The Gender Earnings Gap in the Gig Economy: Evidence from over a Million Rideshare Drivers *

Cody Cook, Rebecca Diamond, Jonathan V. Hall
John A. List, and Paul Oyer

May 2020

Abstract

The growth of the "gig" economy generates worker flexibility that, some have speculated, will favor women. We explore this by examining labor supply choices and earnings among more than a million rideshare drivers on Uber in the United States. We document a roughly 7% gender earnings gap amongst drivers. We show that this gap can be entirely attributed to three factors: experience on the platform (learning-by-doing), preferences and constraints over where to work (driven largely by where drivers live and, to a lesser extent, safety), and preferences for driving speed. We do *not* find that men and women are differentially affected by a taste for specific hours, a return to within-week work intensity, or customer discrimination. Our results suggest that, in a "gig" economy setting with no gender discrimination and highly flexible labor markets, women's relatively high opportunity cost of non-paid-work time and gender-based differences in preferences and constraints can sustain a gender pay gap.

*We thank seminar participants at NBER Labor Studies, Society of Labor Economics, IZA Workshop on Gender and Family Economics, the Utah Winter Business Economics Conference, University of Illinois, Stanford, Chicago, and Yale for valuable input. We also received valuable input from Judith Chevalier, Dan Knoepfle, Libby Mishkin, Ian Muir, and Orie Shelef and valuable data assistance from Lindsey Currier and Bharat Chandar. The editor and anonymous referees provided very insightful comments that substantially improved the manuscript.

Cook: Stanford University Graduate School of Business; Diamond and Oyer: Stanford University Graduate School of Business and NBER; Hall: Uber Technologies, Inc.; List: University of Chicago and NBER. Cook is a former employee of Uber and retains equity in the company. List was Chief Economist at Uber when we started this project and is now Chief Economist at Lyft. Hall is a current Uber employee and retains equity in the company.

1 Introduction

The wage gap between men and women has narrowed throughout the past four decades, with 2010 estimates suggesting women earn 88 cents on the dollar when compared to similar men in similar jobs (Blau and Kahn (2017)).[1] Much of the remaining wage gap can be explained by fewer hours worked and weaker continuity of labor force participation by women, especially for middle-age workers where gender wage gaps are largest (Bertrand et al. (2010) and Blau and Kahn (2017)). Goldin (2014) has suggested that work hours and disruption in labor force participation dramatically lower wages due to a "job-flexibility penalty," where imperfect substitution between workers can lead to a convex hours-earnings relationship. In contrast, the role of on-the-job training (Mincer and Polachek (1974)) is thought to play an economically smaller role (Blau and Kahn (2017)).[2]

It is possible that the growth of the "gig" economy could help narrow the gender wage gap in the economy. Gig economy jobs divide work into small pieces and then offer those pieces of work to independent workers in real-time, allowing for easy substitution of work across workers. This ease of worker substitutability should severely limit a "job-flexibility penalty," and potentially exhibit little to no gender pay disparity. Indeed, Hyperwallet (2017) reports that "86% of female gig workers believe gig work offers the opportunity to make equal pay to their male counterparts" and Cubas et al. (2019) estimate that more than half of the gender wage gap comes from women's inflexibility. Estimates suggest that around 7% of workers are independent contractors in their primary job (Katz and Krueger, 2019). Although most survey estimates of self-employment are flat or declining over the last decade or two, IRS tax forms show an increasing prevalence of self-employment activity that is not reflected in survey responses (Abraham et al., 2018). The online platform economy in particular is growing rapidly, with 1.6% of Chase account holders receiving income from online platforms in the first quarter of 2018 (Farrell et al., 2018b). As more industries gravitate towards using gig work, the importance of the job-flexibility penalty in gender wage inequality could weaken.

[1] See Table 4 Panel B of Blau and Kahn (2017), combining the residual wage gap with the effects of experience.

[2] Blau and Kahn (2017) note that the evidence here is mostly based on older studies (Light and Ureta (1995)). Indeed, data on experience often contain sizable measurement error in traditional datasets (Blau and Kahn (2013)).

In this paper, we make use of a sample of over a million drivers to quantify the determinants of the gender earnings gap in one of the largest gig economy platforms: Uber's platform for connecting riders and drivers. Uber sets its driver fares and fees through a simple, publicly available formula, which is invariant between drivers; neither the pay formula nor the dispatch algorithm for assigning riders to drivers depend on a driver's gender. Further, similar to many parts of the larger gig economy, on Uber there is no negotiation of earnings, earnings are not directly tied to tenure or hours worked per week, and we can demonstrate that customer-side discrimination is not materially important. These job attributes explicitly rule out the possibility of a "job-flexibility penalty." The flexibility of Uber also differentiates it from taxi markets with supply-limiting medallions. To maximize the return on a medallion, contracts are generally structured to make it uneconomical for taxi drivers to work anything less than a very long day (Haggag et al., 2017). Likely due to these differences in flexibility, the share of women drivers on Uber is nearly double the share of women taxi drivers.

We find that male drivers earn roughly 7% more per hour than female drivers on average. We can explain the entire gap with three factors. First, through the logic of compensating differentials (and the mechanisms of surge pricing and variation in driver idle time), hourly earnings on Uber vary predictably by location and time of week, and men tend to drive in more lucrative locations. This is largely because male drivers tend to live near more lucrative locations and because men earn a compensating differential for their willingness to drive in areas with higher crime and more drinking establishments.

The second factor is rideshare-specific human capital. Even in the relatively simple production of a passenger's ride, past experience is valuable for drivers. A driver who has completed more than 2,500 lifetime trips earns 14% more per hour than a driver who has completed fewer than 100 Uber trips, in part because she learns where and when to drive and how to strategically cancel and accept trips. Male drivers accumulate more experience than women by driving more each week and being less likely to stop driving with Uber. Because of these returns to experience and because the typical male driver on Uber has more experience than the typical female—putting them higher on the learning curve—men earn more money per hour.

A unique aspect of our data is our ability to both precisely measure a driver's experience and measure the return to experience through improved driver productivity, holding fixed the compensation schedule. Traditional datasets studying the gender pay gap often have very poor measures of experience (usually just a worker's age, sometimes years of employment). This measurement error in experience leads to attenuated estimates of the return to experience. We show that this measurement error in experience can lead to biased estimates of the job-flexibility penalty. When we remove our precise measure of experience (number of rides completed) and replace it with the typical measure used in other papers (a quadratic in driver age), we find a convex hours/earnings relationship in Uber drivers. Drivers who drive 30+ hours per week for Uber earn a 9% higher hourly wage than those who driver fewer than 10 hours per week. However, once we add in our precise controls for driver experience, we find a *concave* hours/earnings relationship. Drivers working 30+ hours per week earn 7% less per hour than those working fewer than 10 hours per week. For Uber drivers, this is likely due to drivers who work fewer hours per week being able to cherry pick high pay hours, while those working full-time must work some of the less lucrative times.

Because drivers who work long hours also accumulate human capital at a faster rate per week, the importance of the job-flexibility penalty in the gender pay gap might be overstated in studies lacking good measures of worker experience. Separating out the importance of job-flexibility versus the return to experience for the gender pay gap in the broader economy is critical for formulating policy. Policies that improve job-flexibility (such as moving towards gig work) may only have a modest effect on the gender gap if the returns to on-the-job experience are a key driver of the hour-earnings relationship.

The residual gender earnings gap that persists after controlling for experience and where and when drivers work can be explained by a single variable: average driving speed. Increasing speed increases expected driver earnings in almost all Uber settings. Drivers are paid according to the distance and time they travel on trip and, in the vast majority of cases, the loss of per-minute pay when driving quickly is outweighed by the value of completing a trip quickly to start the next trip sooner and accumulate more per-mile pay (across all trips). Men's higher driving speed appears to result from preferences as we see no evidence that drivers respond to the incentive to drive faster. Men's higher average speed and the productive value of speed for Uber and the drivers (and, presumably, the passengers) enlarges the pay gap in this labor market.

We interpret these determinants of the gender pay gap—a propensity to gain more experience, choice of different locations, and higher speed—as preference or constraint-based characteristics that are correlated with gender and make drivers more productive.[3] While much prior work has also shown a relationship between the gender earnings gap and factors that are likely to be related to preferences/constraints, we know of no prior work that fully decomposes the gender earnings gap in any setting. Beyond measuring the gender earnings gap and unpacking it completely in an important labor market, our simple analysis provides insights into the roots of the gender earnings gap and, following the decomposition described in Gelbach (2016), the share of the pay gap that can be explained by each factor. First, driving speed alone can explain nearly half of the gender pay gap. Second, over a third of the gap can be explained by on-the-job learning, a factor which is often almost impossible to evaluate in other contexts that lack high frequency data on pay, labor supply, and output. The remaining gender pay gap can be explained by choices over where to drive. Men's willingness to supply more hours per week (enabling them to learn more) and to target the most profitable locations shows that women continue to pay a cost for working reduced hours each week, even with concavity in the hours-earning schedule.

As the gig economy continues to grow, it will likely bring even more flexibility in earnings opportunities, which is valued by at least some workers (Angrist et al. (2017) and Chen et al. (2019) document the value of flexibility to drivers) if not by all workers (see Mas and Pallais (2017)). However, the returns to experience and the temporal and geographic variation in worker productivity will likely persist and thus sustain a gender earnings gap.

We also show that at least three factors that one might expect to favor men in the labor market and to be relevant for Uber drivers do *not* contribute to men earning more. First, customers do not discriminate by gender of driver. Second, there is not a financial return to work intensity within a period of time. For example, driving forty hours per week does not increase hourly pay (in that

[3] For the purposes of this paper, we often use "preferences" to refer to an individual's optimal choices given his/her constraints. Naturally, men and women may face different constraints that will impact these choices. We discuss whether the results are due to a few specific preferences or constraints in more depth in the Appendix.

week) relative to driving twenty hours. Finally, though men gain from driving *more* hours over time, they do not make more due to the *specific* hours of the day and days of the week that they choose to drive. Men drive more late at night and other times but, on average, they don't drive at more or less lucrative times of the day or week than women drive. The fact that these issues do not penalize women suggest that the non-discriminatory and flexible nature of gig work may help women to achieve pay equity conditional on accumulated experience and some dimensions of preferences and constraints.

Our paper, like a few others that have come before, focuses on gender differences within a single company and/or a narrowly defined set of workers.[4] Several prior papers have established clear empirical connections between gender pay gaps and factors indicative of gender differences in preferences and constraints (especially as they relate to child bearing). Bertrand et al. (2010) show that the gender gap among graduates of a single prestigious MBA program starts small but widens considerably. The growth in the gender gap can be explained almost entirely by differences in hours worked (due to a combination of women working fewer hours per week, conditional on working, and being more likely to have gaps in their careers) which can, in turn, be explained by child rearing. The tight relationship between the pay gap and children indicates gender-based differences in preferences and constraints, though the exact mechanism for the hours/earnings connection is unclear. The authors cannot determine whether the female earnings penalty is due to a convex hours-earnings relationship or a learning-by-doing effect. Our results, though in a very different context, are surprisingly similar and our data enable us to quantify the importance of learning-by-doing and to rule out (at least for drivers on Uber) work intensity as a driver of the gap.

Other papers find broadly similar results. In another paper that looks at the transportation sector, Bolotnyy and Emanuel (2019), find that women make less than men at a large public, unionized employer due to differences in how men and women value certain job attributes. They show that "female workers have greater revealed preference for schedule controllability." Azmat and Ferrer (2017) document a large earnings premium for men among young lawyers in the United

[4] For a general overview of the literature on male/female wage differentials and the factors that lead to them, see Altonji and Blank (1999), Bertrand (2011), and Blau and Kahn (2017).

States that is largely attributable to factors related to hours worked such as hours billed and new clients brought in. They find that differences in preferences and constraints, as captured by small children and by stated aspirations to become law firm partners, explain most of the pay and productivity gaps. Gallen (2018) analyzes a broad sample of Danish workers and finds that mothers are less productive than other women or men, which explains most of the wage difference. Barth et al. (2017) show that the gender gap grows substantially with age for the college-educated due to men's pay rising faster within establishments and tie this to the arrival of children.

The returns to hours worked need not generate a gender gap when earnings are linear (or near linear) in hours worked rather than convex. Goldin (2014) posits that in occupations where workers are perfectly substitutable and there are not switching costs, there is no premium for working additional hours. One example discussed by Goldin and Katz (2016) is the market for pharmacists. Pharmacists have become increasingly female over time, the gender pay gap among pharmacists is merely 4%, and the gap only exists for women who have children. As compared to the MBAs in Bertrand et al. (2010)—who often work in occupations where earnings is highly convex in the number of hours worked—the importance of hours and child rearing is economically weaker for pharmacists and there is no evidence of a "job-flexibility penalty." As we discuss in Section 3.3.1, even when workers are perfectly substitutable and there are no switching costs, a premium for long hours can arise in jobs where workers learn-by-doing; the additional hours worked will cause workers to be more productive by pushing them further up the learning curve.

Other papers have attributed part of the gender gap to factors unlikely to be important in the rideshare market—differences in willingness to bargain and firms sharing rents with employees. See Card et al. (2015) and Hirsch et al. (2010) using matched employee/employer data from Portugal and Germany, respectively, Black and Strahan (2001) studying U.S. banks upon deregulation, and a broader analysis of gender and negotiations in Babcock and Laschever (2003). Lab experiments such as Niederle and Vesterlund (2007) and field experiments such as Flory et al. (2015) suggest that another contributing factor to the gender pay gap is that women disproportionately shy away from competition. Other papers show that some of the gender gap can be explained by differential sorting (which could be at least partially due to differences in preferences and constraints), including

Card et al. (2015), Gupta and Rothstein (2005), and Bayard et al. (2003). Sorting is not relevant in our context, as we study a single firm. However, there clearly is gender-based sorting into rideshare driving given that our sample is overwhelmingly male.

Cullen et al. (2018) and Adams (2020) are the only other papers of which we are aware that examines gender pay differences in the gig economy. Cullen et al. (2018) study a low-skill gig platform where workers agree to one-time gigs for tasks such as delivery, laundry, and carpentry. Adams (2020) studies Mechanical Turk workers. While the work on these platforms is also low-skill, it is far more heterogeneous than Uber. Cullen et al. (2018)'s focus is on sorting rather than pay differences within tasks and they show that women sort into lower paying tasks. The fact that Uber drivers are almost perfectly substitutable for one another allows us to isolate the gender pay gap without fear of unobserved differences driving our results. Adams (2020) finds that women with young children take longer to complete MTurk tasks. The basic finding that preferences and constraints can create a gender pay gap likely applies to other sites (such as TaskRabbit, Handy, and MTurk for relatively low skill work, Upwork and Fiverr for workers spanning a large skill distribution, and TopTal for high skill workers) though it would be much more difficult to measure the gap in those settings. Also, note that analyzing rideshare drivers captures the online portion of the gig economy appropriately, as these drivers comprise the vast majority of online gig workers (Farrell et al. (2018a)).

The paper proceeds with a description of our data and the documenting of a 7% hourly earnings gender gap among well over a million drivers on Uber. Having established that there is a gender earnings gap for drivers, we study the details of how drivers are compensated so that we can break down all components that affect driver pay. We focus on drivers in the Chicago metropolitan area to reveal the primary determinants of the earnings gap, though we also show, using data from several other cities, that our conclusions are invariant to the market we choose. We conclude with implications and summary remarks.

2 Uber: Background and Data

2.1 The Uber Marketplace

Uber's software connects riders with drivers willing to provide trips at posted prices. Riders can request a trip through a phone app, and this request is then sent to a nearby driver. The driver can either accept or decline the request during a short time window after seeing the rider's location. If the driver declines the ride, then the request is sent to another nearby driver. Some products slightly vary this experience. For example, UberPOOL trips may involve picking up multiple riders traveling along a similar route. At the end of each ride, the passenger and driver rate each other on a scale from one to five stars.

Drivers have full discretion regarding when and where they work. Unlike wage and salary workers, drivers do not receive standard employee benefits like overtime or (for many, but not all, wage and salary workers) healthcare. A comprehensive discussion of the classification of drivers as independent contractors is outside the scope of this paper, but driver independence is convenient for this study insofar as we do not need to consider differential value for different kinds of compensation beyond monetary compensation and flexibility.

Drivers are paid according to a fixed, non-negotiated formula. For a given trip, the driver earns a base fare plus per-minute and per-mile rates for the time and distance from pickup to dropoff. In times of imbalanced supply and demand, as manifested by high wait times and few available drivers, a "surge" multiplier greater than one may multiply the time and distance-based fare formula. Importantly, there are no explicit returns to tenure (e.g., promotion), convex returns to hours worked (i.e. higher hourly pay for 50 hours of work than 20 hours of work in a week), or opportunities for earnings discrepancies based on negotiated pay differentials on Uber.[5]

In our analysis, we will essentially be treating earnings as equivalent to productivity. This is a reasonable assumption on any single trip, as driver earnings for a trip are highly correlated

[5] Occasionally, certain promotions will pay for convex hours worked by rewarding drivers for hitting certain thresholds of weekly trips; however, these thresholds are tailored to drivers based on their driving frequency in past weeks and attainable even for infrequent drivers. Further, incentives are a small portion of the average driver's pay and our results hold when considering only "organic" (excluding incentives) pay. See Appendix A.5.

with rider fares.[6] The amount a driver earns on a trip could overstate or understate the driver's marginal product of labor if, for example, a driver takes an action that increases or decreases a rider's demand for future Uber rides. For our analysis, this is only an issue if there are differences by driver gender in how drivers affect future demand. To address this, in Appendix A.4 we look at the ratings passengers provide for drivers at the end of each ride. Reassuringly for our approach, the average of rider ratings of drivers is statistically indistinguishable between genders. When we regress ratings on gender and the control variables used throughout the paper, we find an economically trivial (and, in most specifications, statistically insignificant) relationship between driver gender and ratings. These analyses provide some reassurance that there are not important differences by driver gender in drivers' effects on Uber's reputation or a rider's propensity to take future Uber rides.

In our analysis, we focus on the UberX and UberPOOL products to ensure that drivers in our data were completing comparable work and faced similar barriers to entry; other Uber products may have alternative pay structures (e.g., UberEATS) or stricter car and license requirements (e.g., UberBLACK).

2.2 Driver Earnings

For each trip completed, drivers are paid a base fare plus a per-mile and per-minute rate. In Chicago (as of 2017), drivers are paid a $1.70 base fare plus $0.20 per minute and $0.95 per mile for each UberX trip (which are all, at times of high demand, multiplied by the surge multiplier).[7]

Drivers can also earn money from "incentives." For example, drivers may be offered additional pay for completing a set number of trips in a week. Another type of incentive guarantees drivers a certain surge level for trips taken within a given geography and time (e.g. 1.4x all fares in the

[6] Before Summer 2016, driver pay and rider fares for a trip were directly coupled, with a percentage service fee taken by Uber. However, rider fares are now "decoupled" and, while correlated with driver earnings, are not mechanically tied to earnings. Furthermore, while Uber now allows riders to tip their drivers in-app, this did not become available until June 2017, which is outside the scope of our data. Research on the roll-out of tipping on Uber suggests that it had little impact on *total* driver earnings, but that women are tipped $0.05 more per-trip than men, which would narrow the gender pay gap by about 13% (Chandar et al., 2019a,b). We do not believe that cash tips—which were possible before in-app tipping—had a material impact on driver earnings.

[7] For UberX trips, there is also a minimum fare of $4.60 in Chicago.

Chicago Loop during rush hour). While the use of incentives has varied over time, on average they account for under 9% of a driver's hourly earnings in our data.

With all of these components in mind, we formalize the driver's effective hourly earnings $p(\cdot)$ for a given trip as

$$p(\cdot) = 60 * \left(\frac{SM\left(r_b + m_1 r_m + 60 * \frac{m_1 r_t}{s}\right) + I}{w + 60 * \frac{m_0 + m_1}{s}} \right) \tag{2.1}$$

where r_b, r_m, and r_t respectively represent the base fare, per-mile, and per-minute rates, SM is the surge multiplier, m_0 is the number of miles between accepts and pickup, m_1 is the number of miles on trip, s is speed in miles per hour, w is wait time in minutes for dispatch, and I represents the incentive earnings associated with the trip.

For UberPOOL trips—where multiple riders heading in the same direction can ride together—the pay formula treats the chain of trips as a single trip. The driver still receives a base fare for the initial pickup plus a per-mile and per-minute rate.[8] Importantly, pay does not depend on the number of riders in the car.

2.3 National Data

Our national data include all driver-weeks for drivers in the U.S. from January 2015 to March 2017. We limit the data to drivers for Uber's "peer-to-peer services," UberX and UberPOOL; drivers who have completed a trip on other products such as UberXL, UberBLACK, or UberEATS are excluded.[9] The resulting data include over 1.87 million drivers, about 512,000 of whom are female (27.3%).[10] In total, we observe almost 25 million driver-weeks in 196 cities.[11]

For each driver-week, we track total earnings and hours worked. We compute hourly earnings as the total payout in that week divided by hours worked. For the purposes of this paper, a driver

[8] During the time period of our data, UberPOOL rates were often set marginally lower than UberX rates. In Chicago, the per-mile and base fare are identical to UberX, but the per-minute rate is 6 cents lower. Later versions of the pay formula equalized the rates and added additional pickup-fees for each rider to join the Pool chain.

[9] UberEats has a different pay structure than ride sharing, paying piece-rate for pickups, dropoffs, and miles driven, and has less stringent vehicle requirements for drivers. Results are consistent with or without UberEats drivers (who make up approximately 13% of the sample). UberBLACK drivers are commercially licensed and may face large regulatory barriers to entry depending on the city.

[10] This percentage is higher than the number of active female drivers in a given month due to women having higher attrition (Table 1).

[11] We follow Uber's definition of city, which does not always match canonical definitions. For example, the state of New Hampshire is considered a single city.

is considered to be "working" whenever the app is on and available for trips. That is, a driver is "working" while on a trip, en route to a pickup, or available for a dispatch, but not if, for example, they turn off the app to drive home. All earnings are gross earnings. Costs such as gas, car depreciation, and Uber's service fee have not been subtracted from the earnings we present.[12] We discuss costs in more detail in Appendix A.1.

2.4 National gender earnings gap

Table 1 presents summary statistics of driver pay overall by gender. Active drivers gross an average of $376 per week and $21 per hour. More than 60% of those who start driving are no longer active on the platform six months later (though some of these drivers may be on an extended break). Comparing across gender in Table 1, we find a first hint of differences between male and female drivers. Men make nearly 50% more per week than women, which is primarily a reflection of their choice to work nearly 50% more hours per week. On an hourly basis, men make over $1/hour more than women. Men are also less likely to leave the platform.

Table 1: Basic summary statistics, all US drivers

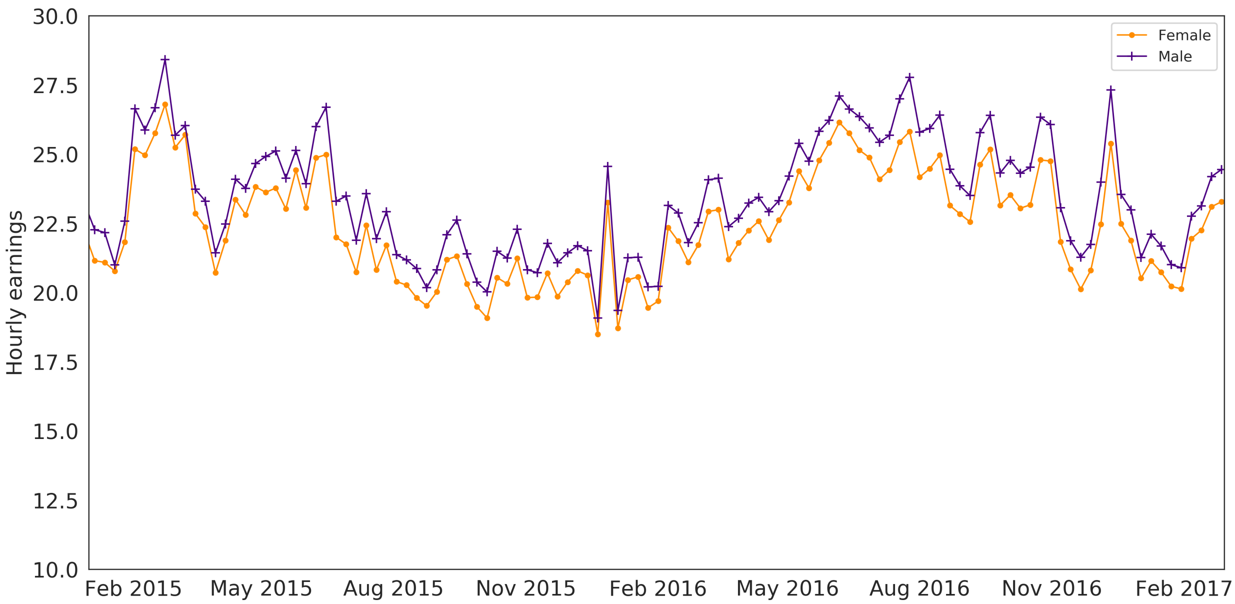
	All	Men	Women
Weekly earnings	$376.38	$397.68	$268.18
Hourly earnings	$21.07	$21.28	$20.04
Hours per week	17.06	17.98	12.82
Trips per week	29.83	31.52	21.83
6 month attrition rate	68.1%	65.0%	76.5%
Number of drivers	1,873,474	1,361,289	512,185
Number driver/weeks	24,832,168	20,210,399	4,621,760
Number of Uber trips	740,627,707	646,965,269	93,662,438

Note: Values are based on all UberX/UberPOOL driver-weeks in the US from January 2015 - March 2017. The percent of drivers who are female varies across city; to mitigate composition effects, we weight averages at the city level by *total* number of drivers in a city, rather than by number of male (or female) drivers. 6 month attrition rate is defined as the percent of drivers who are no longer active 26 weeks after their first trip. We consider drivers to be active on a given date if they complete another trip within another 26 weeks of that date. For calculating the attrition rate, we limit the sample to drivers who completed their first trip between Jan 2015 and March 2016 to allow us to fully observe whether they are inactive, per the definition above, 26 weeks after they join.

[12] Uber increased its service fee from 20% to 25% in September 2015; however, drivers who joined before then were grandfathered in and still pay only 20%. This differentially impacts women, who are more likely to have joined the platform more recently. We look at earnings before the service fee is applied.

Figure 1 provides a graphical view of the hourly earnings gap for all U.S. drivers from early 2015 through early 2017. The gap seen in Table 1 is fairly constant throughout the sample period. Pay of drivers fluctuates, but the changes are generally gender neutral.

Figure 1: Average hourly earnings, US



Note: Data based on hourly earnings averaged across all UberX and UberPOOL drivers who worked in a given week. The percent of drivers who are female varies across city; to mitigate composition effects, we weight averages at the city level by *total* number of drivers in a city, rather than by number of male (or female) drivers. Earnings are gross; costs such as the Uber commission or gas are not subtracted.

Table 2 uses these national data and measures the gender pay gap through a set of standard Mincer regressions. Specifically, we estimate

$$\ln(Earnings_{dt}) = \beta_0 + \beta_1 Male_d + \rho X_{dt} + \epsilon_d \tag{2.2}$$

for driver d in time period t, where $Earnings$ are the gross weekly or hourly earnings in that time period, as described above, $Male$ is an indicator variable for a driver's gender, and X_{dt} is a set of controls such as week and city indicator variables.

Table 2 provides clear evidence that, when examining almost two million drivers across the United States (representing more than one percent of the US workforce) and controlling for the city and the conditions for a given week, there remains a substantial gender pay gap. Men in the

Table 2: Gender pay gap

	All US				Chicago	
	Weekly earnings		Log hourly earnings		Log weekly earnings	Log hourly earnings
isMale	0.4142	0.4092	0.0702	0.0653	0.4315	0.0485
	(0.002)	(0.002)	(0.001)	(0.001)	(0.007)	(0.001)
Intercept	4.9737	4.9208	2.9280	2.8849	5.0487	3.1151
	(0.002)	(0.002)	(0.001)	(0.001)	(0.009)	(0.001)
City	X	X	X	X	X	X
Week		X		X	X	X
N	24,877,588	24,877,588	24,877,588	24,877,588	1,604,627	1,604,627
R^2	0.125	0.136	0.199	0.239	0.038	0.110

Note: This table documents the gender pay gap for all US cities from January 2015 to March 2017. Data are at the driver-week level; weekly earnings is the entire pay for a given week, while hourly earnings is the pay divided by hours worked in the week. Standard errors (clustered at the driver-level) are in parentheses.

US earn about 7% more than women when the analysis is done at the hourly level, indicating that, while a substantial majority of the weekly earnings gender gap is simply due to men driving more hours, there is still a sizable gap when examining hourly earnings.

This gap may seem surprising: men make 7% more per hour, on average, for doing the same job in a setting where work assignments are made by a gender-blind algorithm and the pay structure is tied directly to output and not negotiated. The 7% differential is as large or larger than hourly differentials in other narrowly defined, relatively homogeneous groups such as recent MBAs (Bertrand et al. (2010)) and pharmacists (Goldin and Katz (2016)), but is smaller than the differential in economy-wide samples (Blau and Kahn (2017)).

Throughout the rest of the paper, we focus on drivers in Chicago to decompose the gender gap and analyze its economic roots. This choice reduces the dataset to a more tractable size and allows for more granular data. As shown in Table 2, when the same regressions are done on Chicago drivers alone, the weekly gender earnings gap in Chicago mirrors the national gap and the hourly Chicago gender earnings gap is somewhat lower, at approximately 5%. This small difference between the national and Chicago gap is due to cross-city differences in the factors that explain the gap. We analyze these factors in detail in Chicago, which provides more insight into the roots of the gap than if we were to focus on the generally small differences across cities. In Section 5, we present results for a sample of other large cities and demonstrate that our results are consistent across cities.

3 Decomposing the Wage Gap — Chicago

3.1 Chicago Data

By focusing attention on Chicago drivers, we can examine data at the driver-hour, rather than driver-week, level.[13] A driver-hour is defined as an hour block with some trip activity; for example, 8-9am on a specific Monday. We continue to restrict the data to peer-to-peer drivers in January 2015 to March 2017.

The Chicago dataset includes 120,223 drivers, 36,391 of whom are female (30.2%). In total, we observe 33.0 million driver-hours.[14] As before, we track total gross pay and hours worked for each driver-hour. We compute the implied hourly earnings in a driver-hour as total earnings for trips in that hour divided by minutes worked*60. For trips that span driver-hours, we distribute the pay uniformly between the hours based on the trip time in each hour.[15] In Chicago, certain types of incentive earnings are paid for achieving weekly trips targets, rather than tied to individual trips. We spread these earnings uniformly across minutes worked in the week for which the incentive was earned. Hourly earnings are modestly higher in Chicago than in the national sample; the average driver in Chicago earns $23.81 per hour (gross of commission and other costs).

Before using regressions to decompose the gender earnings gap shown in Figure 1 (which looks nearly identical when looking only at Chicago), we examine average differences across gender in the factors that determine driver earnings. Recall from Equation 2.1 that driver earnings are a function of wait time between trips, distance to the start of the ride from where the driver accepts it, distance of the ride, speed (both on the ride and on the way to pick up the passenger), the surge rate at the time of the ride, and incentive payments.

[13] Analyzing a city at a time allows us to include fine controls for hour-of-week and geography at a driver-hour level. Using national data, this would be computationally impractical.

[14] Regressions are run on a 35% subset of drivers. Results are robust to different samples.

[15] If a driver only works for part of the hour, it is still included as an observation and hourly earnings are based on the fraction of the hour worked and the total earnings. Results are robust to including only full hours.

Table 3 displays the average of these parameters by gender. Note that these averages are presented on a per-trip basis, as that is a more natural way to divide some of the parameters. Table 3 also provides an idea of the sources of the gender pay gap. First, notice that the differences are generally small. Second, while the individual differences are small, nearly every one of the parameters favors men earning more than women. Men have shorter wait times for dispatch, shorter distances to the rider, longer trips, faster speed, and higher surge. The only parameter favoring women (slightly) is incentive pay; per trip, women earn about 3 cents more in incentives. We discuss differences in incentive pay in more depth in Appendix A.5.[16] The remainder of our analysis explores which of these differences in Table 3 are important factors of the Uber gender pay gap and what underlies the differences.

Moving to driver-hour level granularity allows us to control for certain features of a driver's behavior in a given driver-hour. We can now control for where a driver worked, the time of day and day of week, lifetime trips to-date, and whether the driver rejected a dispatch or canceled a trip that hour. To control for driving location, we track the "geohash" where a driver is located when he or she accepts a trip. A geohash is a geocoding system that divides the world into a grid of squares of arbitrary precision. For our case, we use geohashes that are approximately three miles by three miles.[17] We focus on the top fifty Chicago geohashes by trip density, which account for 89.2% of trips. The remaining trips are grouped into an "other" bin. For chains of UberPOOL trips, we only include the geohash of the first trip in the chain; drivers do not have control over where to locate for subsequent trips in the chain.

Table 4 refines the initial Chicago gender pay gap analysis we originally displayed in Table 2. However, whereas Table 2 utilized weekly observations (the hourly rate in that table is the average hourly rate for a driver in a week) to remain consistent with the regression models using the national data, Table 4 uses driver-hour observations. Column 1 of Table 4 reveals a baseline Chicago gender pay gap of 3.6% at the driver-hour level, controlling only for overall conditions in

[16]Equation 2.1 implies that trip distance and speed are ambiguously related to earnings; however, for the values of the other parameters that we observe in the data, earnings are almost always increasing in both distance and speed.

[17]Within busy areas of Chicago, this is a fairly large area and there may be differences in demand and congestion even within these areas that limit our ability to fully control for geographic effects. We have experimented with finer geographic areas and, given the conclusions do not change, we have not found this worthwhile given the additional computational complexity.

a given week.[18]

Before getting to the factors that explain the gender pay gap, we show in Column 2 of Table 4 that customer (that is, passenger) discrimination is *not* creating a gender gap in this setting. While the Uber rider/driver matching algorithm is gender-neutral, *customer* discrimination could contribute to Uber gender pay differences if riders disproportionately cancel trips when paired with a female driver.[19] After requesting a trip, riders see the name and a small image of the driver and can choose to cancel the trip. Drivers also see a rider's name (but not picture) after accepting a dispatch, so a gender pay difference could arise if drivers of one gender canceled rides of certain classes of passengers. Column 2 controls for canceling on both sides of the transaction and shows that discrimination has no effect on the gender coefficient suggesting that discrimination on either side of the market is not a primary cause of the pay gap.[20]

The entire gender pay gap is explained in Column 3 where we include measures of where drivers work (the geohash indicators), when (the hour-of-week indicators), driver experience buckets, and the log of driving speed. We can statistically rule out a gender gap in favor of either gender of greater than 0.6 percentage points. To our knowledge, no other paper has ever estimated such a precise "zero" gender gap in any setting. The remainder of our study focuses on explaining how the various controls in Column 3 contribute to erasing the non-trivial gender gap of approximately 3.6% with which we started.[21]

3.2 Where & When Drivers Work

Men and women drive at different times of the week and different locations across the city. Figure 2 shows the distribution of time spent driving across the 168 hours of a week; men drive more during

[18] This number is lower than the corresponding estimate in Table 2 because the weighting is by driver-hour rather than driver-week, effectively up-weighting drivers who work more hours in a week. This affects the measured gap for reasons similar to those we discuss below as we decompose the gap.

[19] Though many studies have hypothesized about customer discrimination and hypothesized that wage residuals may be due to customer preferences (especially race-based discrimination), prior work has not been able to conclusively establish if or when customer discrimination contributes to gender pay gaps.

[20] In the average driver-hour, total cancellation rates are statistically equivalent between men and women.

[21] Column 4 confirms that the gap is again unaffected by cancellations when controlling for other factors. It also shows that cancellations by either side of the market are, on average, costly for drivers.

Table 3: Parameter averages, Chicago only

	Men	Women	Difference (Men - Women)
w – Wait time (min)	5.857	5.920	-0.063
	(0.00158)	(0.00346)	
m_0 – Accepts-to-pickup distance (mi)	0.569	0.580	-0.011
	(0.00044)	(0.00054)	
m_1 – Trip distance (mi)	5.108	5.070	0.038
	(0.00098)	(0.00223)	
s – Speed (mph)	18.262	17.634	0.628
	(0.00152)	(0.00333)	
SM – Surge multiplier	1.116	1.105	0.011
	(0.00005)	(0.00010)	
I – Incentive payout ($)	0.594	0.624	-0.030
	(0.00026)	(0.00062)	

Note: This table documents averages for men and women of the parameters in Equation 2.1. Averages are per-trip based on trips completed in Chicago. For wait time and accepts-to-pickup distance, averages are based on trips from May 2016 - March 2017 due to limitations in the underlying raw data. All other averages are based on data for the entire sample. Wait time is based on time between either coming online or completing previous trip and picking up passenger for new trip. Trip distance is based on actual route taken; however, accepts-to-pickup distance is the Haversine distance between corresponding coordinates. The gender composition of drivers changes over time; to correct for this, we re-weight observations in each week of data by (total trips)/(trips by that gender). Standard errors are reported in parentheses.

Table 4: Gender pay gap

	(1)	(2)	(3)	(4)
isMale	0.0356	0.0355	−0.0018	−0.0018
	(0.003)	(0.003)	(0.002)	(0.002)
riderCancellations		-0.0091		-0.0238
		(0.000)		(0.000)
driverCancellations		0.0078		-0.0158
		(0.003)		(0.002)
Intercept	3.0862	3.0869	1.7346	1.7452
	(0.003)	(0.003)	(0.003)	(0.004)
Driver experience			X	X
Log driving speed			X	X
Week	X	X	X	X
Hour of week			X	X
Geohash			X	X
N	11,572,163	11,572,163	11,572,163	11,572,163
R^2	0.039	0.039	0.266	0.267

Note: This table documents both the base gender pay gap and the gender pay gap once controlling for experience, location, time, and speed. Data are at the driver-hour level. Further, it includes specifications with rider and driver cancellations. The outcome variable is log of hourly earnings. Hour of week controls for each of 168 hours. Geohash controls are a vector of dummies for whether a driver began a trip in a given geohash. Driver experience is measured by a driver's lifetime trips completed prior to a given date, where lifetime trips is binned into 0-100 trips, 100-500 trip, 500-1000 trips, 1000-2500 trips, and >2500 trips. Driving speed is the speed driven while on trip in a given driver-hour. Standard errors (clustered at the driver-level) are in parentheses.

the late night hours, while women drive substantially more on Saturday and Sunday afternoon.[22] The first panel of Figure 3, which maps the fraction of trips in a given geohash that are completed by men, shows that men are more likely to complete trips in the more Northern parts of the city. These differences in driving habits—whether due to inflexible schedules, preferences and constraints, or differential costs to driving in certain locations (e.g., far from home)—may contribute to the observed gender pay gap.

Table 5 starts to break down the baseline pay gap of 3.6% in column 1 of Table 4. Column 1 adds 168 indicator variables for the hour of week, which eliminates 14% of the gender pay gap. This suggests that, while the variation in preferences or constraints for driving hours documented by Chen et al. (2019) may be correlated with gender, hour-within-week differences are a small part of the gender gap. If female drivers receive more non-pecuniary benefits than men from picking which hours to work, they do not pay a large financial price for this flexibility.

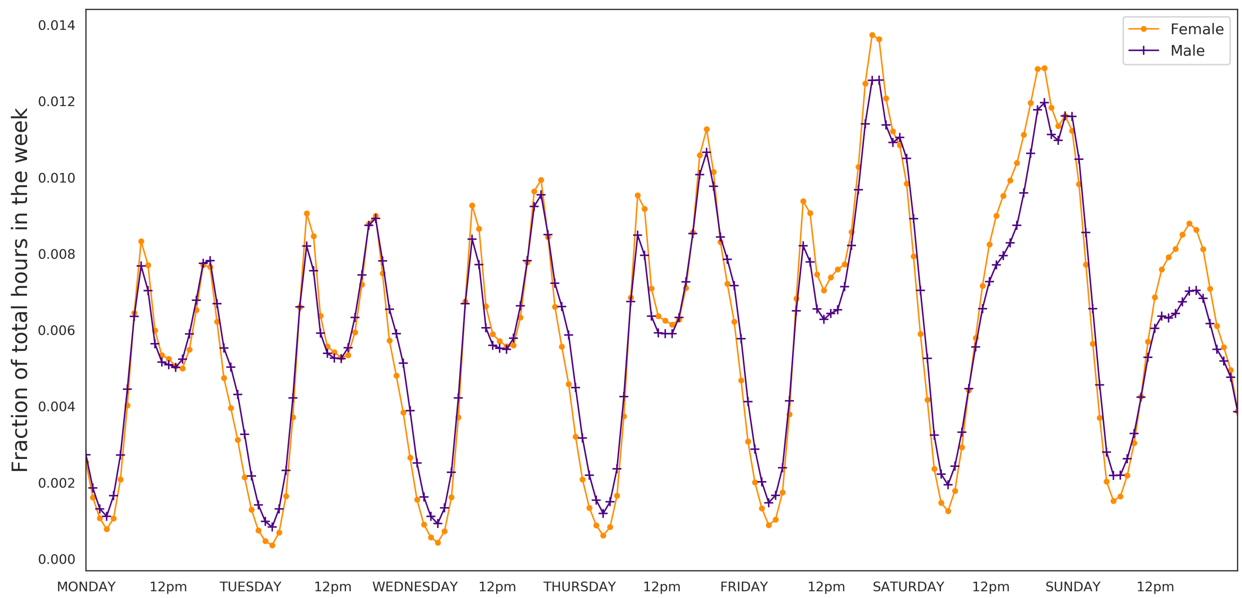
Column 2 of Table 5 adds controls for the top fifty Chicago geohashes. This removes about a quarter of the gender pay gap, indicating that men drive in the parts of Chicago where pay is higher due to factors such as higher surge and shorter waiting times. Per Column 3, the "where and when" variables together attenuate the gender earnings gap by about a third.[23]

Overall, the first three columns of Table 5 show that time and location explain some, but not all, of the gender earnings gap. The remaining gender earnings differential of 2.2% is small compared to overall gender pay gaps measured in the literature, but it is substantial given we are exploring workers doing exactly the same job at the same time and location and being paid by a gender-blind algorithm.

[22] Figure 15 plots the average earnings for men and women across each hour of the week; earnings generally vary between $20-$28. For any hour of the week, men average more per hour than women driving in the same hour.

[23] The order in which variables are added to the regression affects the contribution of the variables to eliminating the gender gap because of collinearity among these variables. Below, we use a method to decompose the gender gap that is invariant to the order in which variables are added and that accounts for collinearity. Also, specifications including the interaction of location and hour had little additional explanatory power, suggesting that the hour of week earnings differentials are fairly consistent across areas of Chicago.

Figure 2: Distribution of hours of the week worked by gender



Note: This figure shows which hours of the week men and women work; each point represents the fraction of their total hours in the week that men (or women) spend working in that specific hour of the week. Data are limited to Chicago UberX/UberPOOL drivers in Chicago, January 2015-March 2017.

3.2.1 Features of Driving Locations

Features of locations—such as safety, likelihood of picking up an intoxicated rider, and proximity to a driver's home—may impact a drivers' propensity to drive there and may do so differentially for men and women. To investigate this, we construct a dataset at the geohash-level with data on crime levels, businesses with liquor licenses, the gender of drivers living nearby, and gender of the overall adult population. Due to limitations in the availability of crime data, we restrict our main data to driver-hours that include a trip within the City of Chicago.[24] Per Column 5 of Table 5, the baseline gender pay gap in this subset is slightly larger than the overall population – 4.3% versus 3.6%. A detailed description of the data construction is available in Appendix A.6.

Figure 3 maps the percentage of trips beginning in a geohash that are by male drivers. There is considerable variance in the percentage of trips completed by men in a given geohash; in the North parts of Chicago, men often complete >85% of trips compared to ∼70-80% of trips in the South and West sides of Chicago. Figure 3 also maps various features of the geohashes that may

[24]This limits us to 68.8% of our original observations.

correlate with driving location: the percentage of drivers living nearby who are men, the gender divide in the adult census population, the number of crimes per 1,000 adult residents, the number of liquor licenses, and the median household income. Most notable are the similarities between home locations and where men and women drive. The locations where Uber trips are predominantly by male drivers are also the locations where the population, both of Uber drivers and of all adults in the Census, is skewed more male.[25] The locations with more female trips (and a higher percentage of female residents) also face higher crime rates. The differences in home location (and the level of segregation in Chicago) suggest that the racial composition of male and female drivers may be substantially different. In Appendix A.7, we show that this is the case—female drivers, for example, are nearly twice as likely to be Black—but differences in race do not qualitatively affect the gender pay gap. See Appendix A.7 for analysis of the gender gap and its contributors while controlling for driver race.

Given the patterns in home location, education may also differ by gender. We do not observe education in our data, but Hall and Krueger (2018) find in a survey of Uber drivers that education has no detectable effect on reported hourly earnings. While we cannot rule education out as a confounding factor in our analysis, we would expect that the (at most) small returns to formal education limit the effect of any unobservable differences in education between men and women on the gender pay gap in our sample.

To further investigate differences in the locations that men and women drive, we regress the log share of male trips against various features of a geohash. Results are presented in Appendix Table 13. Absent controls for home location, women drive in areas correlated with higher crime; however, once controlling for home locations, an increase in either crime or liquor licenses is correlated with a *decrease* in the share of women driving in that location. Women appear to avoid locations that may be unsafe, either due to crime or more intoxicated riders. However, safety considerations are secondary to where drivers live. Controlling for driver home locations alone has far greater explanatory power; drivers work close to where they live. This result is not unique to Uber;

[25]The fact that some neighborhoods are so strongly male or female has not been documented previously to our knowledge. This variation in gender composition by neighborhood, we have found, is not unique to Chicago and has been true for at least several decades. Neighborhoods skew male or female across age and racial groups. While this is an interesting fact with broad implications, further investigation is beyond the scope of our paper.

individuals in traditional labor markets may also work close to home due to the pecuniary costs and disutility associated with commuting and these effects may differ by gender (Madden (1981)).

Columns 6 and 7 of Table 5 show the results of earnings regressions including the average crime rates and number of liquor licenses in the geohashes where drivers begin a trip in a given hour. Surge pricing that incentivizes drivers to go to areas where supply is low relative to demand generates a small compensating differential for working in areas with higher crime or more bars; a 10% increase in the number of crimes is correlated with a 0.43% increase in pay and a corresponding increase in number of liquor licenses is correlated with a 6.75% increase in pay.[26] Further, when controlling only for driver experience, hour of week, driving speed, driver home locations, and the *features* of geohashes instead of the actual geohashes (Column 7), the gender gap is statistically indistinguishable from zero. In fact, controlling just for driver home location—on top of when, experience, and speed—in the full Chicago sample results in a gender gap indistinguishable from zero (Column 4).

These results show that the lower costs associated with driving near one's home are an important factor in where drivers operate. They also show results consistent with women having a stronger preference than men for avoiding areas with a higher incidence of crime or where there is a higher likelihood of picking up intoxicated passengers. This may stem from safety risks that disproportionately affect women. These differences negatively affect earnings for women on Uber, as there are small compensating differentials for driving in areas with higher crime rates or more liquor licenses. Overall, however, residential location of drivers appears to be a much more important determinant than safety considerations for determining where drivers work.

3.3 Returns to Experience

The gender earnings gap generally rises with workers' years of experience (Altonji and Blank (1999)). However, measures of experience in traditional datasets tend to be quite coarse. Often, we only observe years since graduation from school or years employed in a given profession as the best metric of experience. Measurement error may lead to attenuated experience effects. One

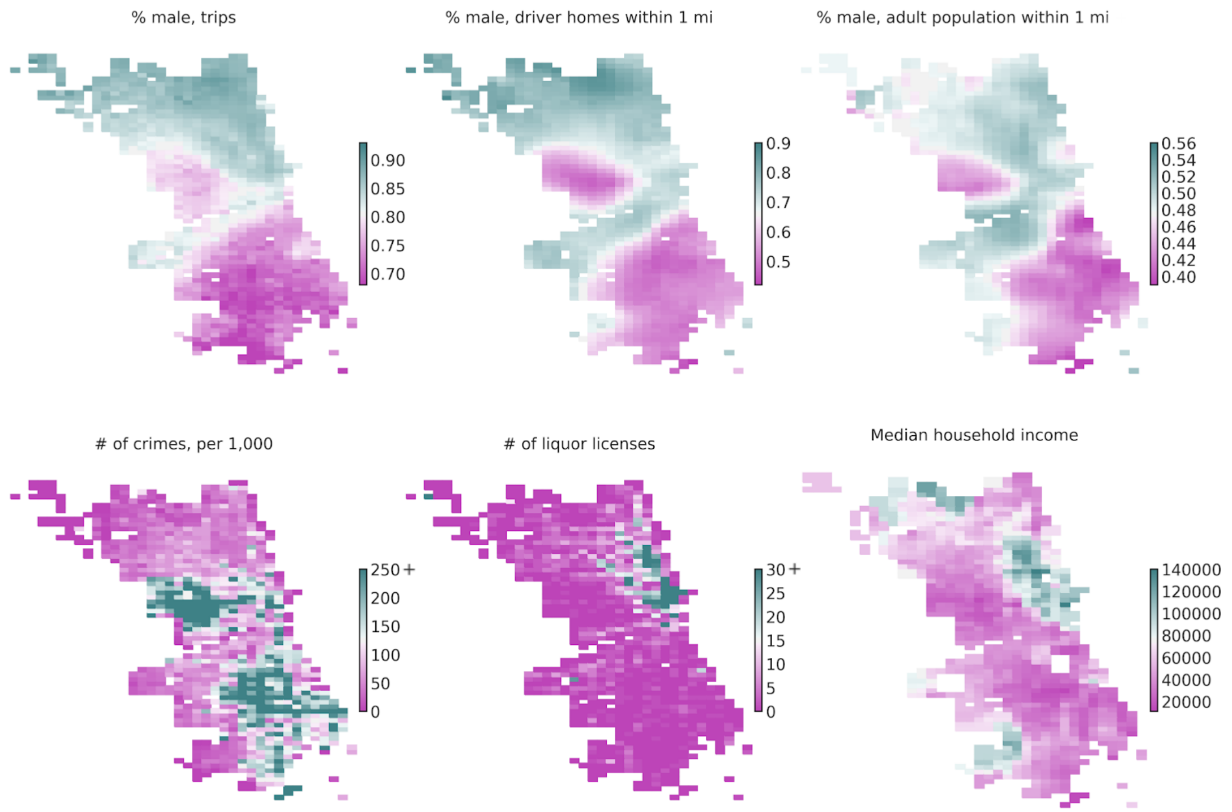
[26] See Appendix Table 14 for the mean and standard deviation of the various geohash features.

Table 5: Returns to driving time and location

	All Chicago data				City of Chicago only		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
isMale	0.0302	0.0261	0.0220	-0.0019	0.0434	0.0080	0.0026
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
Log crimes per 1,000pp						0.0004	0.0043
						(0.001)	(0.001)
Log liquor licenses						0.0699	0.0675
						(0.001)	(0.001)
Intercept	3.0912	3.0946	3.0980	1.7169	3.1199	1.7244	1.7117
	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)	(0.005)	(0.004)
Week	X	X	X	X	X	X	X
Hour of week	X		X	X		X	X
Geohash		X	X		X		
Experience bins				X		X	X
Log speed				X		X	X
Driver home geohash				X			X
N	11,572,163	11,572,163	11,572,163	11,572,163	7,969,988	7,969,988	7,969,988
R^2	0.099	0.092	0.143	0.272	0.062	0.301	0.306

Note: This table documents the evolution of the gender pay gap as time and location covariates are added. Data are at the driver-hour level. The outcome variable is log of hourly earnings. Hour of week controls for each of 168 hours. Geohash controls are a vector of dummies for whether a driver began a trip in a given geohash. The "City of Chicago" refers to the Chicago area for which crime data are available. Standard errors (clustered at the driver-level) are in parentheses.

Figure 3: Features of geohashes



Note: This figure maps various features at the geohash-level for the City of Chicago. The distribution of trip locations is based on where trips originate. The geohashes used are more precise than those used in regressions, measuring about 0.75 miles on each side. Population numbers—both driver home locations as well as total adult population from the 2016 ACS—are smoothed by measuring population within one mile of a given geohash. Crimes include all non-residential crimes and are normalized by the number of crimes per 1,000 adult residents. Liquor licenses are based on number of unique businesses with a liquor license active during our time sample in a given geohash. Median household income is from the 2016 ACS. For crime and liquor licenses, the distributions are winsorized at 250 and 30, respectively, to allow for more informative coloring.

of the unique aspects of working with Uber data is that we can measure a driver's experience level (number of previous rides given) with high precision.

Indeed, there is much to learn being a driver on Uber. Uber pays according to a fixed formula, but many of the parameters of the formula (that is, the variables listed in Table 3) are within the driver's control. For example, drivers can indirectly affect the surge multiplier and wait times by choosing where and when to work and directly affect their driving speed by simply driving faster. As drivers work more, they can begin to learn optimal driving behaviors to maximize earnings.[27]

[27] Another activity that may generate a return to experience is "dual-apping," which is when drivers accept trips from both Uber and a competitor (primarily Lyft). Dual-apping has the potential to increase earnings due to less

As a result, none of the increased earnings with experience comes from a pre-set pay schedule that "mechanically" raises pay with experience. Any experience premium results from learning and increased driver productivity.[28]

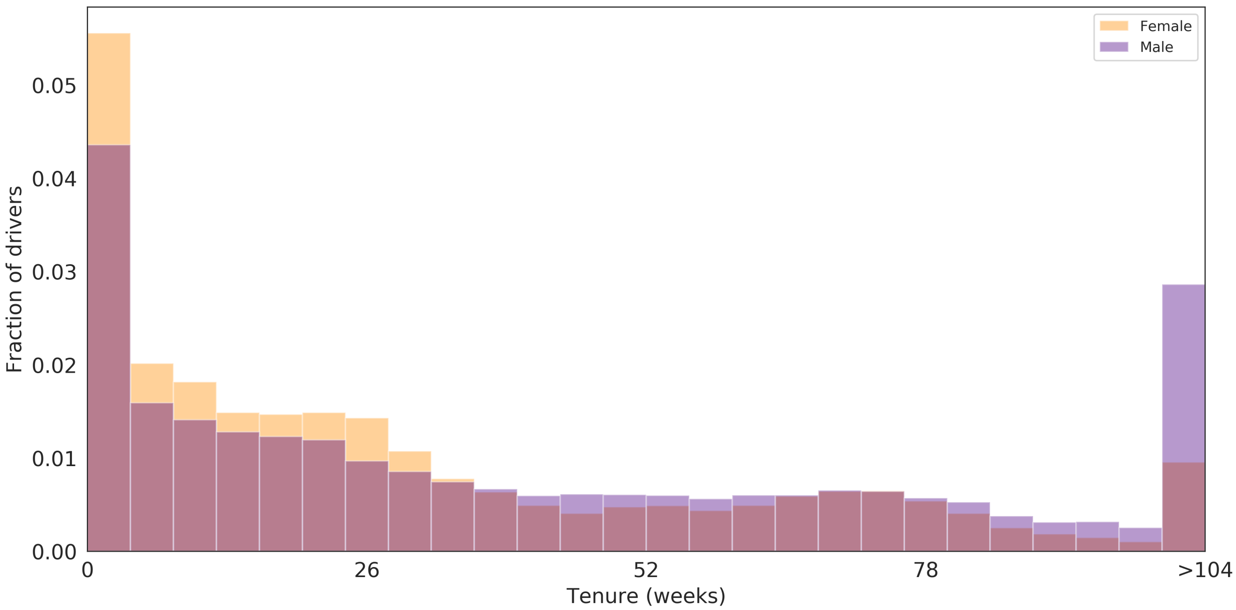
Figure 4 provides a visual indication of how returns to experience can affect the gender earnings gap. The figure, which shows the average tenure of all drivers with a completed trip in January 2017, reveals that men are far more likely to have been driving on Uber for over 2 years. Women are likely to have joined in recent months. Further, Figure 5 shows that men accumulate completed trips at a faster rate per week than women. Since women supply fewer hours of labor per week, they accumulate experience more slowly per week.

Figure 6 demonstrates the raw driver returns to experience as measured by cumulative number of trips driven. There is a clear learning curve, which is especially steep early in a driver's tenure. Drivers continue to learn valuable skills on the job through at least 2,500 trips with a fully experienced driver earning about \$3 per hour (more than 10%) more than a driver in his or her first 500 trips. In principle, the rise in earnings shown in Figure 6 could be a selection effect if drivers' baseline productivity level is correlated with lasting longer on the Uber platform. We investigate this in detail in Appendix A.9. We find that estimating the experience curve using only within-driver variation leads to a slightly *steeper* learning curve, especially in the range of experience over 2,000 rides. This suggests that the riders who choose to work for Uber intensively are lower productivity than average, indicating they may have worse outside employment options than those who do not accumulate a large amount of experience. Since this bias has little impact on the gender wage gap, and including driver fixed effects complicates estimation of the gender gap (it is co-linear with the fixed effects), we keep this analysis in the appendix and focus on prior rides as our measure of experience. In Appendix A.8, we also show that men and women do not learn at different rates as they accumulate experience.

time waiting for a dispatch and the ability to filter higher-value trips if the surge multiplier differs across platforms. We do not have a credible way to determine the degree to which this affects earnings nor whether specific drivers are dual-apping, so we cannot isolate dual-apping's contribution to the return to experience.

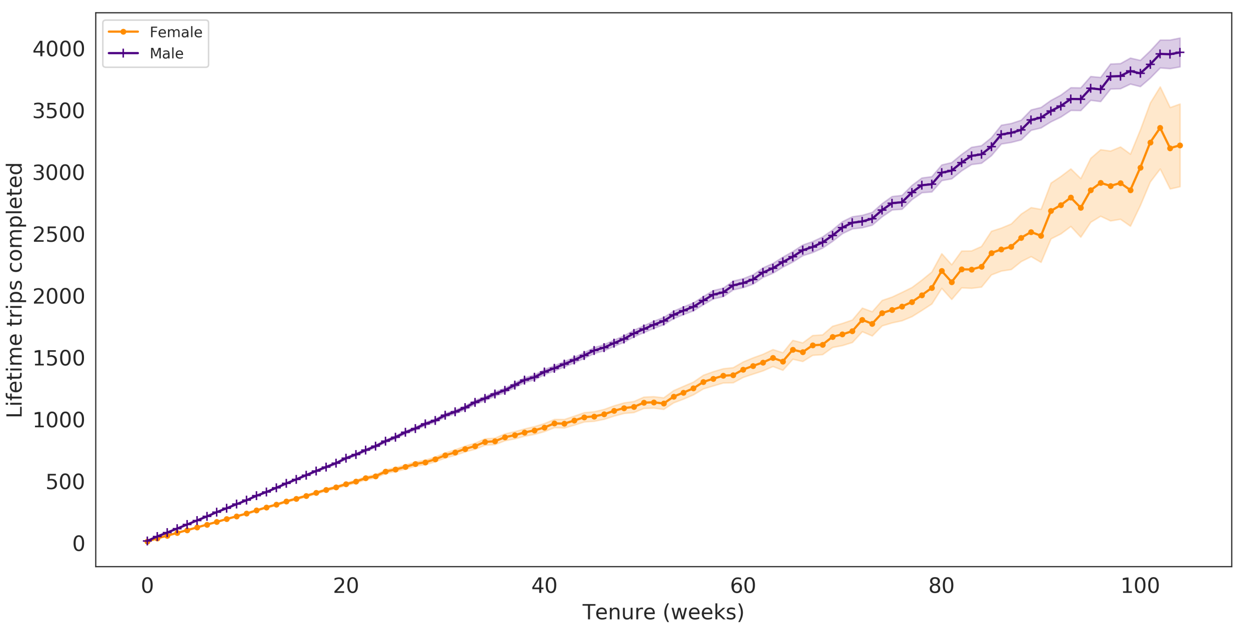
[28]Haggag et al. (2017) show that learning-by-doing and experience are important for New York City taxi drivers. While drivers on Uber may learn in some ways similar to taxi drivers, there are likely important differences. For example, Uber rates fluctuate with surge prices (unlike fixed taxi fares), Uber uses an assignment algorithm to offer trips to drivers, drivers on Uber use in-app GPS, and drivers are not customarily paid a tip on Uber (during the time period of our data).

Figure 4: Distribution of driver tenure, January 2017



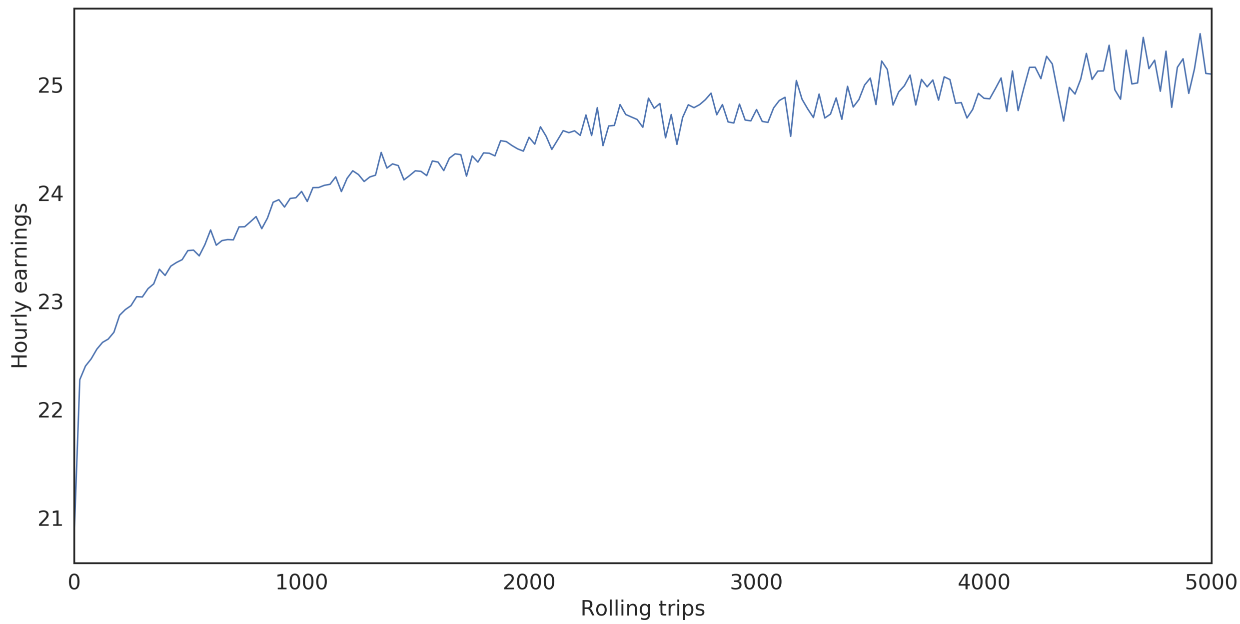
Note: This figure shows the average weeks of tenure for drivers that completed a trip in January 2017; we limit to a single month to avoid composition effects. Tenure is measured as the number of weeks since a driver's first completed trip.

Figure 5: Accumulation of trips over weeks of driving



Note: This figure shows the average number of lifetime trips completed for drivers of a certain tenure. Tenure is based on the number of weeks since a driver completed their first trip. The data only include driver-weeks with >0 trips.

Figure 6: Returns to experience



Note: This figure shows the average earnings of drivers with a given number of rolling trips completed prior to a day of work; rolling trips are binned into buckets of 100 trips completed. Data include all Chicago drivers from January 2015 to March 2017.

In Table 6, we return to our earnings regression and show that there are substantial returns to experience on Uber. Column 1 shows that drivers who have completed over 2,500 trips make nearly 14% more than those in their first 100 trips. Gender differences in average experience are clearly important as, controlling for experience, the gender earnings gap shrinks to 1.4% or roughly a third of the initial earnings gap in Chicago.[29]

With controls for hour of week (Column 2), the gender gap is further reduced to under 1%, but the returns to experience do not change noticeably. On the other hand, controls for driver location (Column 3) do not reduce the gender gap but substantially reduce the returns to experience. Combined, these two columns suggest that the primary effect of experience on earnings comes from learning *where* to drive and that men and women have differences in terms of their preferences/constraints for *when* to drive.[30]

[29]These five bins of experience capture the relevant value of experience. We have experimented with other parametric forms of experience in these regressions and the results are qualitatively similar.

[30]One behavior drivers learn beyond where and when to drive is how to strategically reject and cancel rides. We limit our discussion of this behavior to Appendix A.11 because it does not affect pay differentially by gender.

In all of our analyses, we find no gender differences in the actual learning process. Rather, learning affects the gender gap because, though each additional ride teaches men and women the same valuable skills, men accumulate driving experience faster than women.

Table 6: Returns to experience

	baseline	(1)	(2)	(3)	(4)
isMale	0.0356	0.0138	0.0083	0.0129	0.0085
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Trips completed: 100-500		0.0530	0.0497	0.0357	0.0334
		(0.001)	(0.001)	(0.001)	(0.001)
Trips completed: 500-1000		0.0773	0.0747	0.0512	0.0494
		(0.002)	(0.002)	(0.002)	(0.001)
Trips completed: 1000-2500		0.1001	0.0990	0.0650	0.0648
		(0.002)	(0.002)	(0.002)	(0.002)
Trips completed: >2500		0.1391	0.1390	0.0877	0.0890
		(0.004)	(0.003)	(0.003)	(0.003)
Intercept	3.0862	3.0228	3.0294	3.0528	3.0570
	(0.003)	(0.002)	(0.001)	(0.003)	(0.001)
Week	X	X	X	X	X
Hour of week			X		X
Geohash				X	X
N	11,572,163	11,572,163	11,572,163	11,572,163	11,572,163
R^2	0.039	0.048	0.107	0.096	0.146

Note: This table expands on the regressions in Table 4 by adding controls for a driver's experience. Experienced is measured as trips completed before a given day of work. Drivers with fewer than 100 completed trips are the excluded category. The outcome variable is log of hourly earnings. Standard errors (clustered at the driver-level) are in parentheses.

3.3.1 Experience and the Long-Hours Premium

In other settings, a gender gap has been shown to grow over time as women accumulate fewer hours of on-the-job experience (e.g., Bertrand et al. (2010)). In most of those settings, however, men are working more hours in each week *and* they have accumulated more experience, making it difficult to empirically distinguish between the value of accumulated experience and work intensity. While our setting differs from the more professional settings of prior studies, we do have similar patterns in the value of accumulated experience and we can empirically distinguish between the role of intensity and past experience.

Figure 7 shows the results of pay regressions similar to those in Table 6 but we add a new variable for hours worked in the week. To purge the effects of driving intensity from unobserved demand effects (such as a convention or big event in town), we instrument hours worked in the week with average hours in previous weeks (which requires us to drop drivers in their first week on the Uber platform). We estimate a cubic in hours worked using 2SLS in the pay regression. The graph shows how the driving intensity/earnings relationship changes as we control for different factors.[31]

In our first specification, we control for "potential experience," as is typically done, with driver age and its square.[32] The results of this regression suggest an increasing return to work intensity as hourly earnings increase substantially up to thirty hours per week (which is approximately the 85th percentile of driver weeks) and then flattens out or declines slightly. The upward sloping relationship between hourly wage and hours worked shows what looks like a long hours premium for being an Uber driver. But when we control more accurately for relevant experience by adding our accumulated trips measures, the returns to work intensity turn sharply negative. All else equal, a driver earns a few percentage points lower hourly wage in the thirtieth hour of driving in a given week than in the first twenty. This shows that, at least for Uber drivers, there is significant financial value in accumulated experience and a mildly decreasing return to within-week work intensity. More generally, it shows that what might appear as a convex hours/pay relationship when using conventional controls for experience could be masking a return to true experience when there is no return to work intensity.

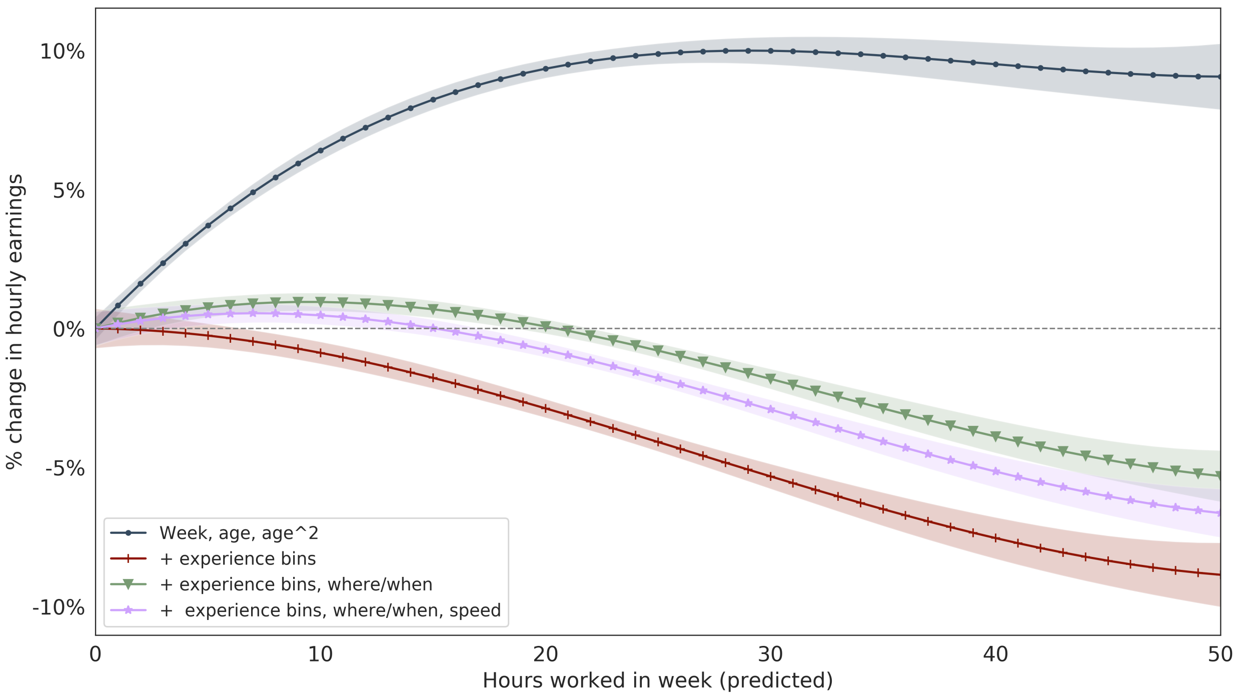
Even in this short-term gig economy environment, experience and gender differences in experience play out in a way that contributes substantially to the gender pay gap and is not related to work intensity per se. On balance, the relationship between experience and the gender pay gap for drivers is surprisingly similar to at least some professional job environments and suggests that, in those settings, it's possible that the convex hours/earnings relationship is overstated because on-the-job learning effects cannot be measured accurately (for examples of convex hours/earnings relationships, see Bertrand et al., 2010; Goldin, 2014). Indeed, distinguishing the difference between

[31]In Appendix Figure 16, we show that the pay gap does not vary in an economically meaningful way across different levels of working intensity. When considering just point estimates, it appears that the gap is marginally smaller for drivers working more hours per week.

[32]Actually, a standard control would be age-education-6 but we do not observe education and do not think it is of first-order importance for Uber drivers.

the long hours premium and on-the-job learning in the broader economy is quite policy relevant. Policies that could make jobs more flexible and lower the premium to working long hours would have little effect on the gender wage gap if a substantial contributor to the correlation of longer hours and high pay is the omitted variable of experience. While Uber drivers are only a tiny slice of the broader labor market, they offer a unique lens into these issues since we can both measure experience and productivity on the job with high precision.

Figure 7: Returns to Work Intensity



Note: This figure graphs the effect on earnings of different values of driving intensity, defined by how often they are predicted to drive in the week of observation (predicted based on hours driven in past weeks) using a regression of log hourly earnings on cubic predicted hours worked. Drivers in their first week of work are not included. “Experience bins” and “where/when” controls are the same as those used throughout the paper. Shaded region represents 95% confidence interval based on standard errors clustered at the driver-level.

3.4 Returns to Speed

As shown in Equation 2.1, drivers earn a per-minute and a per-mile rate on each trip. In some unusual circumstances, there are negative returns to speed as the per-minute rates can be relatively valuable if the driver expects to wait a long time until getting another fare. In general, however, the rates and wait times for Uber drivers are such that there is a positive expected return to driving faster. This return is somewhat higher when driver wait times are shorter. At extreme speeds, the returns to speed net of costs may turn negative if the risk of a collision or a speeding ticket becomes high enough.

Given that prior research suggests men are more risk tolerant and aggressive than women (see Bertrand (2011) and, in the context of driving, Dohmen et al. (2011)) and that Table 3 shows that male drivers drive faster than women, we now investigate how driver speed affects the gender earnings gap. We measure speed as distance on trip divided by time on trip in a given driver-hour. This measure of speed captures both higher driving speed on a given route plus choice of routes that allow for faster speeds (e.g., highways). Table 7 adds the log of speed as an explanatory variable to our earlier hourly pay regressions.

Control variables are important in this regression, because higher pay areas and times of week in Chicago (those areas where there is a more constant stream of fares and where surge is likely to be higher) are also likely more congested, which lowers speed. The coefficient on log speed in Column 1 of Table 8 suggests an elasticity of 27% of speed on earnings; a 1% increase in speed increases earnings by 0.27%. In Column 2, when we control for geohash and hour of week (thus removing the fact that congestion both lowers speeds and increases earnings), this number increases to 46%. Column 2 shows that controlling for speed and neighborhood reduces our original 3.6% gender pay gap all the way to just 1%. Adding the learning-by-doing experience variables to this model fully eliminates the gender pay gap.[33]

We believe that we can describe this speed difference across gender primarily as a difference in preferences that happens to have a productive value on Uber rather than a response by male drivers to the incentive to drive faster.[34] First, as mentioned above, others have shown that men are more risk tolerant, both in general and when driving in particular. Second, when we analyze Uber driver speed as a function of gender, experience, and time/location, we find that men drive 2.2% faster than women. Further, speed is only slightly increasing in experience (and experience

[33]The estimate in Column 4 is very precise; the gender pay gap has a 95% confidence interval of -0.6% to 0.2%.

[34]Speed is productive in that it generates earnings for both Uber and the driver on any given ride. It also may generate at least a small long-term value for Uber (and a positive externality on other drivers) because passenger ratings of drivers are increasing in driver speed, holding other factors constant. This relationship is highly significant statistically but small in magnitude.

does little to close the gender speed gap); if drivers were responding strongly to the incentive to drive faster, we might expect speed to increase substantially with experience.[35]

In addition, we gathered data from the National Highway Travel Survey a nationally-representative survey that gathers demographics, vehicle ownership, and "trip diaries" from 150,000 households. Outside of Uber, there is rarely a pecuniary incentive to drive faster. Despite this, we find that men still drive faster in the NHTS sample (details in Appendix Table 21). Gender differences in the preference for speed are a general population phenomenon that have labor market value to drivers.

Table 7: Returns to speed

	baseline	(1)	(2)	(3)	(4)
isMale	0.0356	0.0256	0.0106	0.0016	−0.0018
	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)
logSpeed		0.2677	0.4552	0.2715	0.4544
		(0.002)	(0.001)	(0.002)	(0.001)
Trips completed: 100-500				0.0563	0.0318
				(0.001)	(0.001)
Trips completed: 500-1000				0.0819	0.0460
				(0.002)	(0.001)
Trips completed: 1000-2500				0.1075	0.0599
				(0.003)	(0.002)
Trips completed: >2500				0.1519	0.0831
				(0.004)	(0.0003)
Intercept	3.0862	2.3084	1.7704	2.2293	1.7346
	(0.003)	(0.003)	(0.004)	(0.005)	(0.004)
Week	X	X	X	X	X
Hour of week			X		X
Geohash			X		X
N	11,572,163	11,572,163	11,572,163	11,572,163	11,572,163
R^2	0.039	0.101	0.263	0.111	0.266

Note: The table expands on earlier regressions by adding log speed as an explanatory variable. Speed is based on total trip distance and duration in a given driver-hour. The outcome variable is log of hourly earnings. Standard errors (clustered at the driver-level) are in parentheses.

[35]We also examined whether speed is correlated with wait times (when wait times are lower, there is a stronger incentive to speed to start the next trip), but found no evidence that men differentially respond to low wait times by increasing speeds. We discuss speed further in Appendix A.12.

4 Summarizing the Decomposition

Using standard pay regressions, we have fully explained the gender earnings gap for drivers on Uber. The raw gap in Chicago of approximately four percent can be attributed to three factors: male preference for faster driving, time and location choices of drivers, and higher average male on-the-job experience.

To measure the extent to which each of these factors contributes to the gender pay gap, we follow the approach described in Gelbach (2016).[36] Conceptually, this approach treats each factor as an "omitted variable" in the relationship between earnings and gender and measures the bias that would result if the factor were excluded. This allows us to disentangle the impact on the gender gap of each factor we controlled for sequentially in the above section, invariant of the order in which we initially added them into our baseline regression specification. This approach is of particular value when our observables are correlated. For example, our measure of driving speed is likely endogenous with where/when a driver works such that the difference in the point estimates of the pay gap with and without controlling for speed is also capturing differences in where/when drivers work.

More precisely, consider a regression of the form

$$\ln(Earnings_{dt}) = \beta \ isMale_d + \gamma_v X_{vdt} + \gamma_2 X_{2dt} + \epsilon_d \tag{4.1}$$

where X_v is a single vector for variable v and X_2 captures all remaining variables in our full model (i.e. speed, experience indicators, time indicators, and location indicators). Now suppose we ignore information contained in X_v. The resulting omitted variable bias is given by $\hat{\pi}_v = \hat{\Gamma}_v \hat{\gamma}_v$ where $\hat{\Gamma}_v$ is estimated using an auxiliary regression of gender on X_v.

Dividing our estimate of omitted variable bias by $\hat{\beta}^{base}$, the baseline relationship between earnings and gender conditioning only on calendar week, provides us an estimate of the variable's

[36]See Grove et al. (2011) for use of the Gelbach decomposition in the context of gender wage gaps. See Allcott et al. (2019), Buckles and Hungerman (2013), or Chandar et al. (2019b) for other examples of the Gelbach decomposition in practice.

contribution to the gender pay gap as a fraction of the baseline, unconditional relationship:

$$\tilde{\pi}_v = \frac{\hat{\pi}_v}{\hat{\beta}^{base}} \tag{4.2}$$

These contributions can be aggregated across vectors of variables, such as each of 168 indicators for hour of week, to obtain the combined contribution of controlling for all hour of week indicators. We do this for hour of week (when), geohash (where), bins of experience, and speed. These are correlations and should not be interpreted as the causal effect on the pay gap of, for example, increasing speed or experience.

A very desirable property of the Gelbach decomposition is that it does not depend on the order in which controls for each attribute are added. Sequentially controlling for more attributes does not allow for a simple decomposition, since the amount the gender gap changes with each additional set of controls depends on what other controls have already been added to the regression. However, the results of the Gelbach or any other decomposition method will depend on *which* covariates are being included in the decomposition. For example, consider the case where experience is a key factor of the gender wage gap, but that the only factor drivers learn about is where the best locations are to drive. Ultimately, experience does not enter the pay formula directly, only driving location does through surge pricing and wait times. Then, if we were to control only for experience in the gender wage regression, the Gelbach method would attribute 100% of the gap to experience differences between men and women. If we then added location fixed effects to the regression and repeated the decomposition (now including both experience and driving location), we would find that location would explain all of the wage gap, and that experience *given driving location* would explain none. Mechanically, we know that the only variables we should need to control for to fully explain the gender gap are those that directly enter the pay formula, along with enough non-parametric flexibility. The mechanical effects of features such as driver wait times and trip distance do not highlight the economic mechanisms at play driving these differences by gender. Indeed, experience does not directly enter the pay formula, but quantifies the conceptual role of learning-by-doing. Our decompositions below focus on the economic forces causing men and women to make different choices in the Uber labor market.

To begin, we first examine the role of fixed worker characteristics (place of residence) and experience in explaining the gender pay gap. By removing controls for where, when, and speed, we allow experience and home location to explain the gender gap through these mitigating channels. The figure, which is the lower graph in Figure 8, indicates that driver experience alone can explain 50% of the gender pay gap, while place of residence can explain 10%. Without additional controls, 40% of the gender gap remains unexplained.

We then return to our baseline regression that includes experience, where, when, and the speed of driving. We repeat the decomposition. The first panel of Figure 8 presents the parameter estimates of Equation 4.2, along with 95% confidence intervals, corresponding to a decomposition of the change in point estimates between our baseline model, which includes only controls for the week of the data (see Column 1, Table 4), and a fully specified model with controls for speed, location, time of week, and experience (see Column 5, Table 7). Speed alone explains nearly half of the gap (48%). Experience can explain the next largest share, at 36%. Without these additional controls, experience explained 50% of the gap, indicating that learning about where, when, and the speed to drive explains about 14 percentage points of the experience effect. Where drivers work can explain a further 28% of the gap, while time of week—once conditioning out the other factors—actually offsets the pay gap (-7%).[37] This suggests that while women may choose to drive at different times of the week than men, they do not pay a steep penalty for this flexibility. The attenuation in the gender pay gap observed when hour of week controls are included (Table 4) is due to factors, such as experience and driving location, correlated with when drivers work. Together, these factors fully explain the driver gender pay gap.[38]

To further unpack the mechanisms of driver location differences by gender, we remove the location fixed effects, and replace them with geohash of driver residence fixed effects and controls for crime rates and liquor licenses in the geohash of pickup. Note that the gender gap also goes to zero with these location controls, even without the geohash fixed effects. The second graph in Figure 8 repeats the Gelbach decomposition using this new regression. We find that 20% of the gender

[37] In Appendix A.7 we repeat this analysis controlling for driver race and find very similar estimates.

[38] The results sum to slightly greater than 100% as the point estimate on $isMale$ is (insignificantly) negative after controlling for each of the covariates.

wage gap can be attributed to differences in where male and female Uber drivers live. As previously discussed, neighborhoods with a high adult female population share tend to be less lucrative areas, and since driver location is skewed towards one's home, women earn less. In addition, womens' decisions to avoid neighborhoods with high crime and many bars contributes an additional 29% to the gender wage gap. Even though Uber pay is "gender blind," if women are disproportionately unsafe in high crime areas or when riders are less likely to be sober, they pay for it through a compensating differential.

Gender differences in drivers' home location can explain 10-20% of the gap. This is consistent with the "spatial mismatch" hypothesis that workers living far from "good jobs" earn less due to the commute cost. While most of the spatial mismatch literature has focused on how racial segregation leads to spatial mismatch, we find evidence that gender segregation across neighborhoods also contributes to the gender wage gap. These decomposition results are robust to controlling for drivers' race and using within-driver variation to estimate the return to driver experience. See Appendix A.7 for the race results and Appendix A.9 for the experience results.

To further identify the underlying sources of the differences in pay by gender, we return to our table of averages of all the parameters that enter into driver earnings, as described in Equation 2.1. Table 8 shows the average of each parameter by gender for drivers of three different levels of experience.

The table highlights three important themes from our analysis. First, both men and women learn in a productive manner and at roughly the same rate in terms of number of rides. The wait times go down by about 15% over 1,500 rides of experience. Surge rates improve and are nearly identical for both genders. Men have slightly longer pickup distances and ride distances throughout, but both genders lower pickup distances and increase trip distances in a similar manner.[39]

Speed is an outlier in that there is not a clear "improvement" over time for drivers. In fact, drivers appear to drive more slowly as they gain experience, though this is likely because drivers learn that more congested areas are more lucrative. As per our regressions, there is a noteworthy (if not huge) difference in speed by gender that is consistent over tenure.

[39]The distance differences seem to be related to men having a stronger preference for airport trips, possibly due to the fact that they work longer shifts and are, therefore, more willing to stray from their base location.

Table 8 captures the important effects of learning. While men and women learn at the same *per-ride* rate, the driving schedules of men mean that they learn, on average, more intensively *per week* of experience, which generates a gender pay gap.

Finally, to ensure our results are not unique to Chicago, we repeat our analyses for drivers between January 2015 and March 2017 in San Francisco, Boston, Detroit, Atlanta, and Houston. Table 9 presents results for each of our main specifications in the different cities.[40] The results tell a similar story: there is a small baseline gender pay gap in each city, which can be explained by differences in where/when drivers work, different levels of experience, and preferences for driving speed. In the case of Houston, the pay gap actually reverses once controlling for those three factors and men make an estimated 1.2% *less* per hour than women.[41] In San Francisco, the baseline gap is nearly double any other city, at 9.8%, and, even after controlling for all factors, there is a 1.65% residual wage gap.[42]

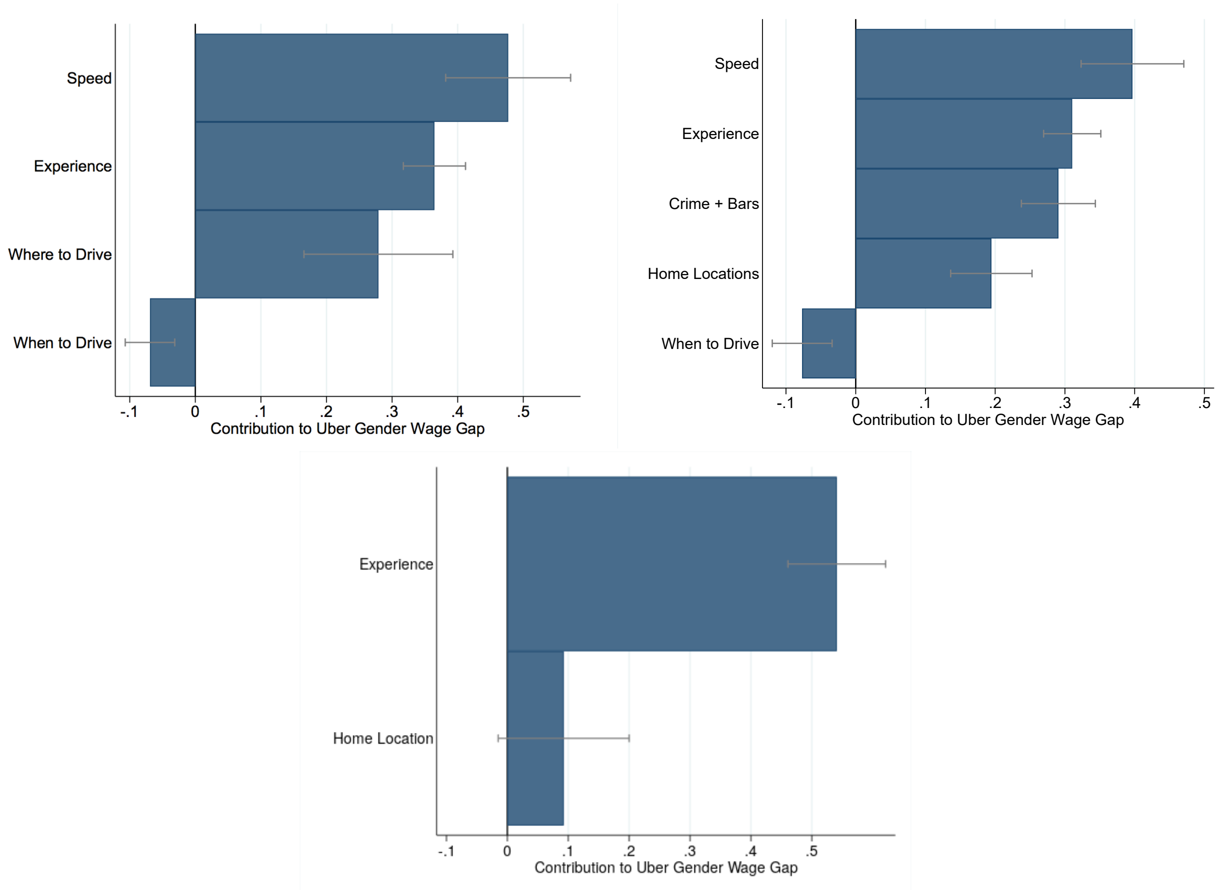
We also conduct a Gelbach decomposition in each city. The results are in Table 9. The ordering of factors by importance is consistent across cities: speed is the most important factor, followed by experience, where, and finally when. The magnitudes, however, differ across cities. In San Francisco, for example, neither when nor where drivers work contribute to the gender wage gap, while in Detroit both when *and* where can explain approximately 10% of the gap.

[40] Cities were generally chosen for a particular feature of the given city. Houston was one of the few cities that Lyft did not operate in during this time frame (it now does). San Francisco had the largest raw pay gap of all major cities. Atlanta had the highest fraction of female drivers. Detroit experimented with different per-time, per-minute rates that reduced the returns to speed. Boston, like Chicago, is one of the cities Uber frequently uses as a testing grounds for new features and research (see, for example, Angrist et al. (2017)).

[41] The Houston gender gap goes to zero if we add in a control for the distance driven on trip. Houston has many highways and men appear to take the highway route more often. This is not relevant in other cities where highways are less important for short distance travel.

[42] The San Francisco gender gap goes to zero if we remove incentive pay from the hourly wage regressions. San Francisco had very large weekly incentive payments for drivers, where they were offered bonus payments if they completed slightly more trips than they had completed per week in the prior month. Analyzing these incentive payments at the weekly level also shows no gender gap once we control for average number of weekly rides given the past month and number of rides given in the current week.

Figure 8: Gelbach decomposition



Note: These figures use the method described in Gelbach (2016) to plot the share of the gender pay gap that can be explained by each factor we consider: speed, experience (lifetime trips controls), where to drive (either geohashes or features of geohashes), and when to drive (hour of week controls).

Table 8: Parameter averages by experience

	Men			Women		
Lifetime trips	0-100	700-800	1400-1500	0-100	700-800	1400-1500
w – Wait time (min)	6.841	5.902	5.740	6.699	5.794	5.630
	(0.00609)	(0.00890)	(0.01091)	(0.01102)	(0.01706)	(0.02434)
m_0 – Accepts-to-pickup distance (mi)	0.716	0.607	0.569	0.647	0.598	0.555
	(0.00433)	(0.00133)	(0.00417)	(0.00078)	(0.00307)	(0.00189)
m_1 – Trip distance (mi)	5.462	5.151	5.062	5.272	5.023	4.982
	(0.00326)	(0.00552)	(0.00693)	(0.00497)	(0.01206)	(0.01723)
s – Speed (mph)	18.915	18.282	18.118	17.907	17.511	17.350
	(0.00472)	(0.00858)	(0.01101)	(0.00706)	(0.01820)	(0.02670)
SM – Surge multiplier	1.090	1.113	1.121	1.090	1.110	1.119
	(0.00013)	(0.00027)	(0.00037)	(0.00020)	(0.00060)	(0.00092)
I – Incentive payout ($)	0.535	0.546	0.552	0.569	0.573	0.641
	(0.00084)	(0.00145)	(0.00188)	(0.00136)	(0.00329)	(0.00529)

Note: This table documents parameter averages from Equation 2.1 by gender and tenure. Drivers are bucketed based on their lifetime trips before a given day. Averages are per-trip based on trips completed in Chicago. For wait time and accepts-to-pickup distance, averages are based on trips from May 2016 - March 2017 due to limitations in the underlying raw data. All other averages are based on data for the entire sample. Wait time is based on time between either coming online or completing previous trip and picking up passenger for new trip. Trip distance is based on actual route taken; however, accepts-to-pickup distance is the Haversine distance between corresponding coordinates. The gender and tenure composition of drivers changes over time; to correct for this, we re-weight by (total trips)/(trips by that gender and tenure group). Standard errors are reported in parentheses.

Table 9: Results from all cities

	Chicago	San Francisco	Boston	Houston	Detroit	Atlanta
Number of drivers	120,223	110,189	72,130	42,194	24,130	64,200
Percent female	30.2%	25.6%	19.8%	26.7%	26.6%	41.9%
Baseline wage gap	0.0356	0.0980	0.0520	0.0327	0.0361	0.0313
	(0.003)	(0.006)	(0.005)	(0.004)	(0.004)	(0.003)
Controls for when, where	0.0220	0.0619	0.0345	0.0156	0.0173	0.0153
	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)
Controls for experience, when, where	0.0085	0.0255	0.0134	0.0022	0.0112	0.0045
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Controls for speed, when, where, experience	-0.0018	0.0165	0.0052	-0.0145	0.0024	-0.0022
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Gelbach – when	-0.0691	-0.0032	-0.0262	-0.0886	0.1108	0.0172
	(0.0193)	(0.0055)	(0.0240)	(0.0296)	(0.0183)	(0.0250)
Gelbach – where	0.2791	0.0183	0.1455	0.1741	0.1270	0.2064
	(0.0579)	(0.0262)	(0.0443)	(0.0169)	(0.0357)	(0.0201)
Gelbach – experience	0.3645	0.4151	0.3829	0.4088	0.2529	0.3681
	(0.0241)	(0.0226)	(0.0536)	(0.0535)	(0.0601)	(0.0381)
Gelbach – speed	0.4770	0.4330	0.4202	0.9923	0.4719	0.5609
	(0.0485)	(0.0395)	(0.0287)	(0.1024)	(0.0536)	(0.0570)

Note: This table includes results from 5 other US cities. All numbers – except for the summary statistics and Gelbach decompositions – are the coefficients on isMale from our standard regressions, with controls: 'when' refers to hour of week, 'where' refers to geohashes, 'experience' refers to bins of lifetime trips, and 'speed' is the log average speed on trip for the given hour. All specifications also control for the calendar week. The outcome variable is log of hourly earnings. Observations are at the driver-hour level. For larger cities, regressions are run on subsets no smaller than 35% so that the full specification is more computationally tractable. Standard errors (in parentheses) are clustered at the driver level.

5 Conclusion

The gig economy has become an important source of earnings for millions of individuals. On Uber alone, there are over 3 million active drivers worldwide completing 15 million trips each day (Bhuiyan (2018)). The growth of flexible work is likely to be especially desirable to women, and could improve the utility (both from monetary compensation and improved time-use) women gain from working. However, our study suggests that jobs offering complete flexibility will likely still contain a gender wage gap, much like the traditional workforce.

Unlike earlier studies, we are able to completely explain the pay gap with three main factors related to driver preferences/constraints and learning: returns to experience, a pay premium for faster driving, and differences in driving locations. Indeed, the contribution of the returns to experience to gender earnings gaps has received little attention in the empirical literature, as it is often quite difficult to measure in traditional work settings. We find that even tracking the number of weeks worked—a common proxy for experience in the literature—does not accurately quantify experience, as men work more hours per week than women and thus accumulate experience more quickly. Measuring the returns to experience is especially important when quantifying the job-flexibility penalty, as omitting experience from the regression overstates the job flexibility penalty.

Importantly, we do *not* find women to be disadvantaged by three factors that one might expect to contribute to a gender wage gap: returns to work intensity, preferences or constraints affecting specific hours worked, or customer discrimination. Overall, our results suggest that on-the-job learning may contribute to the gender earnings gap more broadly in the economy than previously thought. In this spirit, policies that target changes in the time-use choices of men and women may narrow the gender pay gap by helping women move up the learning curve at the same pace as men.

We have shown that, even in the gender-blind, transactional, flexible environment of the gig economy, gender-based preferences/constraints (especially the value of time not spent at paid work and, for drivers, preferences for driving speed) can open gender earnings gaps. Preference/constraint differences that contribute to pay differences in professional markets for lawyers and MBA's also lead to earnings gaps for Uber drivers, suggesting that they are pervasive across the skill distribution, whether in the traditional or gig workplace.

References

Abraham, Katharine G., John C. Haltiwanger, Kristin Sandusky, and James R. Spletzer, "Measuring the Gig Economy: Current Knowledge and Open Issues," NBER Working Papers 24950, National Bureau of Economic Research, Inc August 2018.

Adams, Abi, "The Gender Wage Gap on an Online Labour Market: The Cost of Interruptions," 2020. CEPR Discussion Paper No. DP14294.

Allcott, Hunt, Rebecca Diamond, Jean-Pierre Dubé, Jessie Handbury, Ilya Rahkovsky, and Molly Schnell, "Food Deserts and the Causes of Nutritional Inequality," *Quarterly Journal of Economics*, 2019, *134* (4), 1793–1844.

Altonji, Joseph G. and Rebecca M. Blank, "Race and Gender in the Labor Market," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics, Volume 3*, New York: North-Holland, 1999.

Angrist, Joshua D., Sydnee Caldwell, and Jonathan V. Hall, "Uber vs. Taxi: A Driver's Eye View," 2017. NBER Working Paper #23,891.

Arkhangelsky, Dmitry and Guido Imbens, "The role of the propensity score in fixed effect models," 2018. NBER Working Paper 24,814.

Azmat, Ghazala and Rosa Ferrer, "Gender Gaps in Performance: Evidence from Young Lawyers," *Journal of Political Economy*, October 2017, *125* (5), 1306–1355.

Babcock, Linda and Sara Laschever, *Women don't ask: Negotiation and the gender divide*, Princeton University Press, 2003.

Barth, Erling, Sari Pekkala Kerr, and Claudia Olivetti, "The Dynamics of Gender Earnings Differentials: Evidence from Establishment Data," 2017. NBER Working Paper 23381.

Bayard, Kimberly, Judith Hellerstein, David Neumark, and Kenneth Troske, "New Evidence on Sex Segregation and Sex Differences in Wages from Matched Employee-Employer Data," *Journal of Labor Economics*, 2003, *21*, 887–922.

Bertrand, Marianne, "New Perspectives on Gender," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics, Volume 4b*, North-Holland, Great Britain, 2011.

_, Claudia Goldin, and Lawrence F. Katz, "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors," *American Economic Journal: Applied Economics*, July 2010, *2* (3), 228–255.

Bhuiyan, Johana, "Uber powered four billion rides in 2017. It wants to do more — and cheaper — in 2018," 2018. available at www.recode.net.

Black, Sandra E. and Philip E. Strahan, "The Division of Spoils: Rent-Sharing and Discrimination in a Regulated Industry," *American Economic Review*, 2001, *91*, 814–831.

Blau, Francine D and Lawrence M Kahn, "The feasibility and importance of adding measures of actual experience to cross-sectional data collection," *Journal of Labor Economics*, 2013, *31* (S1), S17–S58.

Blau, Francine D. and Lawrence M. Kahn, "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature*, 2017, *55*, 789–865.

Bolotnyy, Valentin and Natalia Emanuel, "Why Do Women Earn Less Than Men? Evidence from Bus and Train Operators," *Working paper*, 2019.

Buckles, Kasey S. and Daniel M. Hungerman, "Season of Birth and Later Outcomes: Old Questions, New Answers," *The Review of Economics and Statistics*, 2013, *95* (3), 711–724.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women," *Quarterly Journal of Economics*, 2015, *131*, 633–686.

Chandar, Bharat K, Ali Hortacsu, John A List, Ian Muir, and Jeffrey M Wooldridge, "Design and Analysis of Cluster-Randomized Field Experiments in Panel Data Settings," NBER Working Paper 26389 October 2019.

Chandar, Bharat, Uri Gneezy, John A List, and Ian Muir, "The Drivers of Social Preferences: Evidence from a Nationwide Tipping Field Experiment," NBER Working Paper 26380 October 2019.

Chen, M. Keith, Judith A. Chevalier, Peter E. Rossi, and Emily Oehlsen, "The Value of Flexible Work: Evidence from Uber Drivers," *Journal of Political Economy*, 2019, *127* (6), 2735–2794.

Cubas, German, Chinhui Juhn, and Pedro Silos, "Coordinated Work Schedules and the

Gender Wage Gap," 2019. NBER Working Paper 26548.

Cullen, Zoe B., John Eric Humphries, and Bobak Pakzad-Hurson, "Gender and Sorting in the On-Demand Economy," 2018. Working Paper.

Diamond, Rebecca, Tim McQuade, and Franklin Qian, "The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco," *American Economic Review*, 2019, *109* (9), 3365–3394.

Dohmen, Thomas J., Armin Falk, David Huffman, Juergen Schupp, Uwe Sunde, and Gert G. Wagner, "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences," *Journal of the European Economic Association*, 2011, *9*, 522–550.

Farrell, Diana, Fiona Greig, and Amar Hamoudi, "The Online Platform Economy in 2018," 2018. JPMorgan Chase and Company Institute.

_ , _ , and _ , "The Online Platform Economy in 2018," Technical Report, JPMorgan Chase Institute 2018.

Filippas, Apostolos, John Joseph Horton, and Joseph Golden, "Reputation Inflation," in "Proceedings of the 2018 ACM Conference on Economics and Computation" EC '18 ACM New York, NY, USA 2018, pp. 483–484.

Flory, Jeffrey A., Andreas Leibbrandt, and John A. List, "Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job Entry Decisions," *The Review of Economic Studies*, 2015, *82* (1), 122–155.

Gallen, Yana, "Motherhood and the Gender Productivity Gap," 2018. Northwestern University.

Gelbach, Jonah B., "When Do Covariates Matter? And Which Ones, and How Much?," *Journal of Labor Economics*, 2016, *34* (2), 509–543.

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

_ and Lawrence F. Katz, "A Most Egalitarian Profession: Pharmacy and the Evolution of a Family-Friendly Occupation," *Journal of Labor Economics*, 2016, *34*, 705–746.

Grove, Wayne A, Andrew Hussey, and Michael Jetter, "The gender pay gap beyond human capital heterogeneity in noncognitive skills and in labor market tastes," *Journal of Human*

Resources, 2011, *46* (4), 827–874.

Gupta, Nabanita Datta and Donna S. Rothstein, "The Impact of Worker and Establishment-level Characteristics on Male-Female Wage Differentials: Evidence from Danish Matched Employee-Employer Data," *Labour*, 2005, *19*, 1–34.

Haggag, Kareem, Brian McManus, and Giovanni Paci, "Learning by Driving: Productivity Improvements by New York City Taxi Drivers," *American Economic Journal: Applied Economics*, January 2017, *9* (1), 70–95.

Hall, Jonathan V. and Alan B. Krueger, "An Analysis of the Labor Market for UberâĂŹs Driver-Partners in the United States," *ILR Review*, 2018, *71* (3), 705–732.

Hirsch, Boris, Thorsten Schank, and Claus Schnabel, "Differences in Labor Supply to Monopsonistic Firms and the Gender Pay Gap: An Empirical Analysis Using Linked Employer-Employee Data from Germany," *Journal of Labor Economics*, 2010, *28*, 291–330.

Hyperwallet, "The Future of Gig Work is Female," 2017. available at www.hyperwallet.com.

Katz, Lawrence F. and Alan B. Krueger, "Understanding Trends in Alternative Work Arrangements in the United States," *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 2019, *5* (5), 132–146.

Light, Audrey and Manuelita Ureta, "Early-career work experience and gender wage differentials," *Journal of Labor Economics*, 1995, *13* (1), 121–154.

Madden, Janice Fanning, "Why Women Work Closer to Home," *Urban Studies*, 1981, *18* (2), 181–194.

Mas, Alexandre and Amanda Pallais, "Valuing Alternative Work Arrangements," *American Economic Review*, 2017, *107* (12), 3722–3759.

Massie, Dawn L., Kenneth L. Campbell, and Allan F. Williams, "Traffic Accident Involvement Rates by Driver Age and Gender," *Accident Analysis and Prevention*, 1995, *27*, 73–87.

Mincer, Jacob and Solomon Polachek, "Family investments in human capital: Earnings of women," *Journal of political Economy*, 1974, *82* (2, Part 2), S76–S108.

Niederle, Muriel and Lise Vesterlund, "Do women shy away from competition? Do men compete too much?," *The Quarterly Journal of Economics*, 2007, *122* (3), 1067–1101.

Santamarina-Rubioa, Elena, Katherine Péreza, Marta Olabarriaa, and Ana M. Novoaa, "Gender differences in road traffic injury rate using time travelled as a measure of exposure," *Accident Analysis and Prevention*, 2014, *65*, 1–7.

Statista, "How interested, or not, are you in the NFL," 2017. available at www.statista.com/statistics/778854/national-football-league-us-consumer-interest-by-gender/.

A Appendix

A.1 Costs of driving

In our main analyses, we use total earnings for drivers as our primary measure of earnings. But, if costs differ in a way that is correlated with gender, we could have understated or overstated the gender "net" pay gap. In some ways, that is not problematic and is consistent with other work. No studies of the gender gap account for differences in costs of working, though the costs of work vary for reasons that may correlate with gender (such as commuting, clothing, and styling). However, given that a large capital cost is a requirement for independent drivers and that they may deduct some costs from their taxes, it may be appropriate in their case to consider earnings net of direct expenses of driving for hire.

The primary costs drivers face are fuel, maintenance, depreciation, and fines for parking or moving violations. Suppose drivers average 25 cents per mile in costs other than insurance – Uber covers drivers' insurance costs while driving. A typical Uber driver covers about 20 miles in one hour. The driver earns approximately $15.80 net of Uber's current 25% average share of driver gross earnings. As a "raw" gender gap, we use the 3.56% from column 2 of Table 4 which controls only for calendar week. Based on these numbers, men net 56 cents more per hour than women before expenses. Total average costs per hour—based on 25 cents per mile—would be $5. Men's costs would have to be $\sim$11% higher, in terms of fuel, maintenance, depreciation, and fines to erase the gender gap.

Many costs, such as fines and maintenance, are not observable. For gasoline, we can estimate costs by looking at the types of cars men and women drive (although style of driving, including speeding, will also be a factor in determining fuel efficiency). Men and women may drive cars with different average fuel efficiency. At a high-level, this appears to be true – of all miles driven by men in the data, 6.4% of them were in a Toyota Prius compared to only 3.6% of miles driven by women. Men have, on average, more incentive to invest in more fuel efficient vehicles due to their longer driving hours. To further test this, we match drivers' vehicles to fuel economy data from the EPA.[43] On average—weighting by miles driven on Uber—women drive cars that get 25.23 miles per gallon in the city and men drive cars that get 26.80 miles per gallon. Men are getting about 6.4% more miles per gallon on average; controlling for gasoline costs would likely increase the gender pay gap.

Another cost to consider is insurance. Though insurance is a large cost for drivers, Uber pays drivers' insurance when they are working. The costs of insurance are relevant, however, as a proxy for accidents (and the downtime that goes with them) and tickets. Men pay more than women for car insurance, though the rates converge at age 26. Accident rates per mile driven are higher for young men than young women but the difference narrows or disappears around age 25. Fatality rates remain higher for men.[44] However, given the insurance rates converge, it seems that the total costs of dangerous driving are about the same by gender after age 25. This suggests that accidents and fines should only vary by gender for drivers under 26. In our sample of Chicago drivers, 15.8% of female drivers and 14.8% of male drivers are under 26. So our gender gap estimates should not be affected by these costs for the vast majority of our sample.

Overall, we cannot estimate differences in costs by gender nearly as precisely as we can estimate gender differences in earnings. However, we also do not see any evidence that the gender pay differences are offset by cost differences.

[43] Fuel economy data are available at the level of the vehicle make, model, year, and trim. Drivers manually enter these fields on sign-up; there are often typos or abbreviations (e.g., "s-class" instead of the exact model). We fuzzy-match based on the Levenshtein distance between the Uber model and the EPA data's model. Results are based only on matches with a Levenstein distance over 0.7 (about 70% of the data). Results are qualitatively similar for different Levenshtein distance thresholds.

[44] See Massie et al. (1995) and Santamarina-Rubioa et al. (2014).

A.2 Preferences and constraints

Our data on Uber drivers do not include information on a driver's marital status, education, income, or whether they have children. This complicates any effort to isolate the exact nature of how differences in preferences and constraints drive gender differences. If women on average shoulder more of the burden of child rearing, for example, they may be less able to work on days when children (especially young children) are not in school.

To examine how constraints and preferences may underlie certain driving decisions, we look at whether certain events happening on those specific days (or hours) have an affect on the gender composition. We look at two main events. First, school days off, including non-holiday days off, Spring break, holidays, parent-teacher conferences, and Summer break. Second, NFL football games – in a 2017 survey by Statista, 22% of adult men responded that football was one of their top interests compared to only 11% of women (Statista, 2017).

A.2.1 School breaks

We build a dataset of school breaks and parent-teacher conferences for Chicago public schools, which enroll approximately 400,000 children. We focus on three types of days off that affect schools primarily: 1) 'school improvement' days, 2) Spring break each April, and 3) parent-teacher conference days. School improvement days are planned non-attendance days (generally Fridays) for teachers and staff to meet and discuss students and plans. The rest of the city is not affected. Class is in session during parent-teacher conference days, but parents are required to meet with teachers. There are four each of student improvement and parent-teacher conference days each year. Finally, we identify other holidays—such as MLK day and winter break—that are not unique to schools.

We regress the log percent of drivers on a given day who are male on each type of break, with controls for the year-month and day of week. Results are presented in Columns 1 and 2 of Table 10. The gender composition of Uber drivers on a given day does not appear to be affected by school breaks. The lone significant coefficient is for holidays; men drive relatively more during the holidays. School improvement days, parent-teacher conferences, and Spring break have no detectable effect on whether men or women driver disproportionately more. Results are qualitatively similar if we instead use the percent of minutes worked, rather than the percent of active drivers, as an outcome variable.

Next, we look at Summer breaks. We consider them separately due to their length. Using the CPS calendar, we identify the first and last day of school for each year in the data and look at the 21 days before and after school begins, with a dummy variable for "school started" in the 21 days after school starts and similarly for "school ended." The outcome variable is again the log of the percent of active drivers who are male. Results are in Columns 3 and 4 of Table 10. Neither coefficient is statistically significant, but the point estimates would indicate that women drive slightly less than men during the school year.

These results suggest that constraints on drivers stemming from their children's school attendance do not meaningfully impact men and women differently. We caution against generalizing this much further; other unobserved constraints, such as pregnancy, could have a large effect on women's ability to drive during certain times.

A.2.2 Football games

We next look at driving behavior during football games by the Chicago Bears. Football games generally take place on Sundays at Noon, 3-3:30pm, or 7:30pm and last 3-4 hours. There are 32 games that took place during our sample.

We regress the log percent of drivers in a given hour who are male on whether there is a game occurring during that hour, with controls for the week and day of week. We assume games start in the hour of the scheduled start and last for 4 hours. Results are presented in Column 5 of Table 10. On average, relatively fewer men drive during football games: the percent of men decrease by 4.4% (about 3.5 percentage points).

Figure 10 graphs the number of male and female drivers active in the hours before/after a football game begins (normalized to sum to one). Men drive relatively more than women in the morning, but as the game approaches the number of active men decreases. Men appear to begin driving again about two hours after the game has begun, before it has finished – given that the

Bears had a division-worst record of 9-23 these years, fans may have preferred to not watch the end of the game.

Table 10: Driving during school breaks and football games

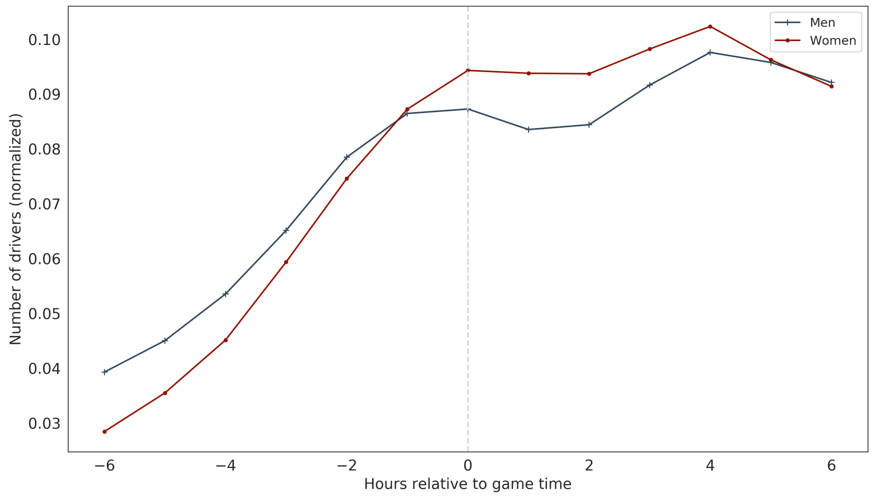
	Daily		School start/end ± 21 days		Hourly
	(1)	(2)	(3)	(4)	(5)
All days off (during school year)	0.0027				
	(0.001)				
'School improvement' day		-0.0010			
		(0.003)			
Spring break		0.0023			
		(0.005)			
Holiday		0.0047			
		(0.004)			
Parent-teacher conference		0.0027			
		(0.003)			
School started			0.0042		
			(0.005)		
School ended				-0.0106	
				(0.006)	
Football game					-0.0440
					(0.003)
Intercept	-0.1094	-0.1095			-0.1048
	(0.001)	(0.001)			(0.000)
Day of week	X	X	X	X	X
Year-week					X
Year-month	X	X			
N	792	792	86	86	18,952
R^2	0.921	0.921	0.216	0.128	0.523

Note: This table documents the relationship between the gender composition of Uber drivers in a given day/hour and school breaks or football games. The outcome variable is the log of the percent of active drivers in that day/hour who are male. We consider an hour to be during a football game if it is between the start of a Chicago Bears game and four hours later. For school start/end, we use a sample of 21 days before/after school starts or ends. Standard errors are in parentheses.

A.3 Taxi drivers

On the surface, taxi driving shares many similarities to Uber driving, although the differences are important. First, taxis are generally required to receive a special license. Second, many cities use medallions to limit the number of taxis. The cost of a medallion makes part-time driving infeasible. Third, most taxi rides are found by observing a 'flag' on the street. In less affluent or less busy areas, availability of taxis is often limited.

Figure 9: Driving during football games



Note: The figure documents driving behavior before and after Chicago Bears football games during the 2015 and 2016 seasons. The lines are normalized to sum to one so that the levels are more easily compared.

We document gender differences in taxi drivers using data from the Annual Social and Economic Supplement (ASEC) from the Current Population Survey (CPS) for 2004-2018 (which corresponds to outcomes in 2003-2017). We limit the data to occupation "taxi drivers and chauffeurs" and split the data into pre- and post-Uber launching in 2012.[45] There are 10 years in the pre-period and 5 in the post-period.

Taxi drivers are predominantly male: 83.7% of taxi drivers in the pre-period and 84.7% of taxi drivers in the post-period are male. The difference is not statistically significant.

We next regress imputed hourly earnings – based on weeks worked in the previous year, usual hours per week, and reported wages – against worker gender, controlling for state. Results are reported in Table 11. We find a negative but statistically insignificant pay gap between men and women. Including a control for pre-/post-Uber does not affect the pay gap. Perhaps surprisingly, the data suggest that taxi pay *increased* in the period after Uber launched. This may be due to Uber drivers being labeled as taxi drivers in the post-period; the sample includes approximately 180 observations per year in the pre-period and 350 per year in the post-period, suggesting some change in which workers were being captured under this label. The difference in earnings pre-/post-may also reflect selection or a shift in the composition of "taxi drivers and chauffeurs" towards the

[45]Uber launched as a black-car-only service in 2009, but UberX was not launched until July 2012.

more generously compensated chauffeurs.

Table 11: Taxi drivers and chauffeurs

	(1)	(2)	(3)
isMale	-0.0513	-0.0573	-0.0511
	(0.052)	(0.052)	(0.072)
postUberLaunch		0.2207	0.2317
		(0.037)	(0.095)
isMale:postUberLaunch			-0.0129
			(0.103)
Intercept	2.4139	2.3298	2.3258
	(0.009)	(0.244)	
State	X	X	X
N	3,566	3,566	3,566
R^2	0.029	0.039	0.039

Note: This table documents the differences in pay for taxi drivers by gender. Data are from the CPS ASEC for 2004-2018 (covering economic activity in years 2003-2018). For the purposes of splitting the data, we consider Uber launch to occur in 2012, the year that UberX launched. Standard errors are in parentheses.

A.4 Driver ratings

At the end of each trip, riders have the option to rate drivers. Ratings are not required; both female and male drivers are rated for 68.4% of trips. The average rating given by riders to both male and female drivers in Chicago is 4.75. The high average rating is consistent with past research showing ratings inflation in many online markets (Filippas et al., 2018). Unfortunately, the inflation in ratings reduces the value of ratings as a measure of driver quality.

Drivers also rate riders at the end of each trip. Here we do find a small gender difference: female *drivers* give higher ratings to their riders than male drivers, on average; riders receive an average rating of 4.81 from female drivers and 4.78 from male drivers.

A.5 Incentive Earnings

Uber offers a number of incentives to drivers in addition to 'organic' earnings (base fare plus per-minute and per-mile rates times any surge multiplier). Both the ubiquity and the structure of incentives have varied over time. For most of our sample, Uber offered drivers two main types of incentives. First, 'Boost' incentives, which guaranteed a surge rate floor for a time and location (e.g., earn at least 1.5x on trips in the Loop, Monday 8-10am). Drivers would receive the higher of the actual surge rate and the incentive surge rate. Boost incentives help Uber spatially and temporally position supply. Second, 'Quest' incentives which would pay a bonus for drivers completing a certain number of trips in a given time frame (e.g., do 10 trips between Monday and Thursday, get $20 extra). Quest incentives encourage loyalty to the Uber platform. The number of trips required is tailored based on a driver's historical number of completed trips and is designed to be a stretch, but attainable. In general, drivers who were active in the past 28 days received both Boost and Quest incentives for a week, announced ahead of time via email.

Our results replicate qualitatively to hourly earnings excluding Quest incentives and hourly earnings excluding *all* incentive pay.[46] In Table 12, we replicate our baseline and fully saturated specifications on two new definitions of hourly earnings, one that excludes Quest incentives and another that excludes all incentives. Without Quest incentives, the gender pay gap is 4.3%, notably larger than the baseline gender pay gap when incentives are included ($\sim$ 3.6%). Women do substantially better than men on these weekly bonuses on a per-hour basis; examining closer, we find that women earn more per-trip conditional on receiving any Quest bonuses that week (intensive margin) with no gender gap in the rate men and women receive any Quest incentive pay in a week (extensive margin). In contrast, men tend to earn more per-trip in Boost incentives, which reward driving during the busiest hours. When we further exclude Boost incentives, the baseline pay gap shrinks 4.0%. In both cases, our standard set of controls—time, location, experience, and driving speed—can entirely explain the pay gap.

A.6 Features of driving location

A.6.1 Data construction

Crime data are provided by the Chicago Data Portal and encompass all reported crimes during our time sample, January 2015 to March 2017. To proxy for the relative safety of a geohash,

[46] Due to data limitations, we cannot perfectly tie incentive pay to the type of incentive (Quest, Boost, and other less commonly used ones). We can identify incentive payments tied to an individual trip, which are likely Boost. The remaining incentives—which we observe being paid out on the weekly basis as bonuses—are almost entirely Quest. Further, we do not observe whether drivers were offered Quest in a given week nor the number of trips they needed to complete to qualify.

Table 12: Gender pay gap, earnings excluding incentives

	Without Quest incentives		Without any incentives	
	(1)	(2)	(3)	(4)
isMale	0.0431	0.0010	0.0401	0.0006
	(0.003)	(0.002)	(0.003)	(0.002)
Intercept	2.9917	2.9739	2.9739	1.4803
	(0.003)	(0.003)	(0.005)	(0.005)
Week	X	X	X	X
Hour of week		X		X
Geohash		X		X
Experience bins		X		X
Log speed		X		X
N	11,572,163	11,572,163	11,572,163	11,572,163
R^2	0.028	0.025	0.251	0.241

Note: This table documents the gender pay gap for hourly earnings without incentives. The outcome variable is log of hourly earnings. Quest incentives are weekly bonuses for completing a certain number of trips that week. Hour of week controls for each of 168 hours. Geohash controls are a vector of dummies for whether a driver began a trip in a given geohash. Experience bins are as defined in the main section of the paper. Standard errors (clustered at the driver-level) are in parentheses.

we measure the number of crimes that occur during our time sample in that geohash per 1,000 residents. We restrict crime data to non-residential crimes; residential crimes, such as domestic violence, are unlikely to be relevant for a driver's sense of safety. The most frequent crimes in the data are theft, battery, and criminal damage.

Population and demographics data come from the 2016 American Community Survey (ACS). The ACS data are at the census tract level; to map to geohashes, we look at all census tracts that intersect a given geohash and assume that the population is uniformly distributed across the census tract such that, if half of a census tract is in a given geohash, we assign half of its population to that geohash. Demographics such as median household income are based on the population-weighted average of all census tracts that intersect a given geohash.

Driver home addresses are based on the address associated with a driver's bank account, which we believe to be more likely up-to-date than the address on their license. We map these addresses to geohashes. For both census population numbers and driver home locations, we smooth the features by measuring the number of drivers (or general population) living in geohashes within one mile of the given geohash.

Finally, we use data on business licenses from the Chicago Data Portal to find number of unique locales with an active liquor license. Bar districts may be differentially unsafe for women, especially during late night hours. In total, there are 3,877 businesses with a liquor license.

Because some of the neighborhood variables we analyze change substantially over small distances and because we are doing the analysis at the geohash level (rather than the driver-hour level), we use finer geohashes for this analysis than we use in our driver pay regressions. Specifically, in this section, we use precision six geohashes; each geohash side is approximately 0.75 miles long. In addition to limiting to the City of Chicago geography, we further limit to geohashes with over 1,000 trips. There are 1,014 geohashes in the resulting data (compared to the fifty larger geohashes we use in the pay regressions), covering 68.9% of our total trips.

A.6.2 Determinants of driving location

To more formally estimate how features of a location relate to male and female drivers propensity to drive there, we regress the difference in log share of trips for men and women against these features. That is, for quantity of trips Q completed by women (w) and men (m) in geohash $g \in G$, the set of all geohashes, we regress

$$\begin{aligned} \log\left(\frac{Q_{g,m}}{\sum\limits_{g\in G} Q_g}\right) - \log\left(\frac{Q_{g,w}}{\sum\limits_{g\in G} Q_g}\right) &= \beta_m X_g - \beta_w X_g + (\epsilon_{g,m} - \epsilon_{g,w}) \\ \log\left(\frac{Q_{g,m}}{Q_{g,w}}\right) &= (\beta_m - \beta_w) X_g + (\epsilon_{g,m} - \epsilon_{g,w}) \end{aligned} \tag{A.1}$$

Results are presented in Table 13.

A.7 Driver Race

Uber does not directly observe driver race. We instead impute race based on a driver's name and home census block following the method described in Diamond et al. (2019), which gives us probabilities of a driver being Hispanic, black, Asian, and white.[47] Driver race using this method

[47] Home census block is available for ∼93% of drivers. When unavailable, we simply use name. Our method deviates slightly from Diamond et al. (2019) in that we use the open source Python package 'ethnicolor' rather than

Table 13: Differences in driving location by gender

	(1)	(2)	(3)
Log share of driver homes male (within 1mi)	0.6667		0.7175
	(0.010)		(0.016)
Log # of crimes per 1,000pp		-0.0866	0.0334
		(0.009)	(0.006)
Log # of liquor licenses		0.1366	0.0653
		(0.011)	(0.007)
Log median household income		0.1495	-0.0374
		(0.027)	(0.016)
Log share adult census pop. male (within 1mi)		0.3833	-0.1396
		(0.053)	(0.032)
Intercept	0.9180	0.0587	1.0748
	(0.011)	(0.311)	(0.180)
N	1,014	1,014	1,014
R^2	0.774	0.435	0.815

Note: This table presents results from various regressions of the form specified in Equation A.1, where the outcome is the log share of trips by male drivers originating in a given geohash. The mean and standard deviation of each feature (as well as the outcome) are available in Table 14 in the appendix. Standard errors are in parentheses.

is available for 98.4% of drivers (98.8% of driver-hours). We then create dummy indicators for a driver's most likely race based on the max of the predicted probabilities.[48]. For the small subset of drivers for whom we cannot identify race, we create an indicator for missing race; this way, our sample is identical to the sample used throughout.

The distribution of driver race by gender is documented in Table 15. The difference is stark.

the proprietry software NamePrism to predict race based on a driver's name. This is due to privacy considerations; NamePrism cannot be run locally so would have involved sending data to a third-party.

[48]For 73.3% of drivers, one race is predited with $> 80\%$ probability. Our results are robust to limiting the data to these drivers as well as to simply using the raw race probabilities rather than dummy indicators

Table 14: Summary statistics, geohash-level features

	Mean	Std. dev.
Share of trips male	4.2845	1.6884
Share of driver homes male (within 1mi)	2.2794	1.1783
# of crimes per 1,000pp	110.68	114.79
# of liquor licenses	3.7968	10.335
Median household income	53,545	25,172
Share adult census pop. male (within 1mi)	0.9025	0.2498

Note: This table presents the mean and standard deviation for geohash level features for geohashes in the City of Chicago area.

Men are about 1.5 times as likely to be white and 5 times as likely to be Asian. Women are over twice as likely to be black.

In Table 16, we replicate our baseline and fully saturated regression specifications controlling for driver race. The baseline gender pay gap with race controls is now 3.0% (compare to 3.6% without controls for race). With controls for time, location, driver experience, and speed, the pay gap again becomes statistically indistinguishable from zero; in fact, it becomes slightly negative (p-value of 0.091). We also interact gender and race to test whether within-race gender gaps exist; as shown in column 4, for each race the within-race gender gap is not significantly different than zero.

Finally, we again conduct a Gelbach decomposition of each factor, now controlling for driver race in the baseline. The results are presented in Figure 10. Compared to results without controls for race, the speed, where, and experience factors now contribute more equally to explaining the pay gap. As before, when to drive has a small negative weight. If we replace where to drive with controls for features of where they drive—crime, number of bars, and home location—we again find that speed, experience, and the rate of crime and bar locations are again the predominant factors. Home location now matters less; however, controlling for home location while also controlling for race imputed in part from home location complicates the interpretation.

In general, despite dramatically different racial compositions across gender, we do not find evidence to suggest that the race of drivers has a substantial impact on the gender pay gap and does not bias the gender pay gap analysis.

Table 15: Distribution of driver race, by gender

	Men	Women
Likely Hispanic	18.52%	14.88%
Likely Black	20.02%	45.20%
Likely Asian	9.40%	2.21%
Likely White	52.06%	37.71%

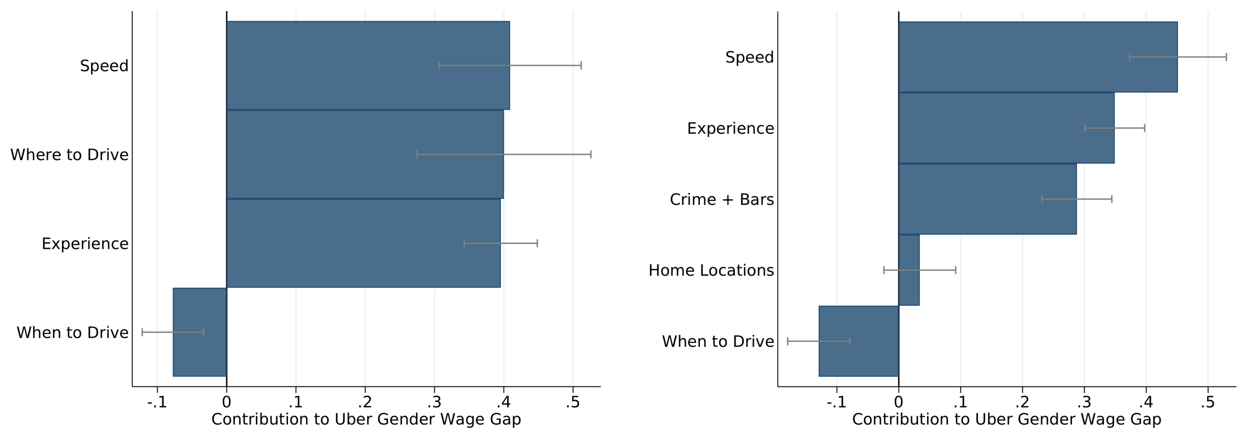
Note: This table presents the distribution of drivers across races. Race is imputed based on a driver's name and home census block; we define driver race based on the most probable race from the imputation.

Table 16: Gender pay gap with controls for driver race

	(1)	(2)	(3)	(4)
isMale	0.0303	-0.0038	0.0326	-0.0036
	(0.003)	(0.002)	(0.005)	(0.004)
Likely Black	-0.0068	-0.0099	-0.0028	-0.0095
	(0.043)	(0.003)	(0.006)	(0.004)
isMale * Likely Black			-0.0052	-0.0005
			(0.008)	(0.005)
Likely Hispanic	0.0180	0.0133	0.0193	0.0157
	(0.008)	(0.003)	(0.002)	(0.005)
isMale * Likely Hispanic			-0.0015	-0.0028
			(0.009)	(0.006)
Likely Asian	0.0355	0.0072	0.0430	0.0060
	(0.016)	(0.004)	(0.002)	(0.012)
isMale * Likely Asian			-0.0081	0.0012
			(0.017)	(0.012)
Missing race	-0.0094	-0.0188	-0.0547	-0.0556
	(0.010)	(0.007)	(0.027)	(0.025)
isMale * Missing Race			0.0512	0.0417
			(0.029)	(0.026)
Intercept	3.0850	1.9418	3.0829	1.9416
	(0.003)	(0.007)	(0.003)	(0.008)
Week	X	X	X	X
Hour of week		X		X
Geohash		X		X
Experience bins		X		X
Log speed		X		X
N	11,572,163	11,572,163	11,572,163	11,572,163
R^2	0.040	0.259	0.040	0.259

Note: This table documents the gender pay gap for hourly earnings after controlling for driver race. The outcome variable is log of hourly earnings. Race is imputed based on a driver's name and home census block; we define driver race based on the most probable race from the imputation. Hour of week controls for each of 168 hours. Geohash controls are a vector of dummies for whether a driver began a trip in a given geohash. Experience bins are as defined in the main section of the paper. Standard errors (clustered at the driver-level) are in parentheses.

Figure 10: Gelbach decomposition, with race controls



Note: These figures use the method described in Gelbach (2016) to plot the share of the gender pay gap that can be explained by each factor we consider: speed, experience (lifetime trips controls), where to drive (either geohashes or features of geohashes), and when to drive (hour of week controls). The baseline includes controls for driver race.

A.8 Differential returns to learning

To test whether men and women learn at different rates, we include an interaction for driver gender in our regressions estimating the returns to experience. As shown in Table 17, there is no evidence of differential learning. We repeat this analysis in the section below where we include driver fixed effects to estimate a learning curve separately for each gender. We again find no statistical difference in learning curves by gender.

A.9 Returns to experience

The returns to experience we document could be driven by selection bias, if drivers' baseline productivity levels are correlated with how much experience they accumulate then cross-section variation in the driver experience would produce a biased estimate of the return to experience. To test this, we will add driver fixed effects to our model. Before adding driver fixed effects, we need to change how we parameterize the return to experience in the regression. In our main analysis we summarize the return to experience using dummy variables representing ranges of experience levels. While this approach provides easy to interpret estimates of the return to experience in the cross-section, it is a poor way to model the learning curve using within-driver variation. By including driver

Table 17: Differential learning

	(1)	(2)
isMale	0.0145	0.0096
	(0.002)	(0.002)
Trips completed: 100-500	0.0529	0.0343
	(0.003)	(0.002)
Trips completed: 500-1000	0.0768	0.0493
	(0.004)	(0.003)
Trips completed: 1000-2500	0.0995	0.0655
	(0.006)	(0.004)
Trips completed: >2500	0.1453	0.0919
	(0.014)	(0.009)
isMale*Trips completed: 100-500	−0.0004	−0.0006
	(0.003)	(0.003)
isMale*Trips completed: 500-1000	0.0004	−0.0002
	(0.004)	(0.004)
isMale*Trips completed: 1000-2500	0.0006	−0.0022
	(0.006)	(0.006)
isMale*Trips completed: >2500	−0.0069	−0.0067
	(0.0012)	(0.012)
Intercept	3.0223	3.0571
	(0.002)	(0.002)
Week	X	X
Geohash*hour of week		X
N	11,572,163	11,572,163
R^2	0.047	0.164

Note: The table expands on results presented in Table 6 by adding interacting gender and experience. Outcome variable is log of hourly earnings. Standard errors (clustered at the driver-level) are in parentheses.

fixed effects, the learning curve will be estimated using changes in experience within drivers. Since few drivers go all the way up the learning curve, a large amount of the data will involve drivers' experience levels changing by only a modest amount. Using dummy variables to represent the experience curve will then be downward biased since they will be estimated from drivers' experience levels jumping from just below the cutoff to just above the cutoff of experience bins. Since this clearly is a tiny accumulation of experience, the dummy variable measure of experience will be biased downward. Instead we use a continue measure of experience presented by a piece-wise linear spline with four kink points. We re-estimate the experience regression shown in Column 4 of Table 6 (including controls for where drivers pick-up, the hour-of-week, week, and log driving speed), but replace the experience bin dummies with our spline. Figure 11 plots the estimated experience curve. As expected, the spline produces very similar experience estimates to those in in Column 4 of Table 6 that use dummies for ranges of experience.

The most common way to estimate the return to experience using only within driver variation in experience is to include driver fixed effects in the regression. Unfortunately, doing this makes the gender dummy drop-out of the regression, as it is co-linear with the driver fixed effects. This is undesirable because it does not allow us to then decompose how much of the return to experience explains the gender gap. However, the main endogeneity problem is that between-driver variation in experience levels may be correlated with the drivers' baseline productivity level. While fixed effects would surely absorb this concern, a simpler alternative is to control for each driver's average level of experience across all his spells of working for Uber. This control will then absorb the between driver variation in average experience, allowing the learning curve to be identified by drivers' experience level at a point in time, given his average level of accumulated experience. Further, adding this control will no longer force the gender dummy to drop out of the regression. Indeed, Arkhangelsky and Imbens (2018) show that controlling for group-specific means of all covariates in a regression is *numerically identical* to including group fixed effects. To estimate a learning curve identical to the one found by including driver fixed effects would require controlling for driver means of all covariates (location dummies, week dummies, hour-of-week dummies, log driving speed) in addition to drivers' mean experience level. However, it is less clear why between-driver variation in these other covariates would lead to a biased learning curve. To assess whether controlling for drivers' mean experience level is sufficient for estimating an un-biased learning curve, we compare the learning curve estimated by including driver fixed effects to the learning curve estimated by controlling for drivers' mean experience level. The fixed effect approach estimates:

$$\begin{aligned}\ln(W_{dt}) &= \sum_{i=1}^{5} \beta_i \left(\left(E_{dt} - \bar{E}_i \right) \mathbb{1}\left(\bar{E}_{i+1} \geq E_{dt} \geq \bar{E}_i \right) + \left(E_{i+1} - \bar{E}_i \right) \mathbb{1}\left(E_{dt} \geq \bar{E}_{i+1} \right) \right) \\ &\quad + \delta_d + \gamma X_{dt} + \varepsilon_{dt}.\end{aligned}$$

E_{dt} measures the number of prior rides completed by driver d at time t. E_i represents the kink points in the piece-wise linear spline, indexed by i. β_i measures the return to experience in region i of the experience curve. δ_d are driver fixed effects and X_{dt} represent the where, when, and driving speed controls. Figure 12 plots the estimated return to experience from the driver fixed effect approach in red. We now compare these learning curve estimates to those where we include each drivers' mean experience level as an additional control, instead of driver fixed effects. We estimate:

$$\begin{aligned}\ln(W_{dt}) &= \sum_{i=1}^{5} \beta_i \left(\left(E_{dt} - \bar{E}_i \right) \mathbb{1}\left(\bar{E}_{i+1} \geq E_{dt} \geq \bar{E}_i \right) + \left(E_{i+1} - \bar{E}_i \right) \mathbb{1}\left(E_{dt} \geq \bar{E}_{i+1} \right) \right) && \text{(A.2)} \\ &\quad + \alpha \bar{E}_d + \gamma X_{dt} + \varepsilon_{dt}, && \text{(A.3)} \\ \bar{E}_d &= \frac{1}{N_d} \sum_{t=1}^{N_d} E_d t, && \text{(A.4)}\end{aligned}$$

where $\bar{E}_d$ represents driver d's average number of completed prior rides, average over all N_d hours that driver d worked. Estimates of the learning curve from this specification are plotted in blue in Figure 12. The learning curve estimated with driver fixed effects is statically indistinguishable from the learning curve estimates that control for drivers' mean experience level. Since controlling for drivers' mean experience levels allows us to keep the gender dummy in the regression and carry this specification through our decompositions methods, we focus on this specification.

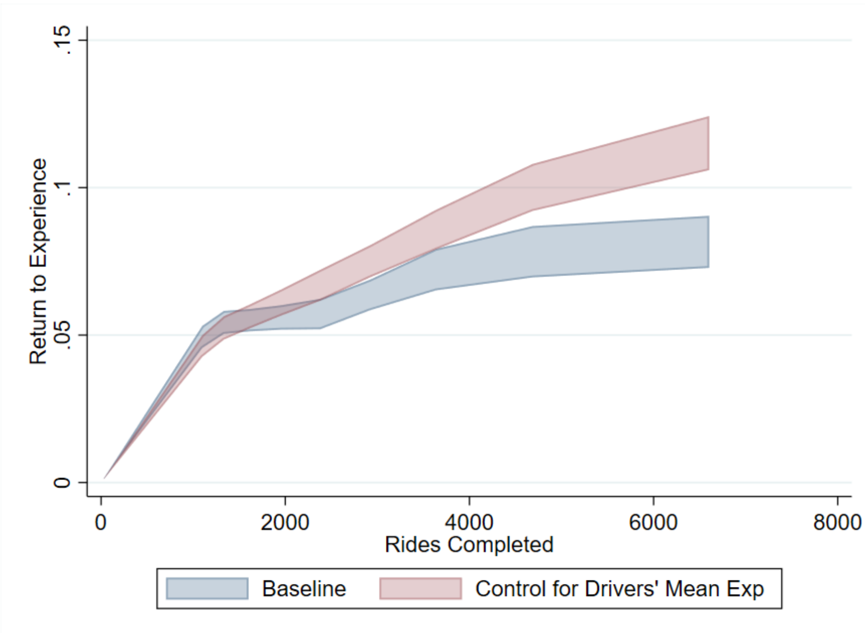
We return now to Figure 11 that compares the learning curve estimated in the simple cross-section, to one that controls for driver average experience. We find that the cross-sectional learning curve was biased *downward*, especially once the number of prior rides goes beyond 2000. This highlights that the drivers who choose to complete a large number of rides tend to be those with *worse* innate productivity levels. This suggests that the drivers with worse outside employment options are those choose to drive for Uber most intensely, enabling them to work a large number of hours and accumulate a high level of experience. The more productive drivers (net of experience) are those who drive for Uber less intensely, possibly supplementing an outside income source.

We repeat the driver fixed effect specification but now include interaction terms between the experience spline and the gender dummy. This identifies the difference in learning curve by gender. We plot the 95% confidence intervals of the difference in the learning curves by gender in Figure 13. We find the point estimates are very close to zero and standard errors can never reject that there is no differential learning by gender.

A.10 Decomposing the gender gap using an unbiased learning curve

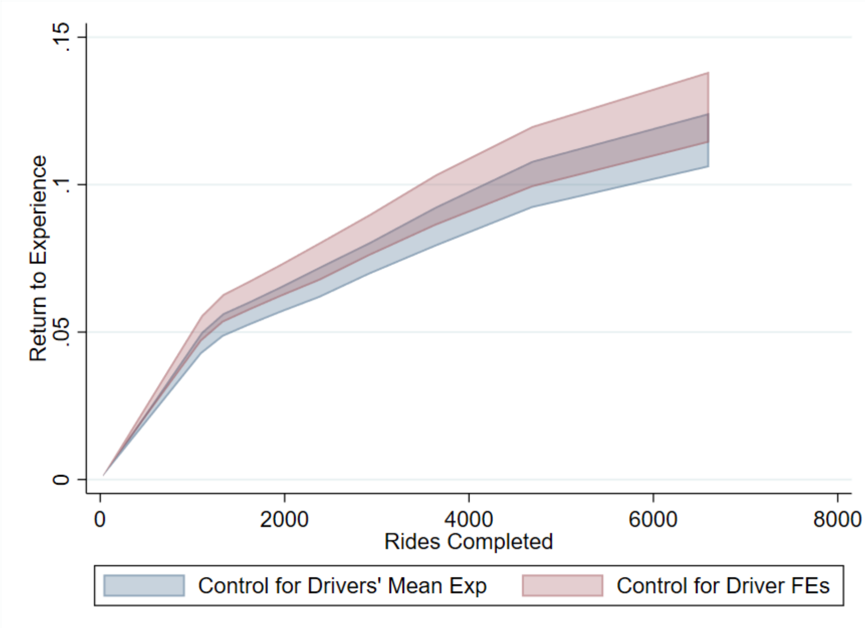
We decompose the contributors to the gender wage gap using the Gelbach decomposition, as done in the main body of the paper. We estimate the learning curve as done above in equation A.2. Figure 14 shows the contributors of the gender gap using the unbiased learning curve estimated above. Consistent with the results in the main paper, which use the simpler learning curve that does not correct for selection, we find that experience and speed each contribute to about 40 percent of the gender gap. Where drivers work contributes about 25 percent and when drivers work slightly mitigates the gap. However, since some of the return to experience is due to learning about where and when to drive, this decomposition under estimates the contribution of experience, since it only attributes experience effects over and beyond speed, where to drive, and when to drive. The right panel of Figure 14 drop these mitigating factors and quantifies how much of the gender gap can be explained solely by experience and where drivers live. Consistent with the results in the main text, experience explains about 65 percent of the gap and differences in drivers' places of residences explains about ten percent.

Figure 11: Return to Experience



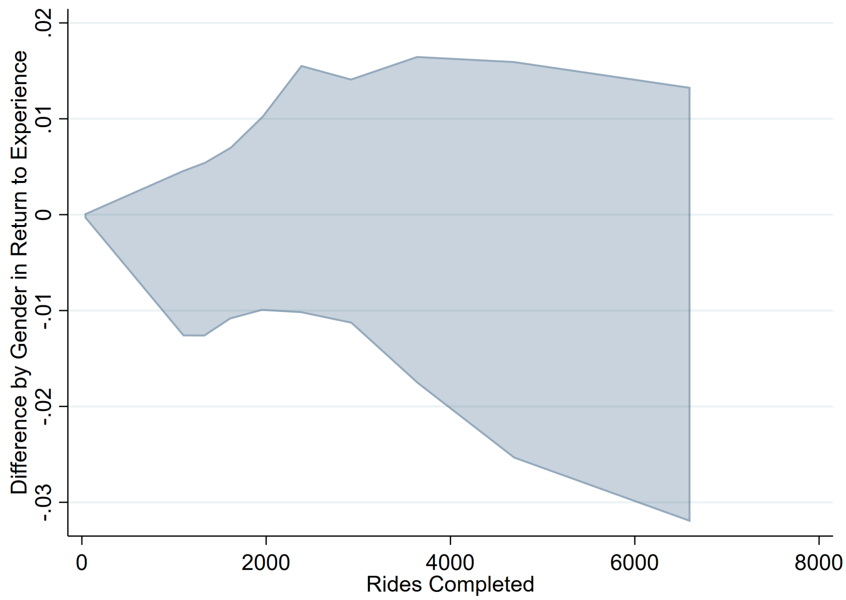
Note: This figure graphs the return to experience using a piece-wise linear spline, controlling for where drivers pick-up, the hour-of-week, week, and log driving speed. Range represents 95% confidence intervals. The "Baseline" estimates use a piece-wise linear spline, instead of dummies for bins of experience levels, but otherwise uses the same specification as the regressions reported in in Column 4 of Table 6. "Control for Drivers' Mean Exp" adds a control of each driver's average experience level across her weeks worked to this regression.

Figure 12: Return to Experience



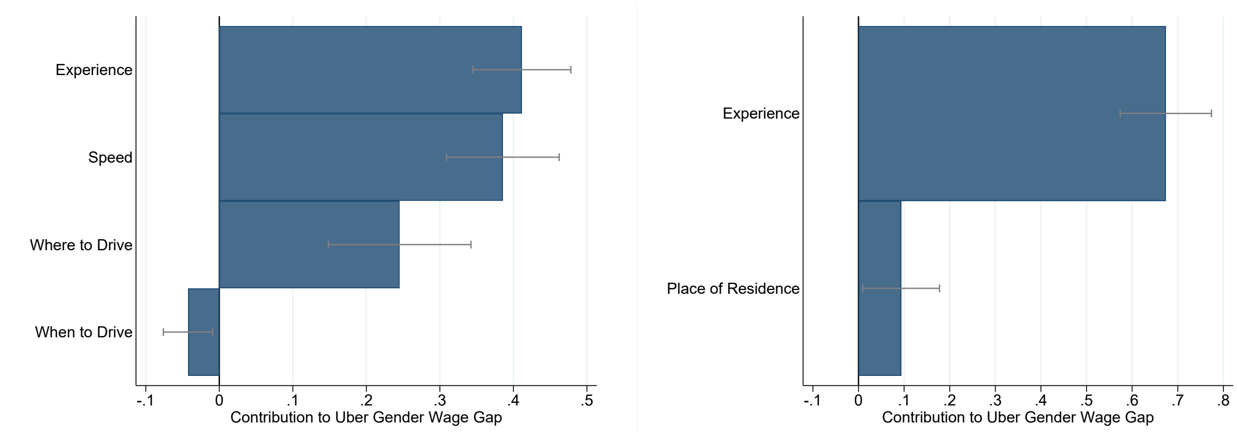
Note: This figure graphs the return to experience using a piece-wise linear spline, controlling for where drivers pick-up, the hour-of-week, week, and log driving speed. Range represents 95% confidence intervals. "Control for Drivers' Mean Exp" estimates use a piece-wise linear spline for experience, instead of dummies for bins of experience levels and controls for each driver's average experience level across her weeks worked, but otherwise uses the same specification as the regressions reported in in Column 4 of Table 6. "Control for Driver FEs" estimates use a piece-wise linear spline for experience, instead of dummies for bins of experience levels and controls for driver fixed effects, but otherwise uses the same specification as the regressions reported in in Column 4 of Table 6.

Figure 13: Difference in Return to Experience between Genders



Note: This figure graphs the difference in return to experience by gender using a piece-wise linear spline, controlling for where drivers pick-up, the hour-of-week, week, driver fixed effects, and log driving speed. Range represents 95% confidence intervals.

Figure 14: Gelbach decomposition



Note: These figures use the method described in Gelbach (2016) to plot the share of the gender pay gap that can be explained by each factor we consider: speed, experience (lifetime trips controls), where to drive (either geohashes or features of geohashes), and when to drive (hour of week controls).

A.11 Learning to accept/cancel trips strategically

Drivers can affect their earnings through strategic actions. We consider two such strategic actions: rejecting dispatches and canceling trips. When drivers receive a dispatch, they are told where the rider is and the estimated time-to-pickup. They can then choose to accept or reject the dispatch. This information can be valuable in assessing the quality of a given dispatch. If a rider is particularly far away, then there is an additional cost; drivers are not compensated for the time it takes to drive to meet a rider.[49] If a driver has reason to think that, by rejecting a ride, he or she will be offered a closer dispatch shortly, that driver may be able to increase expected earnings by not accepting the first dispatch. Savvy drivers will also realize that a high time-to-pickup ride may indicate an imbalance in supply and demand that may soon be corrected by a higher surge.[50]

Once a driver accepts a dispatch, the driver can cancel the trip before picking up the rider. After accepting, drivers are able to contact the rider. Some may do so to learn about the rider destination—for example, calling and asking if the rider is headed to the airport—and canceling if the driver believes the trip will not be worth the time.[51] Experienced drivers may also learn to cancel when they have reason to believe the rider will not show up.

In Table 18, we test whether men and women are more likely to reject a dispatch or cancel an accepted dispatch. In the average driver-hour, men reject 8.95% of dispatches and women reject 7.96% of dispatches. The gap persists when controlling for time and location (column 1) and experience (column 3). There is mixed evidence whether drivers learn to reject dispatches; drivers in their first 100 trips and with trips over 2,500 both cancel more than those with trips between 100-2500. After accepting a dispatch, male drivers are also more likely to cancel (3.12% of total dispatches, compared to 2.77% for women). This gap also persists once controlling for time and location (column 3). Part of the gap is due to experience; drivers cancel more as they gain experience (column 4), and men are more experienced on average.

[49] Effective October 2017, Uber initiated a system where drivers are paid (and riders are charged) for particularly long pickups.

[50] Surge rates update every two minutes.

[51] While this is feasible, it is against Uber's community guidelines, which prohibit "destination discrimination," and may result in deactivation. It is unclear how stringently these guidelines are enforced as identifying true destination discrimination is difficult.

Table 19 tests how these differences affect earnings. We use dummy variables that indicate whether a driver rejected a dispatch or canceled a trip during a given driver-hour to our prior earnings regressions. Controlling for time and geography, there is a negative impact on earnings of rejecting a dispatch or canceling a trip. However, this negative effect decreases as experience increases (while still remaining negative). Receiving a bad draw dispatch can never have a positive effect on earnings. A driver either completes the trip, which likely took longer than it was worth, or recognizes that it was a bad draw, rejects or cancels it, and then must wait for the next dispatch. As drivers gain experience, they can more accurately estimate the trade-off between rejecting and having to wait for a new dispatch versus accepting and completing a potentially low value trip.

These and earlier regressions show that drivers become more productive (and earn more) as they learn where to drive, when to drive, and how to strategically cancel and accept trips. However, even with controls for strategic rejecting and canceling, and when/where to driver, drivers with over 2500 trips make 7.2% more than those in their first 100 trips; there are substantial (but smaller) returns to experience that remain that go beyond these observable measures.

Table 18: Differences in rejection and cancellation rates

	Rejection rate		Cancellation rate	
	(1)	(2)	(3)	(4)
isMale	0.0099	0.0077	0.0028	0.0005
	(0.002)	(0.002)	(0.001)	(0.001)
Trips completed: 100-500		-0.0120		0.0049
		(0.001)		(0.000)
Trips completed: 500-1000		-0.0119		0.0090
		(0.001)		(0.000)
Trips completed: 1000-2500		-0.0052		0.0116
		(0.001)		(0.000)
Trips completed: >2500		0.0091		0.0144
		(0.001)		(0.001)
Intercept	0.0796	0.0853	0.0280	0.0214
	(0.001)	(0.001)	(0.000)	(0.000)
Week	X	X	X	X
Hour of week	X	X	X	X
Geohash	X	X	X	X
N	10,252,847	10,252,847	10,252,847	10,252,847
R^2	0.122	0.123	0.0284	0.038

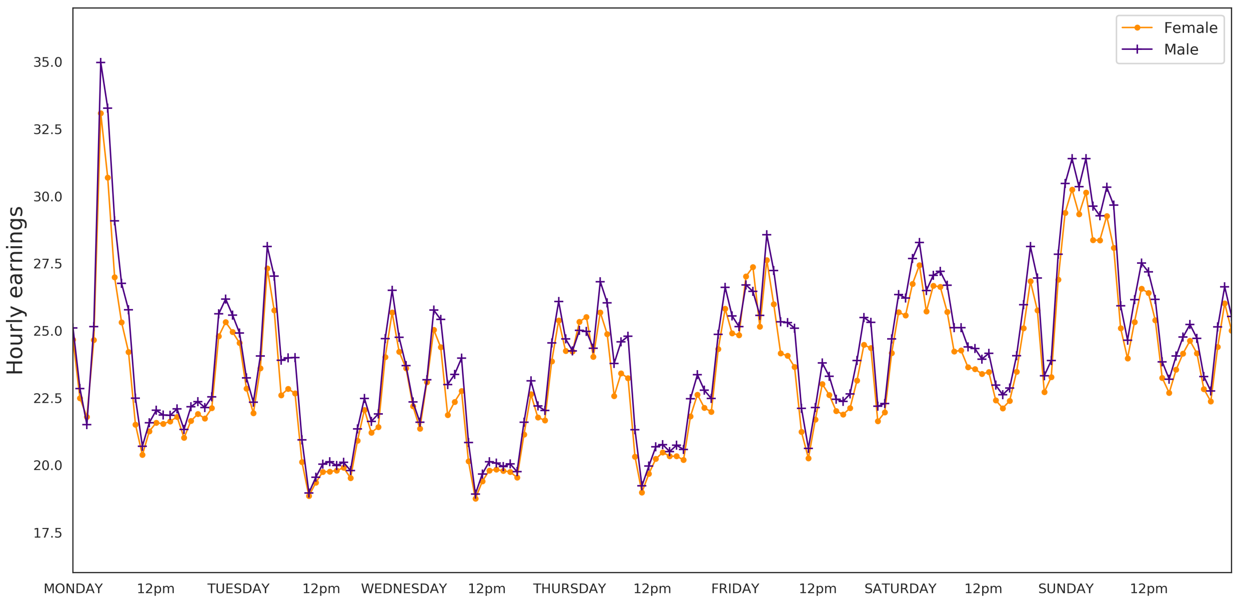
Note: This table documents differences in drivers propensity to reject a dispatch or cancel a dispatch they have already accepted. Data is at the driver-hour level; only hours with some dispatches are included (i.e. we drop driver-hours in which the only trip is a trip begun in the prior hour). Standard errors (clustered at the driver-level) are in parentheses.

A.12 Driving speed

We model driving speed against gender to test whether men drive faster after controlling for experience, location, and hour of week. Results presented in Table 20 show that men drive 2.2% faster after controls. Table 21 presents similar results based on data from the National Household Travel Survey; even in contexts where there is no pecuniary incentive to drive faster, men still do so.

A.13 Additional graphs & tables

Figure 15: Average earnings over course of week



Note: This figure graphs the average earnings by gender for different hours of the week. Data are limited to Chicago UberX/UberPOOL drivers in Chicago, January 2015-March 2017. Earnings include both incentive and organic earnings.

Table 19: Returns to strategic rejecting and canceling

	(1)	(2)
isMale	0.0142	0.0096
	(0.003)	(0.002)
Trips completed: 100-500	0.0471	0.0275
	(0.001)	(0.001)
Trips completed: 500-1000	0.0681	0.0409
	(0.002)	(0.001)
Trips completed: 1000-2500	0.0875	0.0539
	(0.002)	(0.002)
Trips completed: >2500	0.1192	0.0723
	(0.004)	(0.003)
rejectDispatch	−0.0757	−0.1099
	(0.001)	(0.001)
rejectDispatch*Trips completed: 100-500	0.0234	0.0218
	(0.002)	(0.002)
rejectDispatch*Trips completed: 500-1000	0.0367	0.0343
	(0.002)	(0.002)
rejectDispatch*Trips completed: 1000-2500	0.0520	0.0476
	(0.003)	(0.002)
rejectDispatch*Trips completed: >2500	0.0765	0.0705
	(0.004)	(0.004)
cancelTrip	−0.0227	−0.0785
	(0.002)	(0.002)
cancelTrip*Trips completed: 100-500	0.0112	0.0129
	(0.003)	(0.003)
cancelTrip*Trips completed: 500-1000	0.0242	0.0277
	(0.004)	(0.003)
cancelTrip*Trips completed: 1000-2500	0.0206	0.0363
	(0.003)	(0.003)
cancelTrip*Trips completed: >2500	0.0462	0.0571
	(0.004)	(0.003)
Intercept	3.0400	3.0847
	(0.003)	(0.002)
Week	X	X
Hour of week		X
Geohash		X
N	11,572,163	11,572,163
R^2	0.049	0.152

Note: This table expands on the regressions in Table 6 by adding covariates for whether a driver rejected a dispatch or canceled a trip in a given hour. Data are at the driver-hour level. The outcome variable is log of hourly earnings. Standard errors (clustered at the driver-level) are in parentheses.

Table 20: Effect of gender on driving speed

	(1)	(2)
isMale	0.0236	0.0218
	(0.002)	(0.002)
Trips completed: 100-500		0.0039
		(0.001)
Trips completed: 500-1000		0.0075
		(0.001)
Trips completed: 1000-2500		0.0096
		(0.002)
Trips completed: >2500		0.0110
		(0.002)
Intercept	2.9174	2.9119
	(0.001)	(0.001)
Week	X	X
Geohash*hour of week	X	X
N	11,572,163	11,572,163
R^2	0.352	0.352

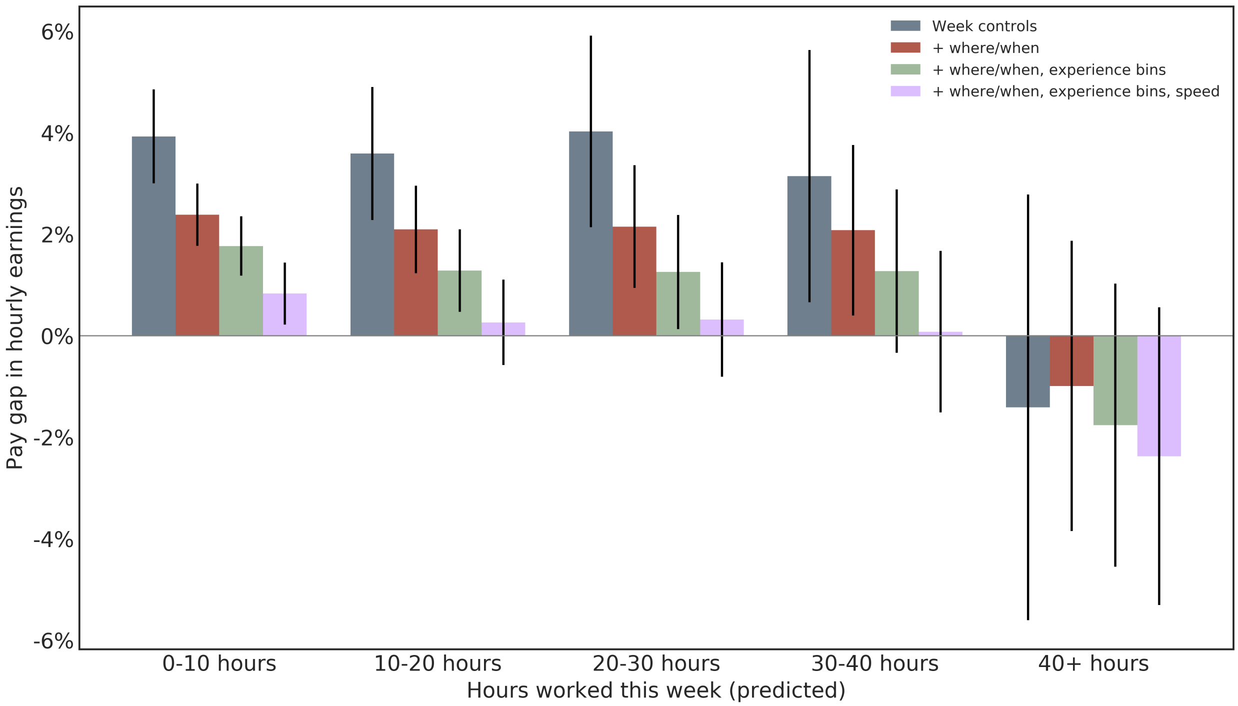
Note: This table regresses log speed in a given driver-hour against the driver's gender and experience. Speed is measured as distance traveled on-trip in an hour over duration on-trip. Standard errors (clustered at the driver-level) are in parentheses.

Table 21: Effect of gender on driving speed, NHTS data

	(1)	(2)	(3)	(4)
Male	0.0881	0.113	0.0494	0.0772
	(0.005)	(0.038)	(0.005)	(0.016)
Intercept	2.973	2.810	3.197	3.103
	(0.004)	(0.024)	(0.068)	(0.125)
N	656,904	3,677	656,904	656,904
R^2	0.004	0.007	0.124	0.582
Nationwide Sample	X		X	X
Chicagoland Sample		X		
Controls			X	X
Vehicle FE				X

Note: The table presents estimates the gender gap in log driving speed using data from the National Household Travel Survey. Dependent variable is average miles per hour driven on a single trip. Column 3 includes controls for household income bins, driver education bins, dummies for why the trip was taken, dummies for why the previous trip was taken, day of the week, hour of day, age dummies, MSA size bins, and whether the interstate was used on trip. Column 4 add individual vehicle fixed effects. Since each household only record trips on a single day, Column 4 only compares male and female speeds driven in the same vehicle on the same day. Standard errors are in parentheses.

Figure 16: Pay gap by levels of driving intensity



Note: This figure graphs the gender pay gap by different different driver types, defined by how often they are predicted to drive in the week of observation (predicted based on hours driven in past weeks). Drivers in their first week of work are not included. The set of controls mimics those used throughout the paper. Lines represent 95% confidence intervals based on standard errors clustered at the driver-level.