge to this exercise is the limited post-treatment observation window. To partly address this shortcoming, in Appendix B, we explain how we can fill in the final two weeks of the experiment using overlapping driver-hour-level data from Chen et al. (2019), albeit with some necessary adjustments. This additional data allows us to track some Uber drivers as far out as 7 weeks after entering the experiment.

To explore treatment effects over time, we construct an event study-like figure that modifies the regression in Table 1, Panel A, Column 1 by interacting the Treatment indicator with indicators for each week (defined as 7 days) prior to and after enrollment. Due to the staggered roll-out, we have at most 8 complete weeks prior to and 7 complete weeks after entering the study. We estimate

all weekly treatment effects relative to 1 week prior to enrollment. The top panel of Appendix Figure B.1 shows that, starting in Week 0 when enrollment occurs, the ITT estimate steadily grows over time and first achieves statistical significance at the 5% level in Week 3. It continues to rise and remains significant through Week 5. The final two weeks produce point estimates similar to week 5, but since there are fewer observed drivers this far out from enrollment (i.e., just the early-enrolling cities), the confidence interval widens.

One challenge to the interpretation of this figure is that the staggered roll-out of the experiment means that any comparison between weeks may conflate actual dynamics in treatment effects with heterogeneity in treatment effects across cities/cohorts, as the composition of which drivers/cities contribute to these estimates varies over time. To address this issue, we also estimate the same specification limiting to a set of drivers who started early in the study and entered on roughly the same date. Specifically, we focus on drivers enrolled on their city's modal experiment start date in the four earliest cities: Seattle (Mar 8), San Diego (Mar 8), Madison (Mar 8), and Boston (Mar 10). While noisier, we again see that the treatment effect does not wane toward the end of this window. Together these results suggest that our treatment effect magnitudes are not simply driven by the initial novelty of the program, insofar as such novelty wears off within these first seven weeks.

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 limitation 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.[8]

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, but we can both ask if responses to Uber wages seem non-normatively high and also 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

[8]The most salient concern is that the labor supply elasticities reflect substitution away from Lyft. We cannot identify whether any drivers in our sample drive for Lyft; however, Koustas (2018), using data from a financial aggregator spanning 18,000 rideshare drivers from 2012 to 2016, finds, "*conditional on being an Uber drivers, 93.3 percent of rideshare earnings come from Uber. Conditional on being a Lyft driver, 33 percent of earnings come from Uber.*" Thus, the scope of Lyft to Uber substitution may be somewhat limited.

adjustment than other classes of workers.

To translate our estimates to a wage elasticity, it is not *a priori* obvious whether to use the ITT or TOT. While wage shifts have perfect compliance as they apply to all drivers, the programmatic costs of Instant Pay (e.g., subsidizing transaction fees) are only borne by the drivers who use it. This consideration 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 low take-up during the short pilot period. On the other hand, there are reasons the TOT may not extrapolate to the majority of drivers who eventually take up Instant Pay. Namely, we know they differ on observable characteristics. Table 2 predicts the probability of treated drivers reaching various take-up stages and shows that drivers who ever applied and activated their cards were more likely to be female (2.9pp and 1.1pp) and younger (-3.6pp and -2.2pp for being older than the median age of 42). For predicting application, being above the median income was also a significant negative predictor of take-up (e.g., -1.1pp for having proxied income exceed $58,750). What may be more relevant are unobservable differences from those who take-up later. For example, early adopters may have had the highest valuation for Instant Pay (and thus possibly the largest LATEs); however, it is also plausible that early adopters are the most attentive to emails or notifications. The correlation between attentiveness and valuation of Instant Pay is unclear, meaning it is also possible that our TOT may underestimate the TOT for individuals who eventually adopt. As a result, both ITT and TOT benchmarks may be useful.

If we use 1.92 as the wage elasticity, this elasticity implies that the ITT estimate of a 1.4% (work minutes) labor supply response to the offer of *Instant Pay* is roughly equivalent to raising wages for all drivers by 0.7% (.014/1.92). Similarly, the TOT using application (Panel B of Table 1) of 10.1% is roughly equivalent to raising wages by 5.2%, while the TOT using activation of the card (Panel C of Table 1) of 21.0% is equivalent to raising wages by 10.9%. While we cannot observe outcomes after the full roll-out of Instant Pay or estimate how much of the labor supply response came from substitution away from other forms of work, all of these wage benchmarks suggest that the EWA makes work on Uber considerably more attractive to Uber drivers.

Table 2: Take-Up Analysis

	Ever Applied		Ever Registered		Ever Activated Card		Ever Utilized	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High Payday Effects	0.0252***	0.0253***	0.0245***	0.0246***	0.0142***	0.0144***	0.0127***	0.0129***
	(0.0030)	(0.0031)	(0.0029)	(0.0030)	(0.0021)	(0.0022)	(0.0018)	(0.0019)
Female		0.0285***		0.0274***		0.0114***		0.0072***
		(0.0045)		(0.0044)		(0.0032)		(0.0027)
Older than Median Driver Age		-0.0357***		-0.0308***		-0.0218***		-0.0213***
		(0.0030)		(0.0029)		(0.0021)		(0.0018)
Above Median Driver Median HH Income (Home Block Group)		-0.0108***		-0.0098***		-0.0008		-0.0015
		(0.0031)		(0.0030)		(0.0022)		(0.0018)
Above Median Uber Experience (Months)		0.0028		0.0031		0.0033		0.0053***
		(0.0031)		(0.0030)		(0.0022)		(0.0018)
Full-Time Driver		-0.0174***		-0.0177***		-0.0124***		-0.0090***
		(0.0044)		(0.0042)		(0.0030)		(0.0025)
N	55,833	53,182	55,833	53,182	55,833	53,182	55,833	53,182
DepVarMean	0.132	0.134	0.121	0.123	0.061	0.061	0.043	0.043

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

Notes: This table reports estimates from OLS regressions predicting various take-up measures among treated drivers who were at least 3-weeks post their enrollment date on April 16, 2016. High payday effects is a proxy for driver's impatience in the baseline pre-period adapted from the measure used by (Kaur et al., 2015). This variable equals 1 if a driver's difference in mean daily minutes on Sundays and Mondays in the baseline divided by the mean daily minutes over the entire baseline period is above the sample mean. We also include indicators for median splits on age (42 years), proxied income ($58,750), and Uber driving experience (16 months). Full-Time Driver corresponds to the 12.2% of treated drivers who average working at least 35 hours per week in the common baseline period of Feb 17 to March 2. The take-up regressions include fixed effects for the randomization cohort (city-by-enrollment-date), and we report robust standard errors.

5 Mechanism

What mechanism can explain the observed labor supply responses to flexible pay? 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 hypothesis is that the response relates to time preferences. However, several potential mechanisms can explain the overall increase in labor supply including – but not necessarily limited to – liquidity constraints and consumption smoothing motives, motivational accounts related to the immediacy of rewards, mental (and actual) accounting, and temporal discounting. Ultimately each of these mechanisms may be more or less operative for different drivers; however, in this section, we present a set of findings that implicate present bias as particularly important for understanding the average response.

First, we consider consumption smoothing motives by revising the benchmarking exercise of the previous section and estimating heterogeneous treatment effects. The TOT (using card activation) is equivalent to the response expected for an 11% increase in wages. Since the product effectively reduces the average delay between effort and pay by 6 days (weighting each day of the week equally), this result implies a 6-day discount factor of 0.9. Thus, an exponential discounting model without binding liquidity constraints would imply an annual discount factor of 0.00164 ($= 0.90^{365/6}$) or an annual discount rate of 60,655%. Either sharply-binding liquidity constraints or present bias could help rationalize this large response without requiring a large long-term discount rate. However, on

the former, recall that all drivers had bank accounts, as it was required to receive a direct deposit. Thus, while many drivers are quite liquidity-constrained (Koustas, 2018), it is plausible they have access to credit at a rate cheaper than 11% for a 6-day repayment period. Nonetheless, we examine whether treatment effects are larger for drivers more likely to be liquidity-constrained. While we cannot observe liquidity directly, we use a proxy for their household income (the Census Block Group median household income of their home address). It is worth noting, however, that even if we did observe liquid assets, it would still be challenging to disentangle time preferences from liquidity, as the former can drive the latter, a point well made by Gelman (2022).[9] More broadly, Figure 3 repeats the ITT Poisson specification reported in Column 1, Panel A of Table 1 for the full sample and various sub-samples. In particular, we find little evidence for heterogeneity across all four outcomes by the sample splits. These include sex (male/female) and median splits on age (42 years), proxied income ($58,750), experience (16 months), baseline Uber usage (0.4 average sessions per day in the common pre-period from February 17 to March 2), and full-time work status on Uber (i.e., splitting on the 12% of drivers who average over 35 hours per week in the common pre-period). Of particular note is the lack of heterogeneity by the median income of the drivers' home block group. Notwithstanding the caveat that observed liquidity constraints themselves could be driven by time preferences, we would expect a larger response among lower-income drivers if liquidity constraints alone drove the labor supply response, but we do not see this result directionally or statistically. More broadly, we find no reliable differences across any of the heterogeneity cuts. While the lack of heterogeneity by income could be partially explained by attenuation bias from using a noisy proxy, these results are at least suggestive evidence against an account that only includes liquidity constraints.

Second, we revisit the take-up analysis with a focus on a proxy for present bias constructed from behavior in the common pre-experiment baseline window (February 17 to March 2). This proxy is similar to the measure of "high payday effects" used by Kaur et al. (2015) to predict take-up of dominated pay contracts among piece-rate data entry workers. For each driver, we take the difference in mean minutes worked on Sundays (earnings paid 3 days later) and Mondays (earnings paid 9 days later) in the baseline period and divide it by her daily mean over the entire baseline period. We then define an indicator for high payday effects that equals 1 if a driver's baseline payday effect was above the sample mean. Unlike in Kaur et al. (2015), Uber drivers' paydays in the baseline are not randomized; however, we believe this scaled difference in means captures the spirit of their payday effects measure: time-inconsistent workers are more productive (i.e., supply more labor) the shorter the time interval between work and the receipt of payment for that work. Table 2 shows that having above average payday effects is associated with increases of 2.53pp for ever applying,

[9] 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).

2.46pp for ever registering, 1.44pp for ever activating a card, and 1.29pp for ever utilizing Instant Pay. In percentage terms, these are increases of 18.9%, 20.0%, 23.6%, and 30.0%, respectively, and all estimates are statistically significant at the 1% level. Thus, the more likely a driver was to supply more labor the closer they were to being paid the higher her likelihood of taking up Instant Pay. The size of these correlations are roughly equivalent to the associations with being female and greater than twice the effect of being above the sample median on the proxied household income measure.[10] In Appendix Table A.9 we use an alternative proxy for present bias constructed by regressing each driver's daily session minutes supplied against the Days-Since-Monday categorical variable over the entire baseline period. This proxy similarly captures how likely a driver was to concentrate their labor closer to payday, and we similarly find that a 1 standard deviation increase in this measure strongly predicts higher take-up.

Third, 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, such as the 1/T hyperbolic discounting model, would predict a behavioral response that shrinks as the counterfactual pay delay gets smaller. Figure 4 shows this pattern when we separately estimate by each day of the week, with a 1.9% ITT on Monday mostly decreasing out to -0.4% ITT on Sunday (Appendix Figure B.2 shows this pattern is a bit more pronounced and tightly estimated in our extended sample). By contrast, liquidity constraints and pressing financial needs suggest either no systematic pattern (if labor supply responses match idiosyncratic financial shocks) or that the pattern will match background variation in liquidity. One challenge to estimating background variation in liquidity is that Uber drivers often have multiple other sources of income, including work on other gig platforms.[11] One liquidity event that we know with certainty is the usual Uber 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 does not show this pattern of increasing effects from Wednesday to Tuesday, we run a horserace through regressions in Table 3. Specifically, we interact $Treatment_{iz} * Post_{izj}$ with categorical variables counting the number of days since Monday (0 to 6) or since Wednesday (0 to 6). Column 2 suggests this linear specification with just the Days-Since-Monday interaction broadly parallels the sub-sample analysis, showing an ITT of 2.6% on Monday and this effect decreasing by 0.4pp each day

[10] We can only speculate, but the strong relationship between sex and take-up may reflect the value of the Uber Debit Card providing a separate account for earnings. For example, if women share their primary direct deposit account with a spouse, then Instant Pay allows them to transfer some of their earnings toward a separate account over which they may have more control.

[11] During our sample period, DoorDash drivers were also paid on Wednesdays, while Lyft drivers could have received their weekly direct deposits any day between Tuesday and Sunday, depending on their banks' processing times (DoorDash, 2016; Lyft, 2016b). Additionally, drivers working for Lyft with access to Express Pay could have cashed out their Lyft earnings at any point after earning at least $50 that week (Lyft, 2016a).

out. By contrast, Column 3 shows that the interaction with Days-Since-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-Since-Monday interaction remaining significant in Panel A. Finally, in Appendix C, we estimate labor supply elasticities in response to randomized wage rates in another experiment by day-of-the-week and do not find the downward-sloping pattern found for Instant Pay. This result suggests the observed temporal pattern in response to Instant Pay is unique and does not simply reflect drivers being more elastic early in the work week in response to *any* incentives.

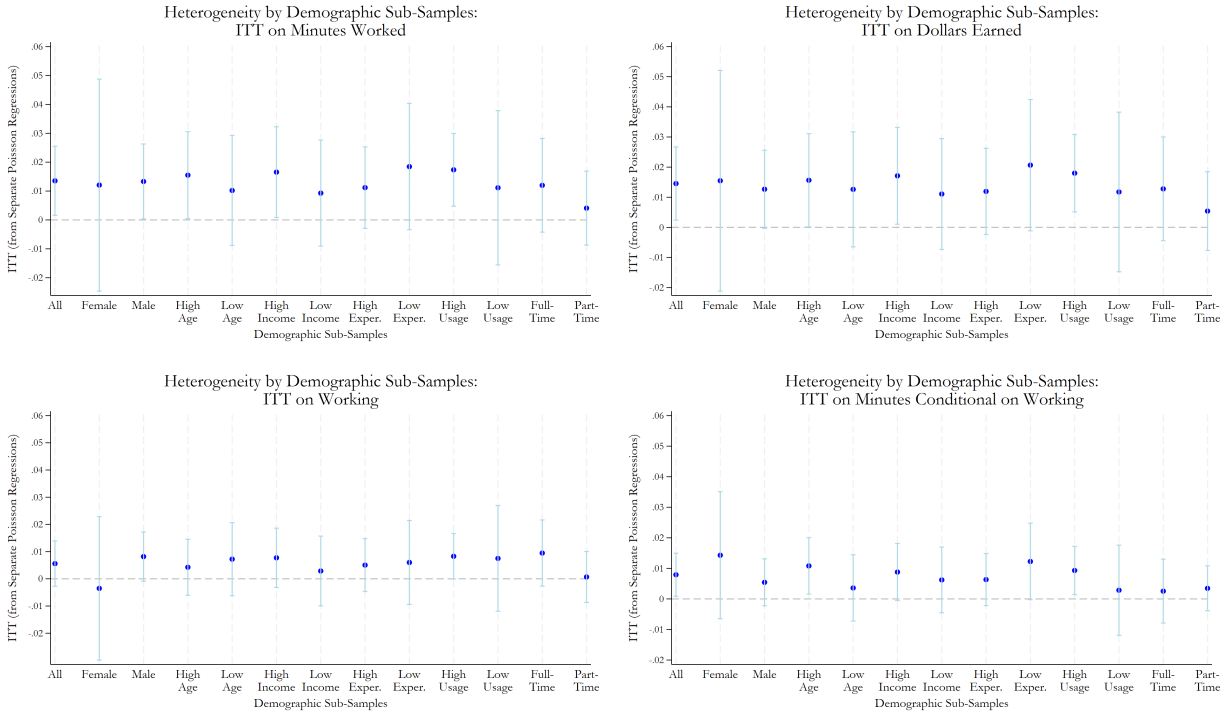
Finally, we examine the daily probability of Instant Pay cash-outs and average daily amounts over time among drivers capable of cashing out. Figure 5 shows a strong weekly pattern emerging by the end of March: cash-outs were the least likely and lowest in average value on Mondays and grow increasingly likely and greater in value throughout the week before peaking on Sundays. This increasing weekly pattern suggests two possibilities. First, drivers may have present-biased preferences over *cash* itself (rather than simply *consumption*), a possibility experimentally demonstrated by Andreoni et al. (2018) using electronic bank transfers. As noted by Andreoni et al. (2018), there is also physiological evidence that the receipt of cash acts the same as a primary reward Lempert et al. (2015, 2016); Löckenhoff et al. (2011). Under this account, drivers may be increasing labor supply in response to the now immediate option-value of accessing their earnings, even if they do not have immediate consumption needs. As a result, patterns in labor supply may be completely decoupled from patterns in cashing out. Second, to understand why cash-outs increase out to Sunday rather than just being idiosyncratic, recall the deadline to cash out falls on that date. Drivers may have procrastinated pushing the button within the driver app to transfer their earnings until Sunday night, after which any remaining funds were unavailable until Wednesday.[12] In summary, if liquidity constraints and immediate consumption needs were the only mechanism driving the response to the EWA, we would expect the patterns of labor supply and cash-outs to move in lockstep. Instead, we see exactly opposite trends in driving and Instant Pay use over the course of a week.

In sum, the evidence presented in this section suggests that present bias over *access to cash* provides a parsimonious explanation for many of the patterns observed in the data. Our results on the relevance of time preferences for explaining the labor supply response to Instant Pay are also consistent with findings on the consumption side by De La Rosa and Tully (2022) who find a

[12]One concern with looking at cash-out probabilities by day-of-the-week is that drivers may have a heuristic of only cashing out once their accumulated earnings have reached a particular threshold. While there is no fee to cash out, it is possible that the small effort of pushing the cash-out button is enough friction to preserve such a rule. Since accumulated earnings are automatically sent for direct deposit at the end of the weekly earnings cycle, accumulated earnings will be mechanically small on Mondays, and thus, this heuristic on its own could produce the observed pattern. To imperfectly account for this, in Appendix Table A.10 and Figure A.5 we reproduce these daily probability figure while controlling for accumulated earnings prior to any cash-outs and show that the pattern is virtually unchanged, suggesting that this pattern is not simply reflecting differences in accumulated earnings by day-of-the-week

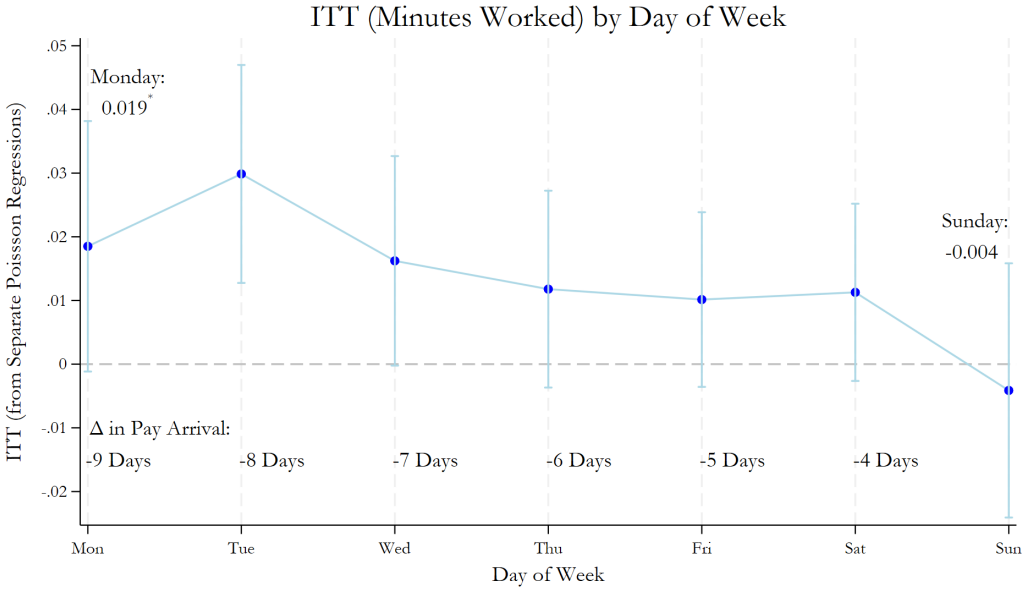
correlation between intertemporal discount rates and increases in consumption among lab participants given access to an EWA product. However, to be clear, several mechanisms could be present simultaneously. For example, Smith (2022), an ethnography of Los Angeles rideshare drivers between 2018 to 2020, found that liquidity-constrained drivers often struggle with the asynchronous timing of pay and expenses and use "a range of techniques, apps, and obstacles both mental and physical to help themselves split or aggregate their income and expenses, so that they have the right amount of money at the right time". He notes the varied ways that drivers find Instant Pay helpful for managing these challenges. Moreover, there are some other psychological accounts that could be relevant. For example, Smith (2022) discusses gamificiation – while the driver app already displayed earnings after each trip and weekly cumulative totals, EWA products enabled drivers to "start at $0 each day so that [they] can easily track [their] progress". In other words, coupling the usual daily-level feedback with the immediate option-value reward could increase intrinsic work motivations (Woolley and Fishbach, 2018; Liu et al., 2022). It is also possible that the separate Instant Pay account either allowed drivers to segment earnings in a way that provided a higher effective savings rate (either due to mental accounting or actual accounting in the intrahousehold bargaining we speculate may drive the sex correlation). While these additional accounts may play some role in the overall labor supply effects, they do not predict the observed day-of-the-week patterns. Altogether, while our evidence is indirect, it collectively suggests an important role for present bias.

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. 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. Regressions include fixed effects for the interactions of randomization cohort and date. Standard errors are clustered by driver.

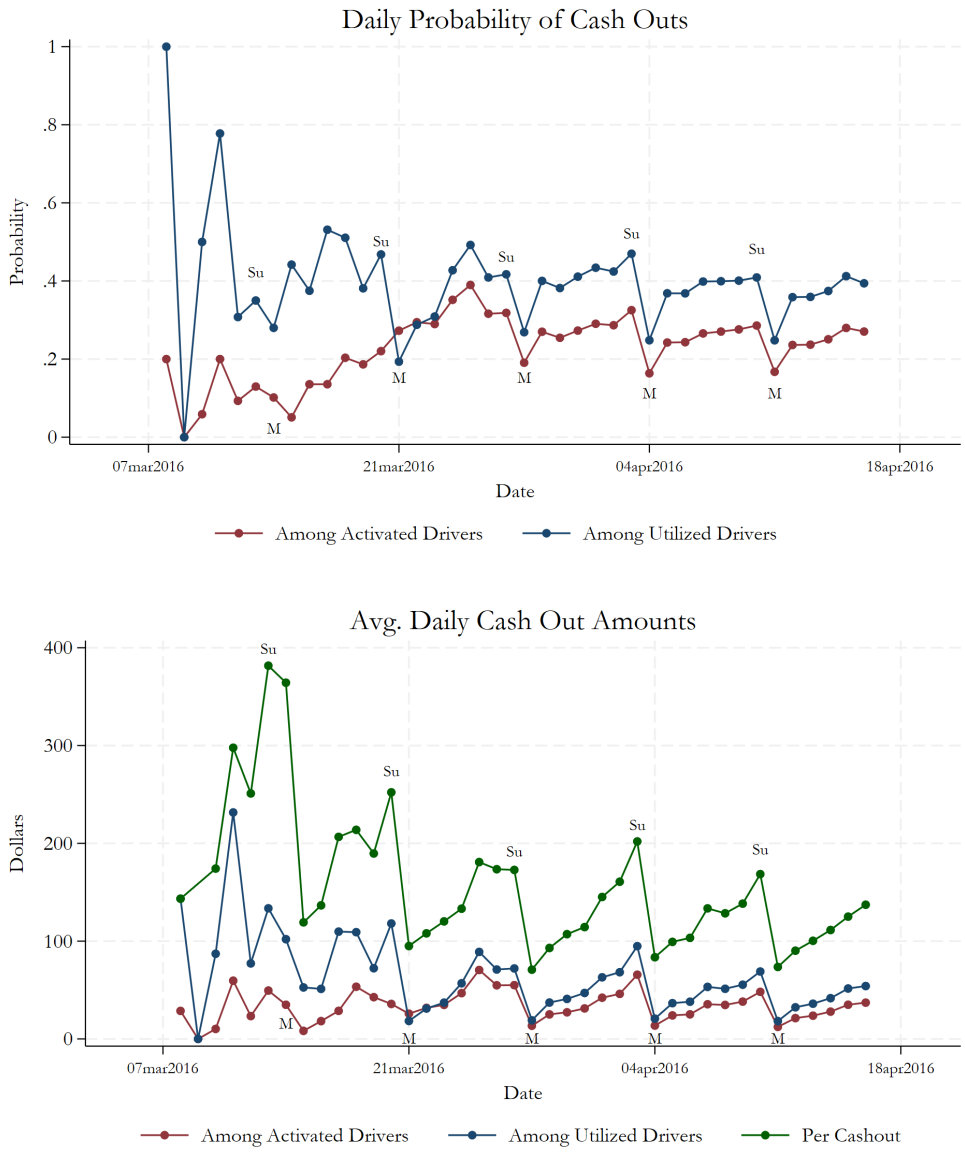
Table 3: Days-Since-Monday vs. Days-Since-Wednesday

	Minutes			
	(1)	(2)	(3)	(4)
Treat X Post	0.014**	0.026***	0.010	0.024**
	(0.006)	(0.009)	(0.007)	(0.010)
Treat X Post X Days Since Monday		-0.004**		-0.004**
		(0.002)		(0.002)
Treated X Post X Days Since Wednesday			0.001	0.000
			(0.001)	(0.001)
N	13,242,030	13,242,030	13,242,030	13,242,030
DepVarMean	121.752	121.752	121.752	121.752

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

Notes: Column 1 repeats the specification used in Table 1 to estimate the ITT on minutes worked (Panel A), earnings (Panel B), the probability of working (Panel C), and minutes conditional on working (Panel D). Column 2 modifies this by including an interaction with Days-Since-Monday which goes from 0 (Monday) to 6 (Sunday), while Column 3 instead does so for Days-Since-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.

Figure 5: Daily Cash Out Probability and Average Amounts



Notes: This figure plots the daily probability of cashing out earnings using Instant Pay and average amounts. Daily likelihoods and amounts cashed out among activated and utilized drivers are restricted to the groups of drivers capable of cashing out on that day (e.g., March 28 is the first date on which there were at least 1,000 drivers who had activated their cards). We say that a driver is capable of initiating a cash out on a given day if they have either activated their Uber debit card or cashed out ("utilized") at any point up to and including that day. By contrast, the average daily amount per cash out is the total dollars cashed out divided by the number of cash-outs on that day.

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, 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 11%. Moreover, the increase in work hours is concentrated in days further away from the counterfactual payday, in line with previous research documenting payday effects in labor supply (Kaur et al., 2015).

Estimating the welfare and possible general equilibrium 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 both the optimal base fare and surge multipliers – all of which should increase consumer welfare. While we do not see evidence of significant price changes, such price changes could have occurred after the feature was available to all drivers. Second, increasing pay frequency may improve ride quality for consumers by reducing the psychological effects of financial concerns for drivers (e.g., improving attention and focus as in Kaur et al. (2022)), especially if they respond to the option-value rather than actual receipt of early wages.[13]

The effects of Instant Pay on driver welfare largely depend on whether drivers are sophisticated or naïve about their time preferences. If drivers do not suffer from present bias, then Instant Pay may increase welfare by relaxing very tightly binding liquidity constraints and by allowing them to withdraw pay at a cadence better matching their expenses. If drivers do display present bias and are sophisticated (i.e., they correctly forecast their future periods' labor-leisure trade-offs) then they will only enroll in Instant Pay if it is welfare-improving because enrollment has a hassle cost and only facilitates early access to payment after getting the card in a future period. 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

[13] Kaur et al. (2022) test the effects of varying pay timing 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 that this is through fewer costly mistakes by alleviating the psychological effects of financial strain. However, whether these financial strain effects would be relevant among Uber drivers in the US 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).*"

increases present discounted expected utility.[14]

There is, however, ambiguity on welfare in the case of full or partial naivete.[15] 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.[16] Tempering these concerns, recall that drivers' day-of-the-week cashouts pattern does not match the high early-in-the-week pattern displayed by the additional trips induced by Instant Pay; a synchronicity we would have expected if money earned through Instant Pay went predominantly to finance immediate temptation spending.

Additionally, while we've focused on present bias and beliefs about present bias (naivete), it is 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). Finally, the welfare effects all of these considerations would likely depend on the potential of Instant Pay to substitute for high-interest vehicles like payday loans, and the welfare impacts of reductions in the utilization of high-cost debt (Allcott et al., 2022).

This paper shows both that (a) there is demand for increased pay flexibility even among weekly-paid workers, and that (b) increased pay flexibility can benefit the firm by increasing labor supply through a relatively inexpensive lever. As such, firms are likely to increasingly offer such pay flexibility, particularly in contexts in which the firm benefits from increased labor supply. While our limited observation window precludes an analysis of pay flexibility on Uber driver retention, other research suggests that such flexibility may also benefit firms through lower employee turnover (Baker, 2018; Murillo et al., 2023).

This shift towards greater pay flexibility is growing as more financial companies provide EWA solutions to companies and workers directly. For example, the largest US employer, Walmart, allows

[14] 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.

[15] 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.*"

[16] 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.

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), and *Even* also provided financial planning and savings tools which may help address some of the consumption misforecasting risks of higher pay frequency, though more research is needed on possible consumption effects (Crosman, 2018). Similar EWAs are also being offered and utilized by employees of McDonalds, Outback Steakhouse, DoorDash, GrubHub, among others (Gee, 2017), and in 2022 over 7 million U.S. workers used EWAs to access around $22 billion in earnings (Consumer Finance Protection Bureau, 2024).

It is important to note several risks associated with the use and growth of EWA products as identified by the Consumer Finance Protection Bureau (CFPB). Workers can face relatively high transaction fees, especially for expedited payments, under the direct-to-consumer model or under the employer-integrated model when the employer does not subsidize such fees. These high fees can be equivalent to APRs over 100% and place additional financial burden on frequent EWA users. The CPFB also raises concerns that workers may financially overextend themselves if they use multiple EWA products simultaneously, increasing the likelihood of incurring additional overdraft or nonsufficient funds fees. Finally, workers' limited understanding of the terms and conditions of these products, especially since the EWA products are not credit and lack regulatory oversight, may also lead to unintentional misuse and unexpected fees, potentially increasing the financial instability of EWA users (Consumer Finance Protection Bureau, 2024). Finally, while Instant Pay was free to use during the experiment and for several years after, that is no longer the case for most drivers, so caution should be taken in extrapolating to the present version of the feature.

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 to better understand the full financial impact of these policies. 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.

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.

Allcott, H., J. Kim, D. Taubinsky, and J. Zinman (2022): "Are High-Interest Loans Predatory? Theory and Evidence from Payday Lending," *The Review of Economic Studies*, 89, 1041–1084.

Andreoni, J., C. Gravert, M. A. Kuhn, S. Saccardo, and Y. Yang (2018): "Arbitrage Or Narrow Bracketing? On Using Money to Measure Intertemporal Preferences," Working Paper 25232, National Bureau of Economic Research.

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.

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.

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.

Baker, T. H. and S. Kumar (2018): "The Power of the Salary Link: Assessing the Benefits of Employer-Sponsored FinTech Liquidity and Credit Solutions for Low-Wage Working Americans and Their Employers," Tech. rep., M-RCBG Associate Working Paper Series, unpublished manuscript.

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*.

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.

Consumer Finance Protection Bureau (2024): "Data Spotlight: Developments in the Paycheck Advance Market," Tech. rep., Consumer Finance Protection Bureau.

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.

DoorDash (2016): "Payment Details," https://web.archive.org/web/20160303200309/http://doordash.squarespace.com:80/payment-details, accessed: 2024-09-14.

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.

GoBank (2022): "Uber Debit Card Account Updates and GoBank Debit Card Benefits," https://www.gobank.com/uber-debit-card-program-changes?uclick_id=dab2fce1-40b1-4d48-9887-81da67785974, accessed: 2024-09-14.

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.

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

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*.

Koren, J. R. (2016): "Uber drivers don't need Green Dot accounts for Daily Payments Anymore," Accessed: 2024-09-14.

Koustas, D. K. (2018): "Consumption Insurance and Multiple Jobs: Evidence from Rideshare Drivers," *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.

Lempert, K. M., P. W. Glimcher, and E. A. Phelps (2015): "Emotional arousal and discount rate in intertemporal choice are reference dependent." *Journal of Experimental Psychology: General*, 144, 366–373.

Lempert, K. M., E. Johnson, and E. A. Phelps (2016): "Emotional arousal predicts intertemporal choice." *Emotion*, 16, 647–656.

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

Liu, Y., Y. Yang, X. Bai, Y. Chen, and L. Mo (2022): "Do Immediate External Rewards Really Enhance Intrinsic Motivation?" *Frontiers in Psychology*, 13.

Lyft (2016a): "Express Pay," https://web.archive.org/web/20160317052140/https://help.lyft.com/hc/en-us/articles/213830188-Express-Pay-, accessed: 2024-09-14.

——— (2016b): "How to Set Up Your Account To Get Paid," https://web.archive.org/web/20160421051934/https://help.lyft.com/hc/en-us/articles/214216927-How-to-Set-Up-Your-Account-To-Get-Paid, accessed: 2024-09-14.

Löckenhoff, C. E., T. O'Donoghue, and D. Dunning (2011): "Age differences in temporal discounting: The role of dispositional affect and anticipated emotions." *Psychology and Aging*, 26, 274–284.

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.

Montgomery, J. M., B. Nyhan, and M. Torres (2018): "How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do about It," *American Journal of Political Science*, 62, 760–775.

Murillo, J., B. Vallée, and D. Yu (2023): "Fintech to the (Worker) Rescue: Access to Earned Wages, Financial Health and Employee Turnover," Tech. rep., Working Paper, unpublished manuscript.

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.

Smith, B. (2022): "Learning and Earning in This New Economy: A Study of Rideshare Drivers in Los Angeles," PhD dissertation, Pardee RAND Graduate School, Santa Monica, CA.

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

Uber (n.d.): "Instant Pay," https://www.uber.com/us/en/drive/driver-app/instant-pay/, accessed: 2024-09-14.

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

Woolley, K. and A. Fishbach (2018): "It's about time: Earlier rewards increase intrinsic motivation." *Journal of personality and social psychology*, 114, 877.

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)

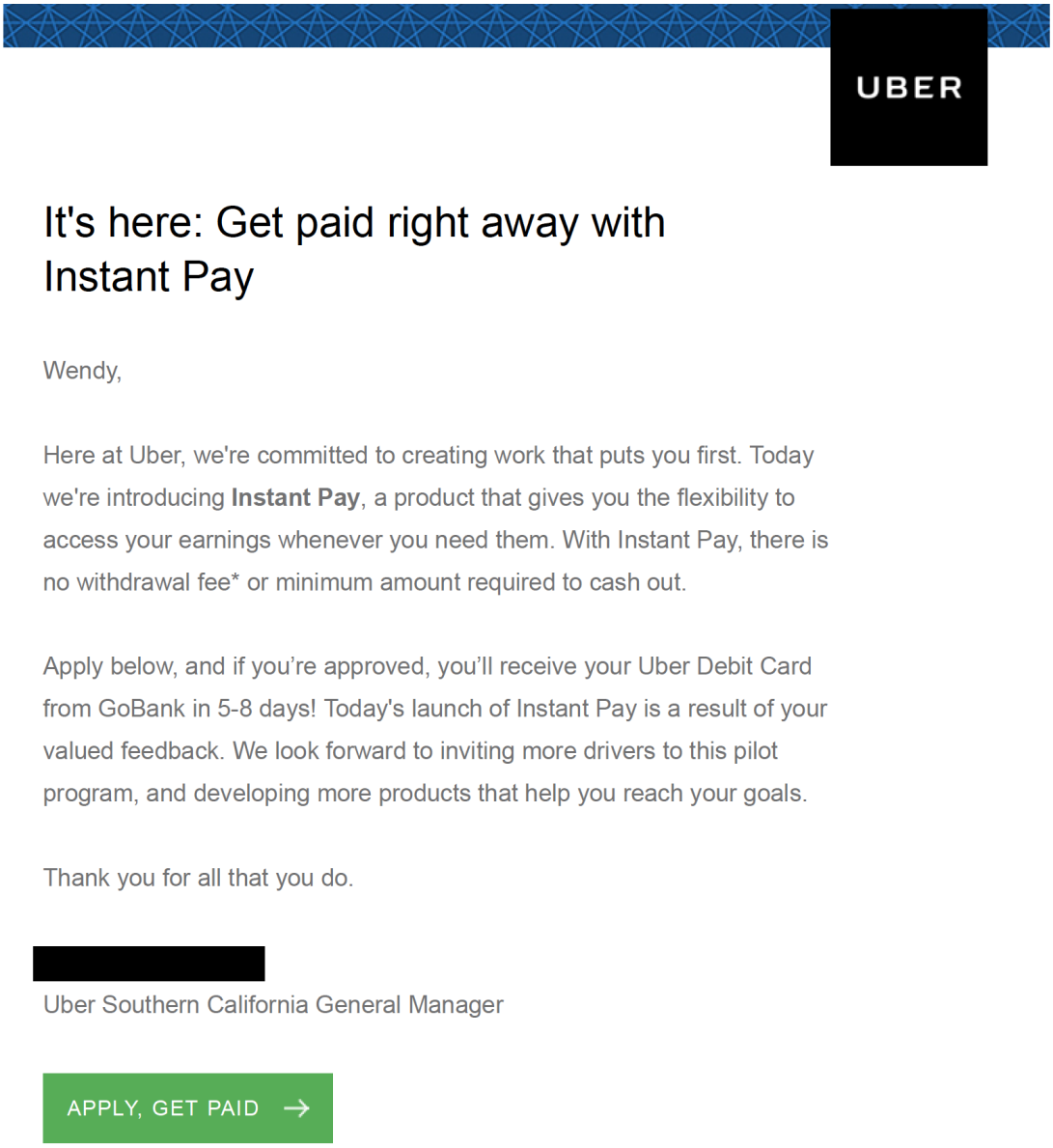
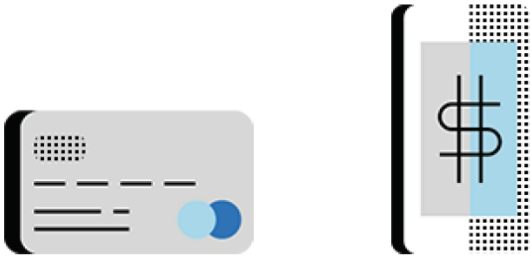


Figure A.2: Launch Date Email (Treatment, Details)

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 waved 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)



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)

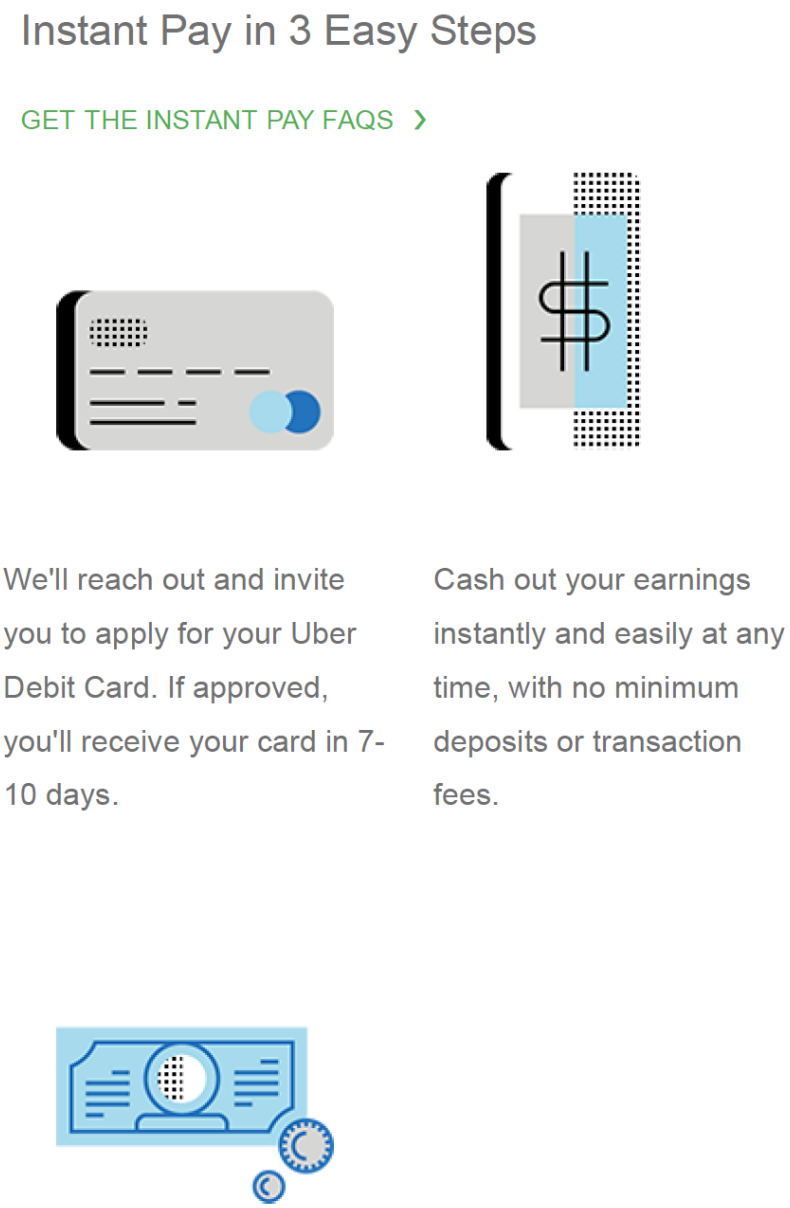


Table A.1: Balance Table (Pre-Experiment Driver Characteristics)

	(1) Treatment		(2) Control		T-test P-value
Variable	N	Mean/SE	N	Mean/SE	(1)-(2)
Female	105686	0.16	105411	0.16	0.36
		(0.00)		(0.00)	
Age (Years)	105634	43.00	105360	42.93	0.12
		(0.04)		(0.04)	
Median HH Income (Home Block Group)	104034	64,588.14	103666	64,679.43	0.59
		(97.18)		(98.79)	
Uber Experience (Months)	105686	20.36	105411	20.29	0.21
		(0.05)		(0.05)	
Avg Number of Shifts (Daily)	111227	0.53	110935	0.53	0.87
		(0.00)		(0.00)	
Total Number of Shifts (Baseline)	111227	7.91	110935	7.94	0.87
		(0.02)		(0.02)	
Avg Minutes Worked (Daily)	111227	110.87	110935	110.89	0.62
		(0.37)		(0.37)	
Total Minutes Worked (Baseline)	111227	1,663.04	110935	1,663.33	0.62
		(5.56)		(5.56)	
Avg Earnings (Daily)	111227	38.29	110935	38.30	0.74
		(0.13)		(0.13)	
Total Earnings (Baseline)	111227	574.32	110935	574.48	0.74
		(1.96)		(1.95)	
Total Days Worked (Baseline)	111227	5.90	110935	5.93	0.62
		(0.01)		(0.01)	
F-test of joint significance (F-stat)					1.02
F-test, number of observations					207551

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	55841	0.13	55662	0.00	0.13***
		(0.00)		(0.00)	
Ever Registered	55841	0.12	55662	0.00	0.12***
		(0.00)		(0.00)	
Ever Activated Card	55841	0.06	55662	0.00	0.06***
		(0.00)		(0.00)	
Ever Withdrew Cash	55841	0.04	55662	0.00	0.04***
		(0.00)		(0.00)	
Panel B: Outcomes (Driver-Day Level)					
Total Minutes Worked (Daily)	641298	141.56	639602	139.67	1.89***
		(0.26)		(0.26)	
Total Earnings (Daily)	641298	52.45	639602	51.64	0.81***
		(0.10)		(0.10)	
Number of Shifts (Daily)	641298	0.63	639602	0.63	0.00***
		(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 May 2. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels.

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

	Poisson				Levels			
	Minutes	Dollars	Work	Minutes | Work	Minutes	Dollars	Work	Minutes | Work
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: IV (TOT, Take-Up = Applied)								
Applied	0.137	0.133	0.018	0.118*	17.505	5.845	0.011	28.072*
	(0.104)	(0.105)	(0.064)	(0.065)	(12.792)	(4.521)	(0.028)	(16.039)
N	13,242,030	13,241,824	13,242,030	5,795,535	13,329,180	13,329,180	13,329,180	5,795,535
DepVarMean	121.752	42.082	0.435	280.039	120.885	41.781	0.432	280.039
Panel B: IV (TOT, Take-Up = Activated)								
Activated	0.507	0.483	0.064	0.442*	66.756	22.292	0.043	94.825*
	(0.396)	(0.401)	(0.245)	(0.247)	(48.783)	(17.239)	(0.107)	(54.214)
N	13,242,030	13,241,824	13,242,030	5,795,535	13,329,180	13,329,180	13,329,180	5,795,535
DepVarMean	121.752	42.082	0.435	280.039	120.885	41.781	0.432	280.039

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

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				Levels			
	Minutes	Dollars	Work	Minutes | Work	Minutes	Dollars	Work	Minutes | Work
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Intent-to-Treat (ITT)								
Treatment	0.013**	0.016**	0.006	0.008**	1.891**	0.811**	0.003	2.359**
	(0.006)	(0.006)	(0.004)	(0.004)	(0.867)	(0.329)	(0.002)	(1.087)
N	1,280,900	1,280,900	1,280,900	600,393	1,280,900	1,280,900	1,280,900	600,393
DepVarMean	139.668	51.638	0.467	298.808	139.668	51.638	0.467	298.808
Panel B: IV (TOT, Take-Up = Applied)								
Ever Applied	0.096**	0.110***	0.041	0.055**	14.043**	6.023**	0.019	16.780**
	(0.042)	(0.043)	(0.031)	(0.025)	(6.443)	(2.444)	(0.015)	(7.738)
N	1,280,900	1,280,900	1,280,900	600,393	1,280,900	1,280,900	1,280,900	600,393
DepVarMean	139.668	51.638	0.467	298.808	139.668	51.638	0.467	298.808
Panel C: IV (TOT, Take-Up = Activated)								
Ever Activated Card	0.195**	0.223***	0.085	0.104**	30.005**	12.869**	0.042	32.714**
	(0.082)	(0.082)	(0.063)	(0.046)	(13.763)	(5.221)	(0.032)	(15.099)
N	1,280,900	1,280,900	1,280,900	600,393	1,280,900	1,280,900	1,280,900	600,393
DepVarMean	139.668	51.638	0.467	298.808	139.668	51.638	0.467	298.808

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

Notes: This table repeats Table 1 but limits the sample to "post-treatment" observations (i.e., 21 days after 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: Robustness: Only Cohort Fixed Effects

	Poisson				OLS (Levels)			
	Minutes	Dollars	Work	Minutes | Work	Minutes	Dollars	Work	Minutes | Work
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Intent-to-Treat (ITT)								
Treat X Post	0.014**	0.015**	0.006	0.008**	1.959**	0.789**	0.003	2.375**
	(0.006)	(0.006)	(0.004)	(0.004)	(0.857)	(0.322)	(0.002)	(1.077)
N	13,329,720	13,329,720	13,329,720	5,800,486	13,329,720	13,329,720	13,329,720	5,800,486
DepVarMean	120.884	41.781	0.432	279.956	120.884	41.781	0.432	279.956
Panel B: IV (TOT, Take-Up = Applied)								
Applied X Post	0.103**	0.113**	0.043	0.060**	14.548**	5.860**	0.021	16.895**
	(0.045)	(0.046)	(0.032)	(0.027)	(6.366)	(2.396)	(0.015)	(7.668)
N	13,329,720	13,329,720	13,329,720	5,800,486	13,329,720	13,329,720	13,329,720	5,800,486
DepVarMean	120.884	41.781	0.432	1.335	120.884	41.781	0.432	279.956
Panel C: IV (TOT, Take-Up = Activated)								
Activated X Post	0.216**	0.238**	0.084	0.130**	31.085**	12.522**	0.044	32.942**
	(0.097)	(0.098)	(0.067)	(0.057)	(13.600)	(5.118)	(0.032)	(14.962)
N	13,329,720	13,329,720	13,329,720	5,800,486	13,329,720	13,329,720	13,329,720	5,800,486
DepVarMean	120.884	41.781	0.432	279.956	120.884	41.781	0.432	279.956

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

Notes: This table repeats Table 1 but includes an indicator for the 3-week post period and uses cohort (city-by-enrollment-date) fixed effects for all specifications.

Table A.6: Robustness: Including Minneapolis-St.Paul

	Poisson				Levels			
	Minutes	Dollars	Work	Minutes | Work	Minutes	Dollars	Work	Minutes | Work
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Intent-to-Treat (ITT)								
Treat X Post	0.024***	0.024***	0.012***	0.013***	3.327***	1.229***	0.006***	3.740***
	(0.006)	(0.006)	(0.004)	(0.004)	(0.844)	(0.316)	(0.002)	(1.068)
N	13,403,588	13,403,370	13,403,588	5,863,932	13,498,680	13,498,680	13,498,680	5,863,932
DepVarMean	121.524	41.999	0.434	279.775	120.586	41.674	0.431	279.775
Panel B: IV (TOT, Take-Up = Applied)								
Applied X Post	0.176***	0.176***	0.088***	0.093***	24.587***	9.085***	0.042***	26.421***
	(0.045)	(0.045)	(0.031)	(0.026)	(6.243)	(2.336)	(0.015)	(7.562)
N	13,403,588	13,403,370	13,403,588	5,863,932	13,498,680	13,498,680	13,498,680	5,863,932
DepVarMean	121.524	41.999	0.434	279.775	120.586	41.674	0.431	279.775
Panel C: IV (TOT, Take-Up = Activated)								
Activated X Post	0.369***	0.372***	0.180***	0.201***	52.263***	19.312***	0.089***	51.313***
	(0.095)	(0.096)	(0.066)	(0.056)	(13.273)	(4.967)	(0.031)	(14.710)
N	13,403,588	13,403,370	13,403,588	5,863,932	13,498,680	13,498,680	13,498,680	5,863,932
DepVarMean	121.524	41.999	0.434	279.775	120.586	41.674	0.431	279.775

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

Notes: This table repeats Table 1 but includes 2,825 drivers assigned to Minneapolis-St.Paul whom we omit from our main analyses due to randomization issues in that city market.

Table A.7: Robustness: Transformed Log(1+Y) Outcomes

	Log(1+Y)			
	Minutes	Dollars	Work	Minutes | Work
	(1)	(2)	(3)	(4)
Panel A: Intent-to-Treat (ITT)				
Treat X Post	0.019*	0.016*	0.003	0.010**
	(0.011)	(0.009)	(0.002)	(0.004)
N	13,329,180	13,329,180	13,329,180	5,795,535
DepVarMean	2.285	1.848	0.432	5.293
Panel B: IV (TOT, Take-Up = Applied)				
Applied X Post	0.141*	0.118*	0.019	0.074**
	(0.084)	(0.070)	(0.015)	(0.032)
N	13,329,180	13,329,180	13,329,180	5,795,535
DepVarMean	2.285	1.848	0.432	5.293
Panel C: IV (TOT, Take-Up = Activated)				
Activated X Post	0.302*	0.251*	0.042	0.145**
	(0.180)	(0.150)	(0.032)	(0.062)
N	13,329,180	13,329,180	13,329,180	5,795,535
DepVarMean	2.285	1.848	0.432	5.293

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

Notes: This table repeats Table 1 but estimates treatment effects for log(1+Y) transformations of daily minutes worked, earnings, and minutes conditional on working. We also reproduce estimates for the untransformed "Work" indicator for having driven any sessions on a particular day.

Table A.8: 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, FL	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, MO	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	732	11apr2016	0.89	0.51	0.00	3.61
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.

Table A.9: Predicting Take-Up using Baseline Regressions for Minutes Worked on Days Since Monday

	Ever Applied		Ever Registered		Ever Activated Card		Ever Utilized	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Baseline Avg Minutes Worked (Daily)	-0.0001***	-0.0000*	-0.0001***	-0.0000	-0.0001***	-0.0000**	-0.0000***	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Days Since Monday Slope (Std.)	0.0085***	0.0082***	0.0076***	0.0073***	0.0049***	0.0048***	0.0043***	0.0041***
	(0.0013)	(0.0013)	(0.0012)	(0.0013)	(0.0009)	(0.0009)	(0.0008)	(0.0008)
Female		0.0284***		0.0274***		0.0111***		0.0071***
		(0.0045)		(0.0044)		(0.0032)		(0.0027)
Older than Median Driver Age		-0.0342***		-0.0295***		-0.0208***		-0.0206***
		(0.0030)		(0.0029)		(0.0021)		(0.0018)
Above Median Driver Median HH Income (Home Block Group)		-0.0109***		-0.0100***		-0.0010		-0.0016
		(0.0031)		(0.0030)		(0.0022)		(0.0018)
Above Median Uber Experience (Months)		0.0033		0.0037		0.0037*		0.0056***
		(0.0031)		(0.0030)		(0.0022)		(0.0018)
Full-Time Driver		-0.0044		-0.0077		-0.0013		-0.0029
		(0.0071)		(0.0069)		(0.0050)		(0.0042)
N	55,833	53,182	55,833	53,182	55,833	53,182	55,833	53,182
DepVarMean	0.132	0.134	0.121	0.123	0.061	0.061	0.043	0.043

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

Notes: This table reports estimates from OLS regressions predicting various take-up measures among treated drivers who were at least 3-weeks post their enrollment date on April 16, 2016. Baseline Avg Minutes Worked is the mean of daily session minutes over the common pre-period before any city launches its experiment (February 17 - March 2). Days Since Monday Slope are driver-specific OLS coefficients estimated over the same period and standardized to have mean of 0 and standard deviation of 1. We also include indicators for median splits on age (42 years), proxied income ($58,750), and Uber driving experience (16 months). Full-Time Driver corresponds to the 12.2% of treated drivers who average working at least 35 hours per week in the common baseline period of Feb 17 to March 2. The take-up regressions include fixed effects for the randomization cohort (city-by-enrollment-date), and we report robust standard errors.

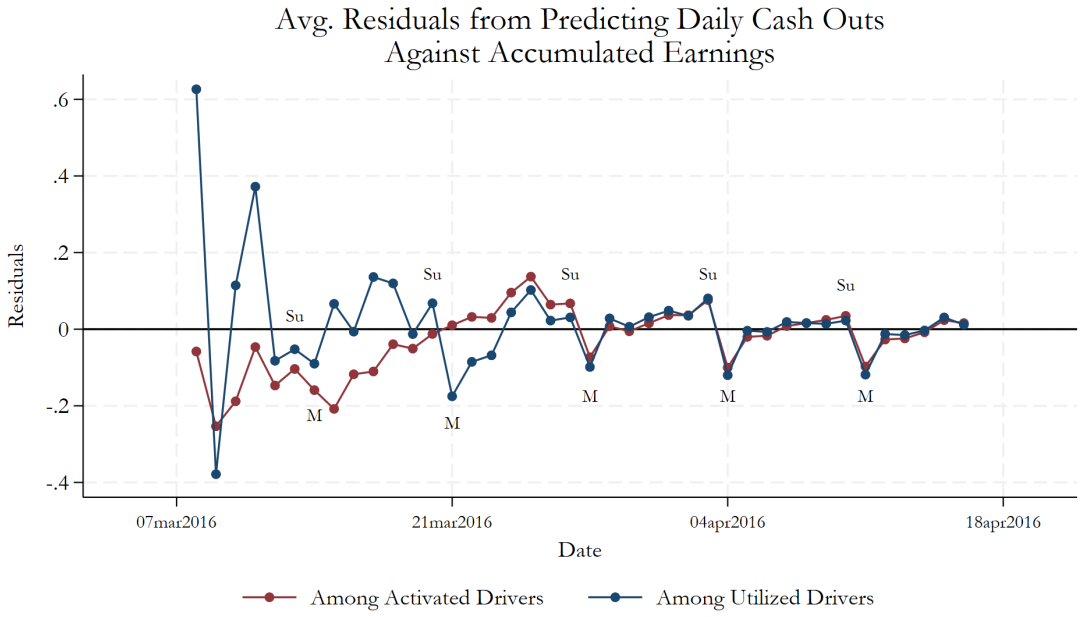
Table A.10: Predicting Daily Cash-Outs

	Activated Card			Utilized		
	(1)	(2)	(3)	(4)	(5)	(6)
Days Since Monday	0.0174***		0.0216***	0.0236***		0.0214***
	(0.0010)		(0.0011)	(0.0013)		(0.0015)
Accumulated Earnings		-0.0001***	-0.0001***		0.0002***	0.0001***
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
N	59,286	59,286	59,286	40,998	40,998	40,998
DepVarMean	0.257	0.257	0.257	0.379	0.379	0.379

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

Notes: This table reports estimates from OLS regressions predicting the daily probability of treated drivers ever cashing out. Columns 1-3 are restricted to drivers who have activated their Uber Debit card prior to or on a given date, and Columns 4-6 are restricted to drivers who have utilized Instant Pay prior to or on a given date. Days Since Monday is a categorical variable that goes from 0 (Monday) to 6 (Sunday). Accumulated Earnings captures a drivers' earnings at the time of potential cash-out, which resets to $0 after drivers either initiate a cash-out or enter a new weekly earnings cycle at 4am on Mondays. This measure equals either their cumulative earnings at the end of the day on on days without cash-outs, their actual balance at the time of realized cash-outs, or the mean of the latter on days with multiple cash-outs. Standard errors are clustered by driver.

Figure A.5: Residuals from Predicting Daily Cash-Outs



Notes: This figure plots residuals from predicting the daily probability of drivers ever cashing out on a given date for drivers who have activated their cards or utilized Instant Pay up to and including that date. The residuals are calculated using OLS regressions for predicting any cash-outs against drivers' accumulated earnings at the time of potential cash-out (Columns 2 and 5 in Table A.10.

Appendix B: Extended Sample and Dynamics

The staggered roll-out of Instant Pay across city markets, Uber's decision to end the RCT early, and our preferred 3-week *post*-period definition all result in a short observation period and a limited set of cities and drivers that effectively contribute to our treatment effect estimates. To increase this subset of drivers and explore potential treatment dynamics over a longer time frame, we merged our experimental assignments with an overlapping driver-hour-date data extract pulled from Chen et al. (2019). This alternative source allows us to extend our sample period through the end of the RCT on April 16, 2016, and we can identify our main estimates off 181,746 drivers across 27 city markets, up from 111,503 drivers across 12 markets. In this appendix, we discuss the differences across the two data sources and replicate our main regression estimates using outcome measures derived from: 1) the trips-level data for February 17-April 16 and 2) the hourly data for February 17-May 2. We also take advantage of the additional weeks of data in the latter panel to estimate treatment dynamics over time and check robustness of the weekly pattern of ITT effects on daily session minutes.

First, we describe how we constructed our daily labor supply outcomes using the driver-hour-date data for dates spanning mid-February to early May. Chen et al. (2019) drop drivers whose cities cross time-zones (which complicates the computation of driving time) and who were active in fewer than 16 weeks out of their 36 week sample period, so we must drop 3,193 Instant Pay drivers omitted from this source. The hourly data summarizes the total minutes a driver was "active" and their earnings accumulated during that hour on that calendar date for UberX trips. While we use the same definition of driving sessions as blocks of times where drivers are active with fewer than 120 minutes of inactivity, we lack more granular timestamps for activity within an hour block. As such, we treat all recorded minutes as occurring at the start of the hour to measure inactivity and create sessions. This decision introduces some measurement error in demarcating sessions as we likely both under- and over-estimate the length of dormant periods, leading to under- and over-estimating the number of sessions, respectively.

Minutes worked per session is then defined as the length of the entire session and measured as the sum of active minutes in the first session hour, active minutes in the last session hour, and 60 minutes per intermediate session hour. For example, a driver with 30 active minutes from 10 to 11am, 40 active minutes from 12pm to 1pm, 15 active minutes from 1pm to 2pm, and 60 active minutes from 4pm to 5pm on the same day is treated as having 2 sessions that were 135 and 60 minutes long. We measure session earnings by summing payouts over the corresponding hours. Again, we assign sessions to the day in which they began in order to measure daily minutes worked and earnings. Finally, there is an issue with data quality for April 21, where that date has fewer than 40% of driver-hour-date observations as directly adjacent days (and was missing entirely from the trips-data). Since we cannot differentiate whether these missing observations are non-active days for those drivers, for April 21 we set all drivers' outcomes to missing instead of zero if they

do not have a session that day.

Appendix Table B.1 reproduces our main Poisson regression estimating the treatment effects using just the trips data (Columns 1-4) and hourly data (Columns 5-8). Across the board for all labor supply outcomes, we estimate weakly larger treatment effect magnitudes when using only the trips data. These estimates are relatively similarly precise across the two sources, despite the differences in Uber services captured (UberX, Uber pool, and Uber Eat vs. only UberX) and potential measurement error from moving to a less granular source of driver activity.

Table B.1: Robustness: Poisson Regressions Across Different Labor Supply Data Sources

	Trips Data				Hourly Data			
	Minutes	Dollars	Work	Minutes | Work	Minutes	Dollars	Work	Minutes | Work
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Intent-to-Treat (ITT)								
Treat X Post	0.014**	0.015**	0.006	0.008**	0.013**	0.011**	0.006*	0.007**
	(0.006)	(0.006)	(0.004)	(0.004)	(0.005)	(0.005)	(0.003)	(0.003)
N	13,242,030	13,241,824	13,242,030	5,795,535	16,388,437	16,388,359	16,388,437	7,067,408
DepVarMean	121.752	42.082	0.435	280.039	110.762	40.881	0.430	257.726
Panel B: IV (TOT, Take-Up = Applied)								
Applied X Post	0.101**	0.108**	0.040	0.060**	0.077**	0.067*	0.032	0.044**
	(0.045)	(0.046)	(0.032)	(0.027)	(0.032)	(0.034)	(0.022)	(0.019)
N	13,242,030	13,241,824	13,242,030	5,795,535	16,388,437	16,388,359	16,388,437	7,067,408
DepVarMean	121.752	42.082	0.435	280.039	110.762	40.881	0.430	257.726
Panel C: IV (TOT, Take-Up = Activated)								
Activated X Post	0.210**	0.228**	0.078	0.130**	0.154**	0.133*	0.064	0.089**
	(0.097)	(0.098)	(0.067)	(0.057)	(0.066)	(0.069)	(0.045)	(0.038)
N	13,242,030	13,241,824	13,242,030	5,795,535	16,388,437	16,388,359	16,388,437	7,067,408
DepVarMean	121.752	42.082	0.435	280.039	110.762	40.881	0.430	257.726

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

Notes: This table repeats the Poisson specifications in Table 1 across different data sources for measuring labor supply. Columns 1 - 4 use trips-level data ranging from February 17 - April 16 for all drivers who were enrolled prior to April 16 and reproduce our main paper estimates. Columns 5-8 use session-level outcomes derived from the hourly data in Chen et al. (2019) for all dates from February 17 - May 2 and exclude 3,193 drivers who were not present in the hourly data due to time-zone and activity restrictions discussed in Chen et al. (2019).

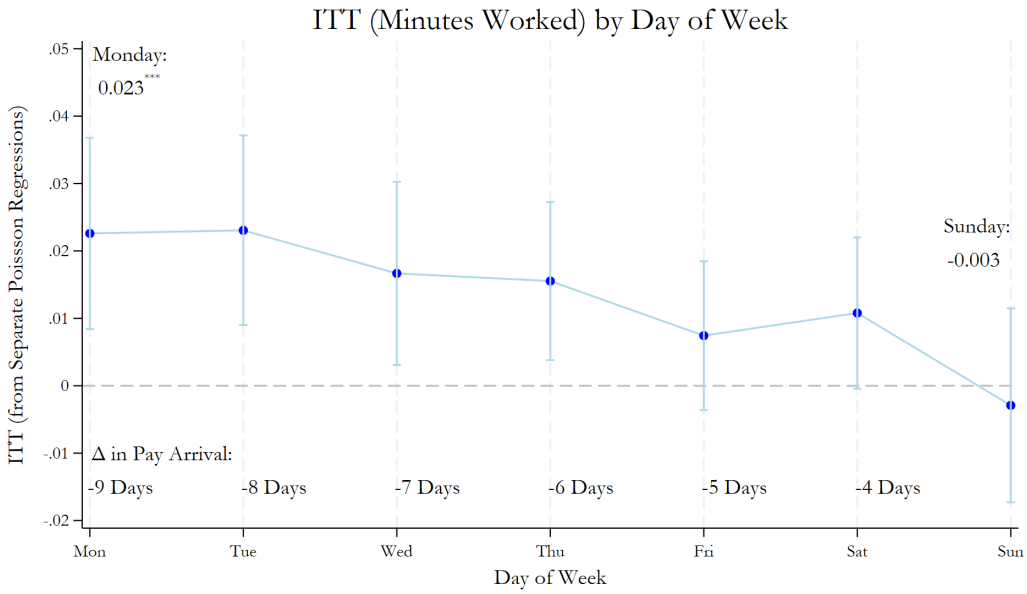
Appendix Figure B.1 examines dynamics as discussed in Section 4.3.4, and Appendix Figure B.2 repeats the day-of-the-week ITT analysis in Figure 4 using the hourly data for labor supply outcomes.

Figure B.1: Treatment Dynamics Over Time: ITT on Minutes Worked



Notes: This figure displays event-study estimates for the weekly ITT effects of Instant Pay for the full sample (top) and then for a sample restricted to the modal enrollment cohorts in four cities who enrolled drivers early into the RCT (bottom): Seattle (Mar 8), San Diego (Mar 8), Madison (Mar 8), and Boston (Mar 10). We estimate Poisson regressions that interact treatment with indicator variables for each week (7 days) before and after the driver-specific enrollment date and include fixed effects for the interactions of randomization cohort (city-by-enrollment-date) and date. We omit the interaction term for 1 week prior to enrollment, so the weekly-treatment effects are estimated relative to the 7 days just before drivers enter the RCT. We plot the point estimates and along with dashed lines for the 95% confidence intervals (+/- 1.96*SE). Standard errors are clustered by driver.

Figure B.2: Robustness: ITT by Day of Week



Notes: This figure repeats Figure 4 using the balanced panel derived from the hourly data used in Chen et al. (2019). The panel covers 218,969 drivers for all dates spanning February 17-May 2.

Appendix C: Day-of-Week Elasticities

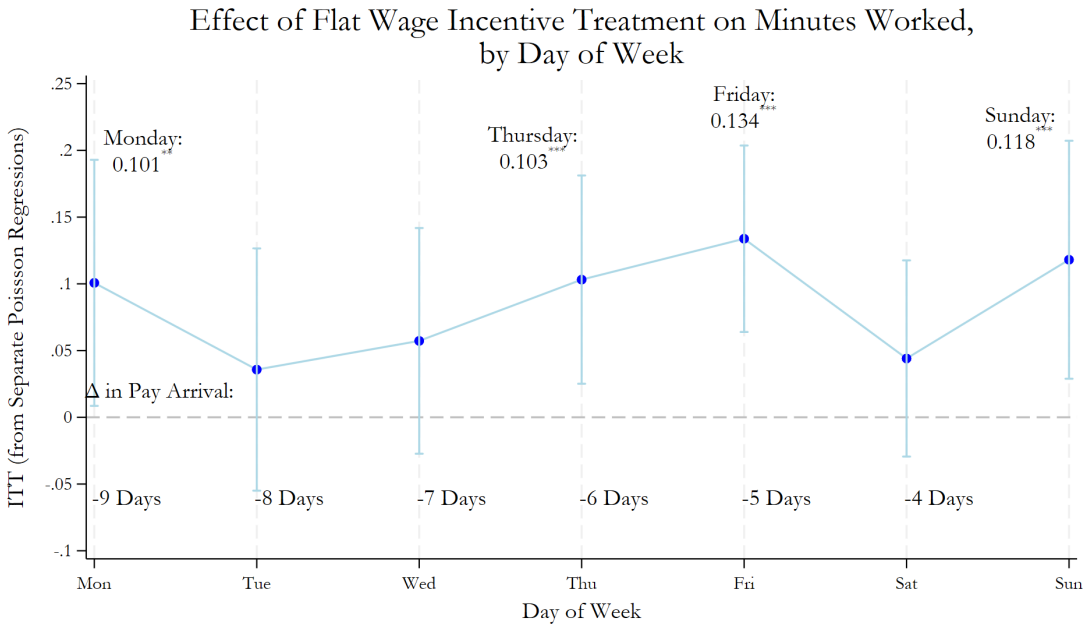
In this Appendix, we briefly describe a separate experiment that allows us to assess whether labor supply elasticities (in response to randomized wage rates) vary by day-of-the-week. The aggregate elasticities were reported in (Chen et al., 2019); however, we report a new analysis that separately estimates treatment effects by day of the week. This analysis allows us to ask whether the day-of-the-week pattern found in response to the Instant Pay may simply reflect drivers being more elastic early in the work week in response to *any* type of incentive, rather than the focal feature that distinguishes these days in the context of Instant Pay (i.e., the counterfactual lag between work effort and pay). This analysis is important because payday is not randomly assigned, and thus heterogeneous treatment effects by day-of-the-week could reflect other correlates of these days.

In April 2016, Uber ran a separate, concurrent experiment in Orange County by randomizing 16,197 drivers evenly into control and two treatment groups, where treated drivers were offered either a $1 bonus for every $10 earned ("flat treatment") or $2 bonus for every $10 earned past $300 earned ("progressive treatment") for 3 weeks from April 4-April 25. As discussed in Chen et al. (2019), Uber imposed eligibility criteria that affected both which drivers and which trips' earnings were eligible for these incentives. Drivers had to meet specific thresholds for accepting and completing trips, and trips had to originate in Orange County for the bonus to apply. Using the first week of this experiment where drivers did not appear to respond to the offer of these incentives, Chen et al. (2019) estimate that the 10% incentive effectively boosted eligible drivers' wage rate between 2 and 5%.

Random assignment in the Orange County experiment was orthogonal to random assignment in the Instant Pay RCT. There were 4,316 drivers enrolled in both experiments and balanced across the 2 X 3 experiment conditions. We prefer comparing the Orange county flat treatment group to the control group because the incentive applies to earnings for eligible trips equally throughout the week, allowing a direct test for day-of-week-varying labor-supply elasticities. Meanwhile, under the progressive treatment the wage rate increase is more likely to apply to eligible trips that occur later in the week because drivers must first reach the $300 weekly earnings threshold.

We re-estimate Figure 4 on drivers in the flat treatment and control groups (N=10,798 drivers) – replacing the offer of Instant Pay as the treatment with this flat wage incentive – using separate Poisson regressions for each day-of-the-week subsample, and plot point estimates and 95% confidence intervals below:

Figure C.1: Orange County Experiment



Notes: This figure repeats Figure 4 using data from an experiment where Uber offered drivers in Orange County additional incentives. Each dot corresponds to a coefficient from a separate Poisson regression for each "Day of the Week" sub-sample for weeks 2 and 3 of the experiment. "Treated" is defined as drivers who were offered a 10% bonus on their weekly earnings and who effectively received a 2.3% wage increase due to restrictions on eligibility for the incentives. The bars correspond to the 95% confidence interval (+/- 1.96*SE) on that point estimate. The Poisson regressions include fixed effects for the calendar date. Standard errors are clustered by driver.

The labor supply elasticities in response to the flat wage increase appear relatively flat across the weekly cycle. If anything, they slightly increase moving from Tuesday to Friday, although our limited sample size prevents more precise estimates. This pattern of day-of-the-week estimates runs counter to the downward trend in Figure 4 and suggests that the latter is driven (at least in part) by Instant Pay, specifically. That is, Figure 4 is unlikely to merely reflect existing patterns in drivers' labor supply responses. However, an important caveat to this analysis is that this experiment was only run with Orange County, CA drivers, so we need to assume that other cities do not meaningfully differ in how labor supply elasticities vary by day-of-the-week.