The Impact of Behavioral and Economic Drivers on Gig Economy Workers

INFORMS 2019

Park Sinchaisri (Wharton)
Gad Allon (Wharton), Maxime Cohen (McGill)

Gig Economy

Gig Economy

freelancing upwork fiverr

Gig Economy

freelancing Upwork fiverr

local tasks TaskRabbit handy

Gig Economy

freelancing — Upwork, fiverr

local tasks — TaskRabbit, handy

ride-hailing — lyft, Uber, Grab

delivery — instacart, DOORDASH

Gig Economy

freelancing Upwork fiverr

local tasks TaskRabbit handy

retail (allwork) snag.work on-demand

ride-hailing lyft Uber Grab

delivery instacart DOORDASH

Gig Economy

2017

57.3 Million

= 36% of US workforce

JP Morgan Chase Institute / WSJ / Staffing Industry Analysts / McKinsey & Company

Gig Economy

2017

57.3 Million

= 36% of US workforce

2027

60% of work

+ $2.7 trillion global GDP

JP Morgan Chase Institute / WSJ / Staffing Industry Analysts / McKinsey & Company

Who are Gig Workers?

70% by choice

44% primary income

~50% millennials/gen z

2016 McKinsey Global Institute

Who are Gig Workers?

70% by choice

44% primary income

~50% millennials/gen z

when to work?

how long?

which platforms?

Workers decide work schedules

2016 McKinsey Global Institute

Gig Company

Supply $\xleftrightarrow[\text{real-time}]{\text{match}}$ Demand

independent/flexible
on-demand

Gig Company

Supply ⟵ match / real-time ⟶ Demand

independent/flexible
on-demand

Workforce planning is challenging

Research Questions

How do gig economy workers make labor decisions?

How can the platform influence their decisions?

Research Questions

Econometrics

How do gig economy workers make labor decisions?

How can the platform influence their decisions?

Simulation

Structural

Outline

- What has been done

  Practice / labor economics / OM / hypotheses
- Data and empirical strategy

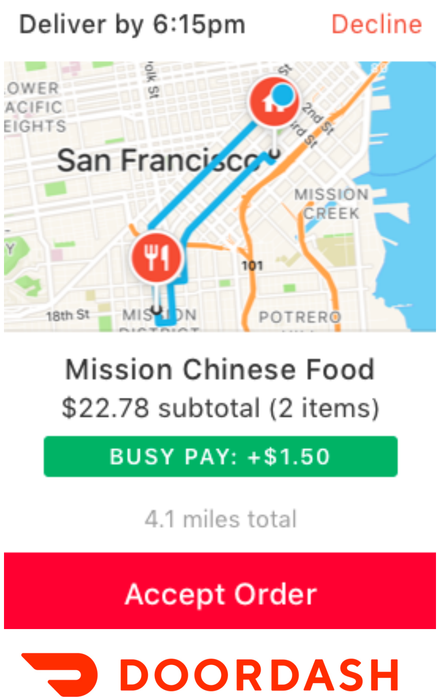
  Dealing with endogeneity and selection bias
- Results

  Impact of incentive and behavioral elements on labor decisions
- Implications

  Simulation of optimal incentive re-allocation

In Practice

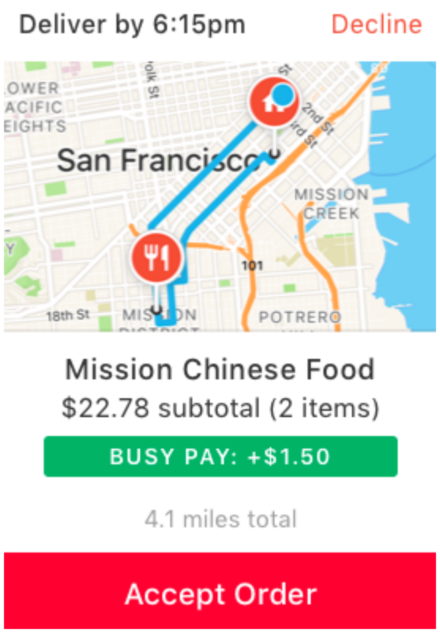
Real-time "surge pricing"



https://dasherhelp.doordash.com/busy-pay

In Practice

Real-time "surge pricing"



DOORDASH

Pre-announced bonus

5:00 PM–6:00 PM
+10% (5:00pm - 5:30pm)
+30% (5:30pm - 6:00pm)

6:00 PM–7:00 PM
+30% (6:00pm - 6:30pm)
+40% (6:30pm - 7:00pm)

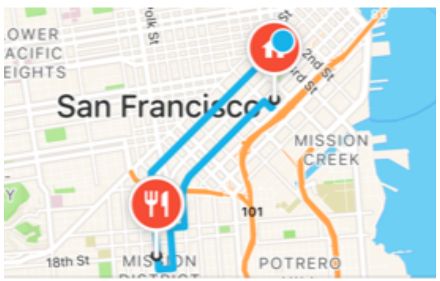
caviar

https://dasherhelp.doordash.com/busy-pay
https://courierhelp.trycaviar.com/customer/en/portal/articles/2821000-peak-hour-pay?b_id=9619/

In Practice

Real-time “surge pricing”

Deliver by 6:15pm   Decline



Mission Chinese Food
$22.78 subtotal (2 items)

BUSY PAY: +$1.50

4.1 miles total

Accept Order

DOORDASH

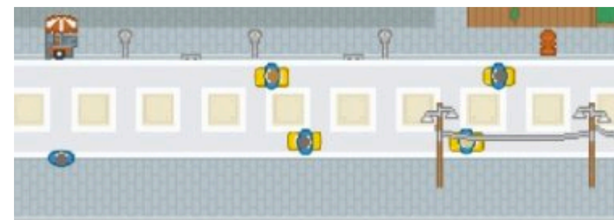
Pre-announced bonus

5:00 PM–6:00 PM
+10% (5:00pm - 5:30pm)
+30% (5:30pm - 6:00pm)

6:00 PM–7:00 PM
+30% (6:00pm - 6:30pm)
+40% (6:30pm - 7:00pm)

caviar

“You’re so close to your precious target”



How Uber Uses Psychological Tricks to Push Its Drivers’ Buttons

https://dasherhelp.doordash.com/busy-pay
https://courierhelp.trycaviar.com/customer/en/portal/articles/2821000-peak-hour-pay?b_id=9619/
https://nytimes.com/interactive/2017/04/02/technology/uber-drivers-psychological-tricks.html

Theories of Labor Supply

Theories of Labor Supply

Neoclassical

- Maximize lifetime utility
- **Positive** income elasticities

Wage $\uparrow$

Work more

Theories of Labor Supply

Neoclassical

- Maximize lifetime utility
- **Positive** income elasticities

Wage $\uparrow$
Work more

Behavioral

- Reference-dependence, targets
- **Negative** income elasticities

Wage $\uparrow$
Work less

Theories of Labor Supply

Neoclassical

- Maximize lifetime utility
- **Positive** income elasticities

Behavioral

- Reference-dependence, targets
- **Negative** income elasticities

Carrington (1996)

Camerer et al. (1997)

Oettinger (1999)

Farber (2005)

Fehr/Goette (2007)

Farber (2008)

Stafford (2013)

Farber (2015)

Chen/Sheldon (2016)

Sheldon (2016)

Thakral/To (2017)

Wage ↑
Work more

Wage ↑
Work less

Recent Work in OM

Theoretical

Dong & Ibrahim (2018)
Taylor (2018)
Cachon, Daniels & Lobel (2017)
Hu & Zhou (2017)
Ibrahim (2017)
Bimpikis, Candogan & Saban (2017)
Gurvich, Lariviere & Moreno (2016)
Tang et al. (2016)
Banerjee, Riquelme & Johari (2016)
Benjaafar et al. (2015)
…

Empirical

Sun, Wang & Wan (2019)
Kabra, Belavina & Girotra (2018)
Karacaoglu, Moreno & Ozkan (2018)
Chen, Chevalier, Rossi & Oehlsen (2017)
Cui, Li & Zhang (2017)
Li, Moreno & Zhang (2016)
…

Our Paper

- Behavioral drivers of decisions
- Rich data with complete description of the supply side

Drivers of Work Decisions

We are interested in three effects

Hourly Wage

Income Target

Time Target

on two work decisions:

Work or not?

If so, how long?

Drivers of Work Decisions

We are interested in three effects

	"ISF"	"HSF"
Hourly Wage	Income So Far /Income Target	Hours So Far /Time Target

on two work decisions:

Work or not?	If so, how long?

Drivers of Work Decisions

We are interested in three effects

Hourly Wage	"ISF" Income So Far /Income Target	"HSF" Hours So Far /Time Target
H1: Positive	**H2:** Negative	**H3:** Negative
Carrington (1996), Oettinger (1999), Stafford (2015)	Farber (2008), Thakral & To (2017)	Crawford & Meng (2011), Farber (2015), Agarwal et al (2015), Brachet et al (2012), Collewet & Sauermann (2017)

on two work decisions:

Work or not?

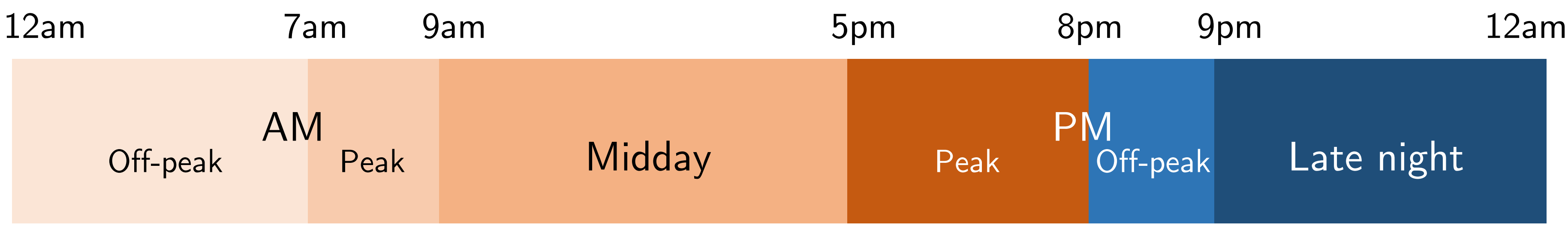
If so, how long?

Data US ride-hailing firm

Drivers are guaranteed an hourly $\underbrace{\text{Base Rate} + \text{Promotions}}_{\text{“Offer”}}$

Data US ride-hailing firm

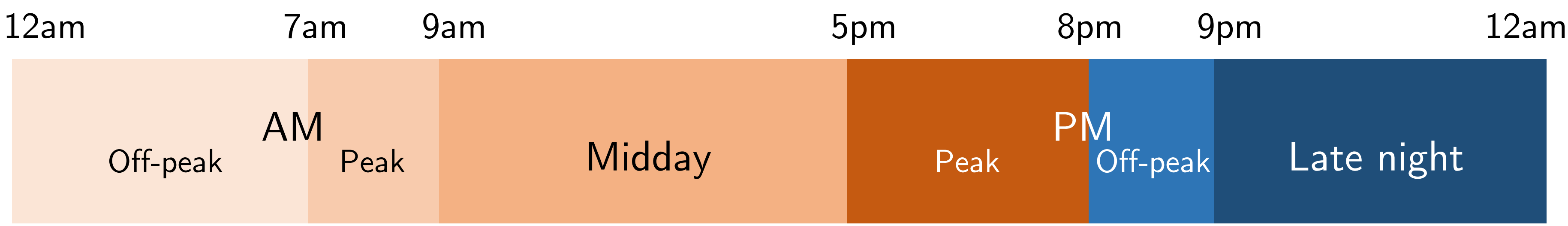
Drivers are guaranteed an hourly **Base Rate** + **Promotions**



Shift-level financial incentives and driving activity *for all*

Data
US ride-hailing firm

Drivers are guaranteed an hourly **Base Rate** + **Promotions**



Shift-level financial incentives and driving activity *for all*

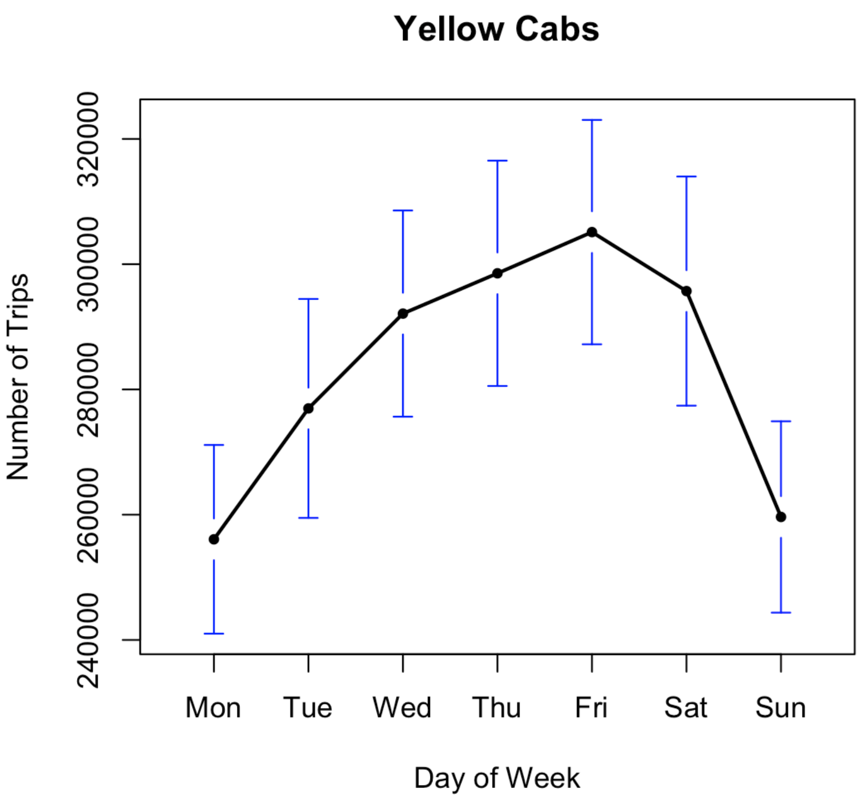
5.5M
Observations

358
Days
Oct 2016 – Sep 2017

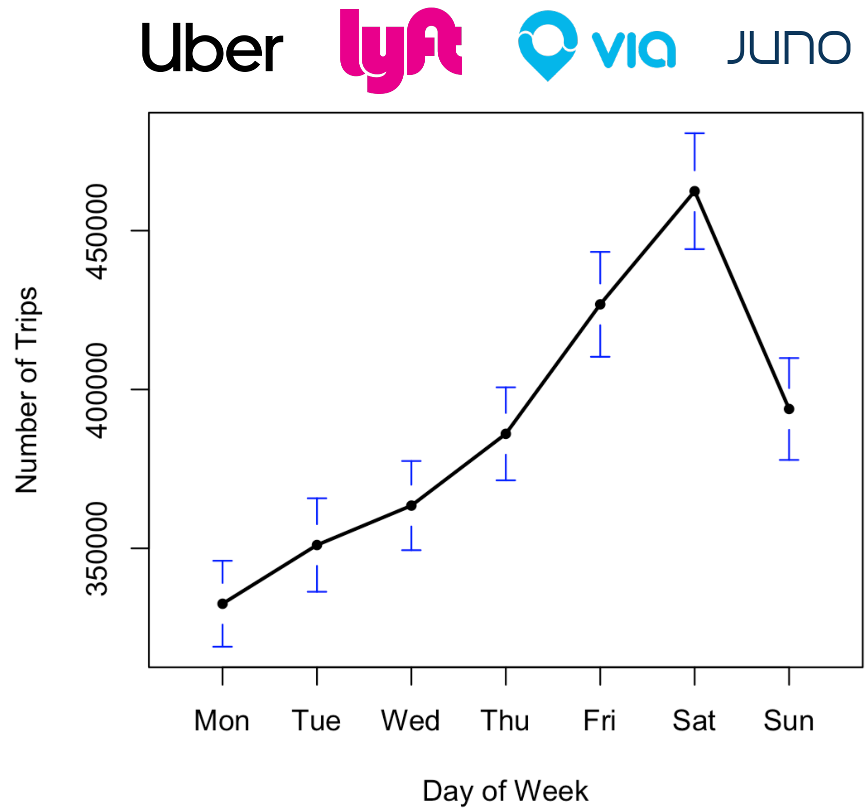
7,826
Unique drivers
SUV/Sedan/Van

+ TLC Data

Trip records collected by Taxi & Limousine Commission



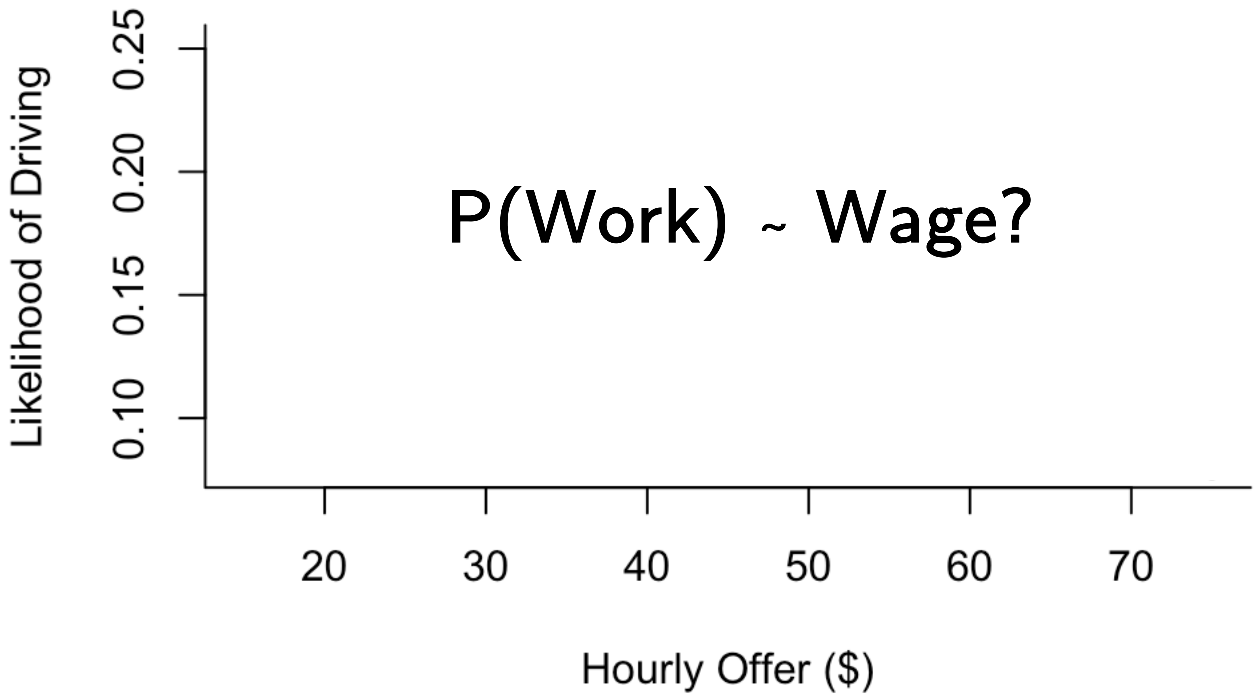
101M yellow cab trips



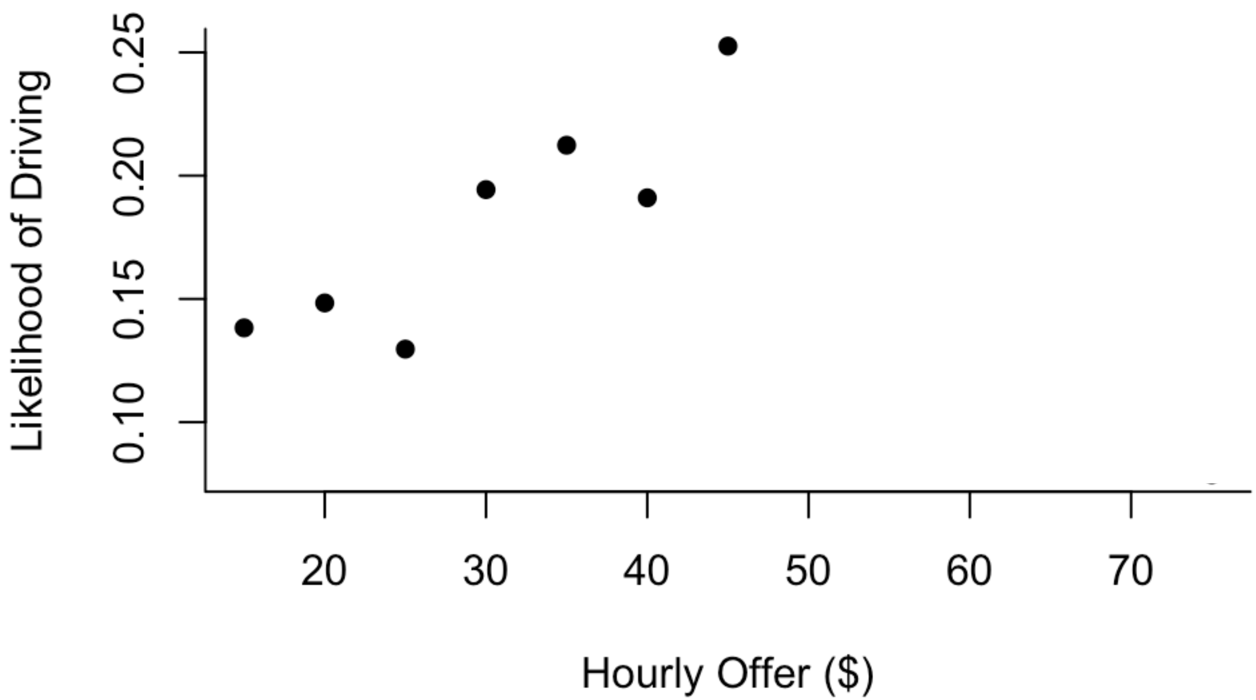
139M ride-hailing trips

Empirical Strategy Challenges

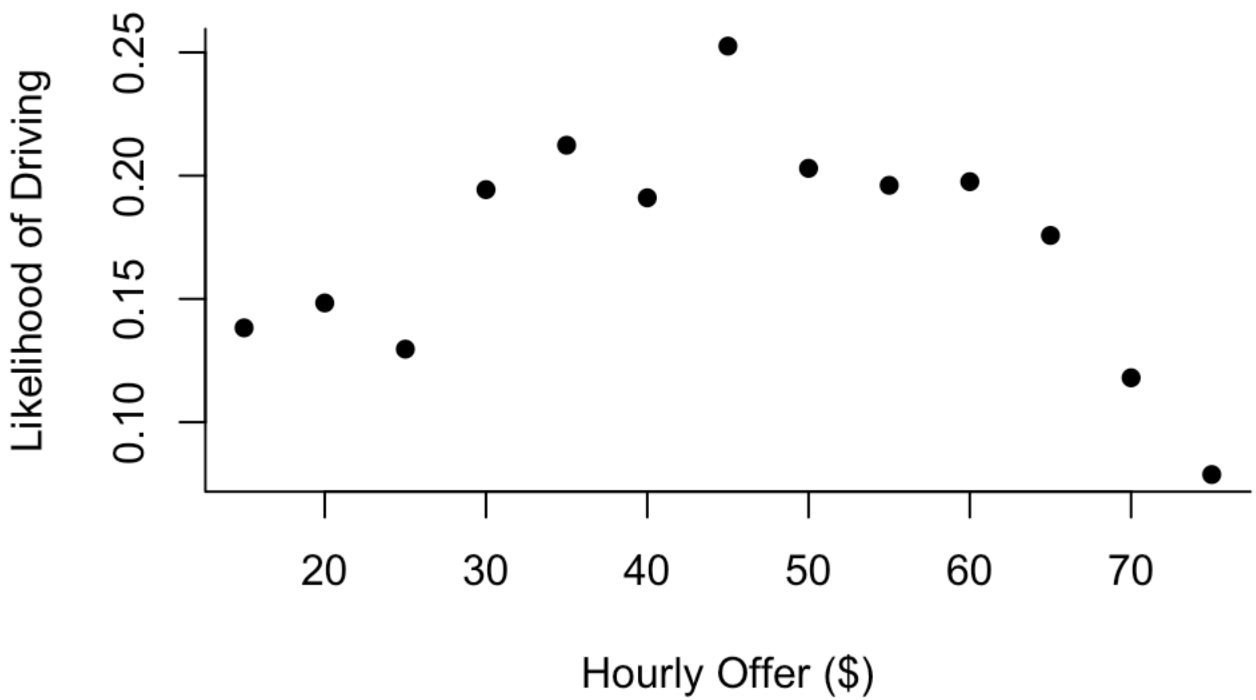
Empirical Strategy Challenges



 Challenges


Likelihood of Driving
0.10
0.15
0.20
0.25
20
30
40
50
60
70
Hourly Offer ($)

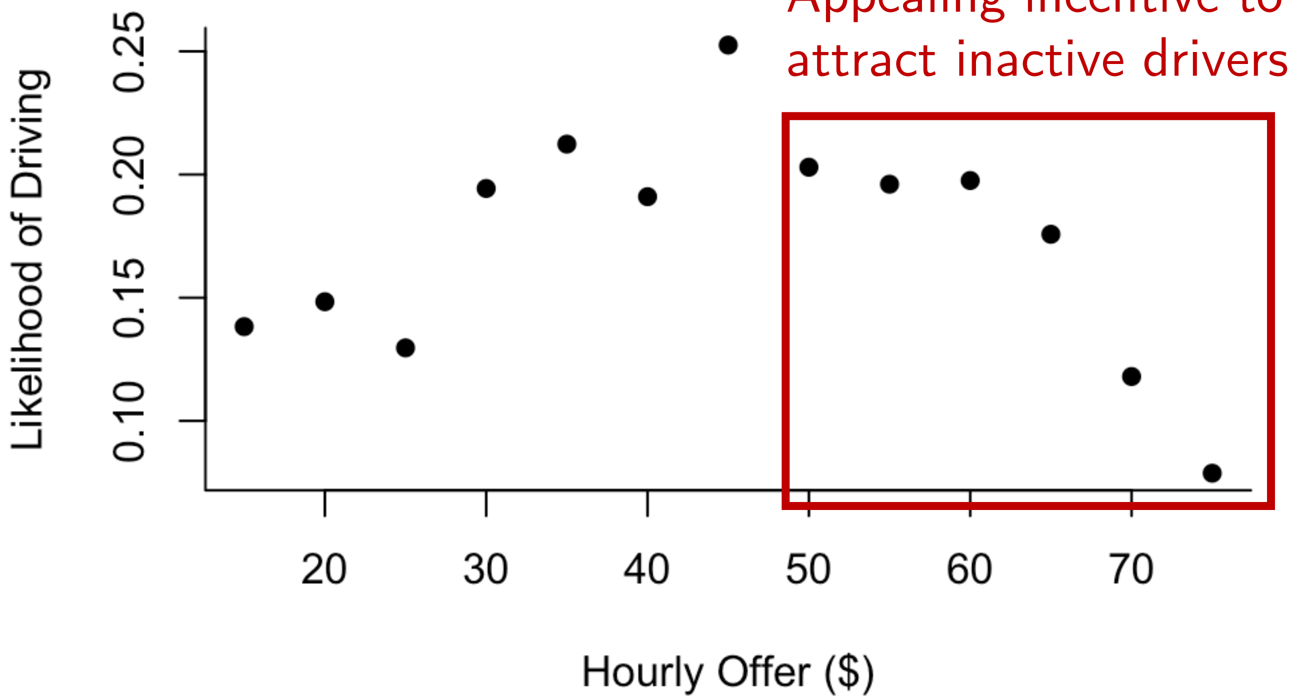
Empirical Strategy Challenges



Empirical Strategy Challenges

Higher wage, less likely to drive?

Appealing incentive to attract inactive drivers?

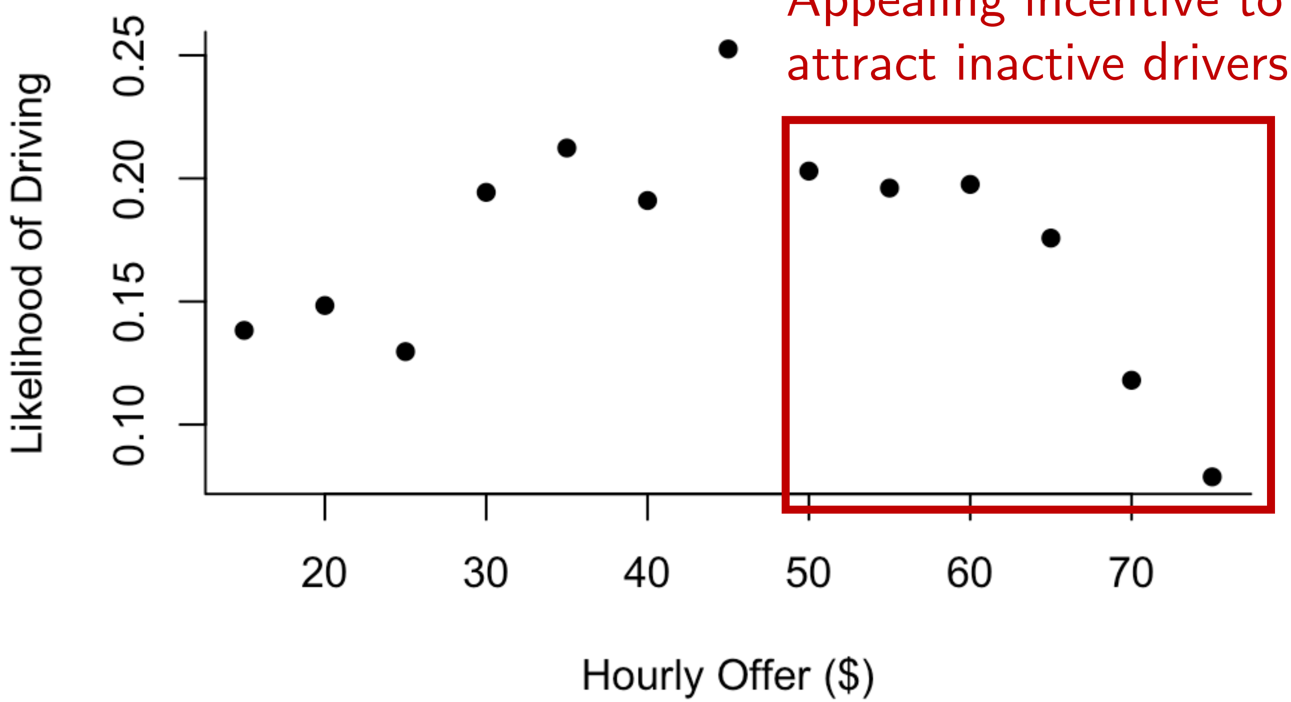


Empirical Strategy Challenges

Simultaneity

Higher wage, less likely to drive?

Appealing incentive to attract inactive drivers?



Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Endogenous Variable

Hourly offer

Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Endogenous Variable	Instrument
Hourly offer	Average offers of “co-workers”

Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Endogenous Variable	Instrument
Hourly offer	Average offers of "co-workers"

= currently available
+ made similar decisions
+ different vehicle type

Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

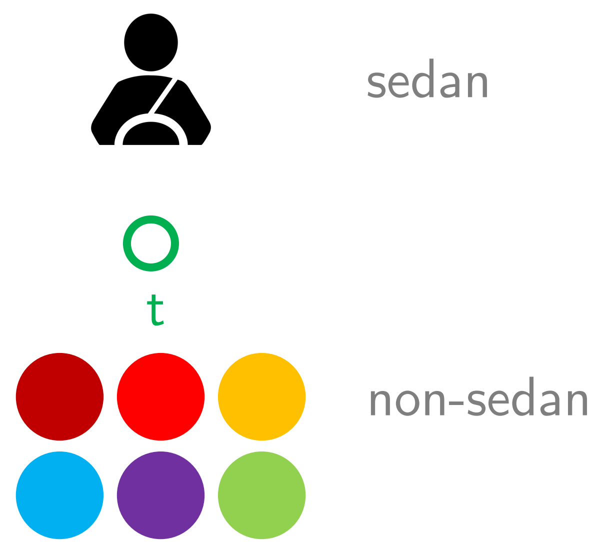
Endogenous Variable	Instrument
Hourly offer	Average offers of "co-workers"

Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Endogenous Variable	Instrument
Hourly offer	Average offers of "co-workers"

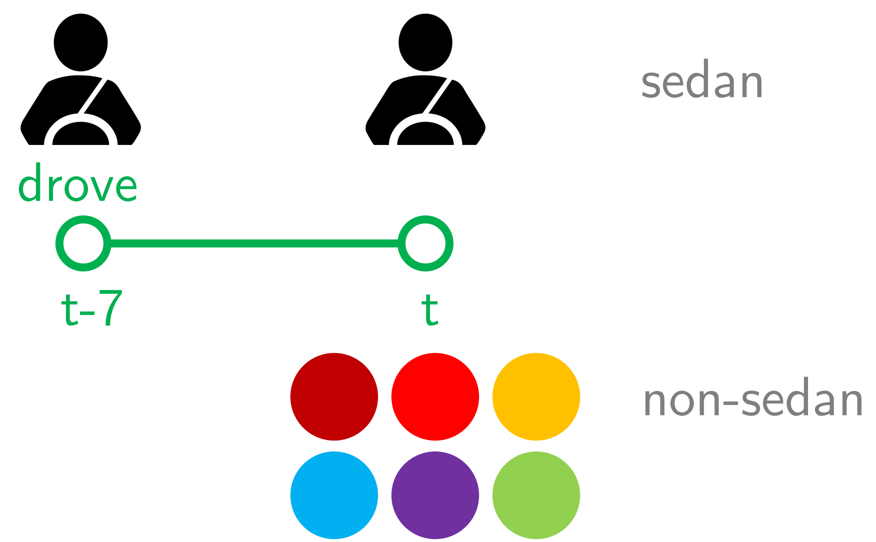


Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Endogenous Variable	Instrument
Hourly offer	Average offers of “co-workers”



Empirical Strategy Challenges

Simultaneity

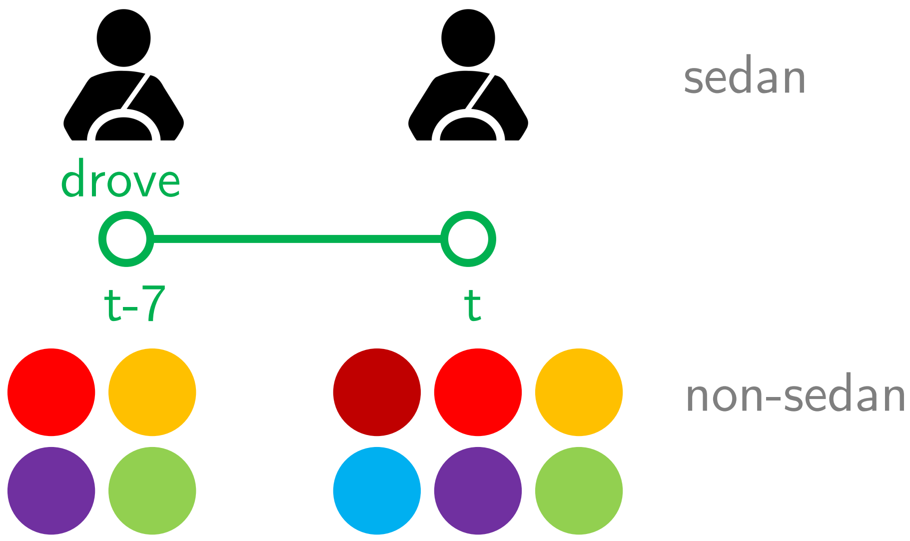
Solution: Instrumental Variables

Endogenous Variable

Hourly offer

Instrument

Average offers of "co-workers"



Empirical Strategy Challenges

Simultaneity

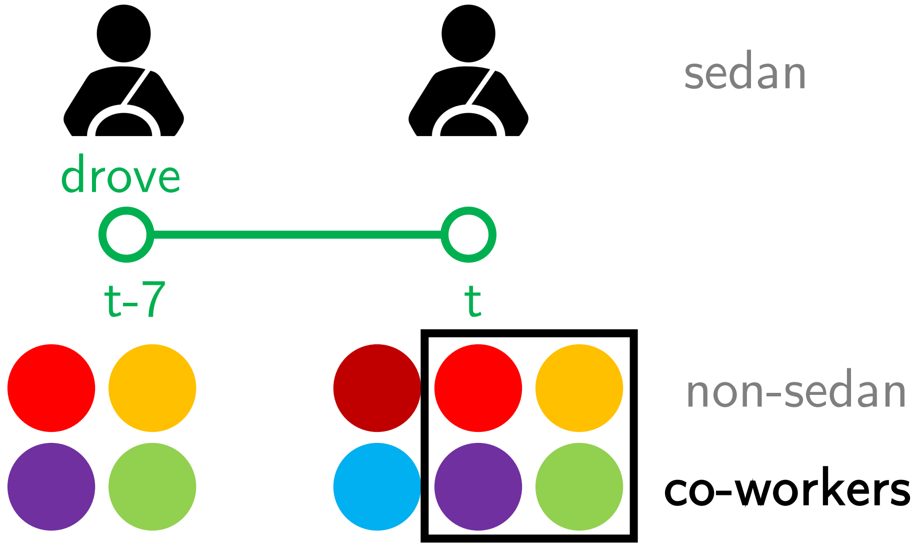
Solution: Instrumental Variables

Endogenous Variable

Hourly offer

Instrument

Average offers of "co-workers"

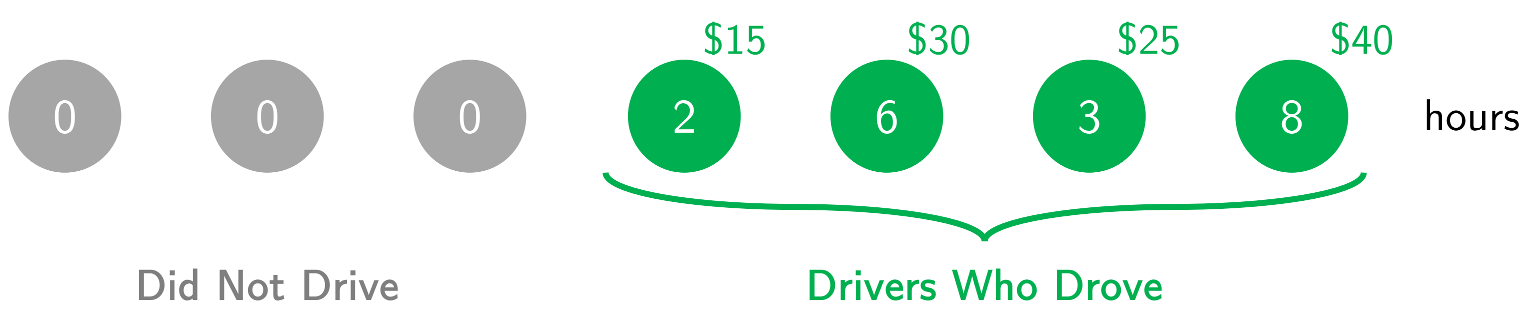


Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Hours ~ Wage?



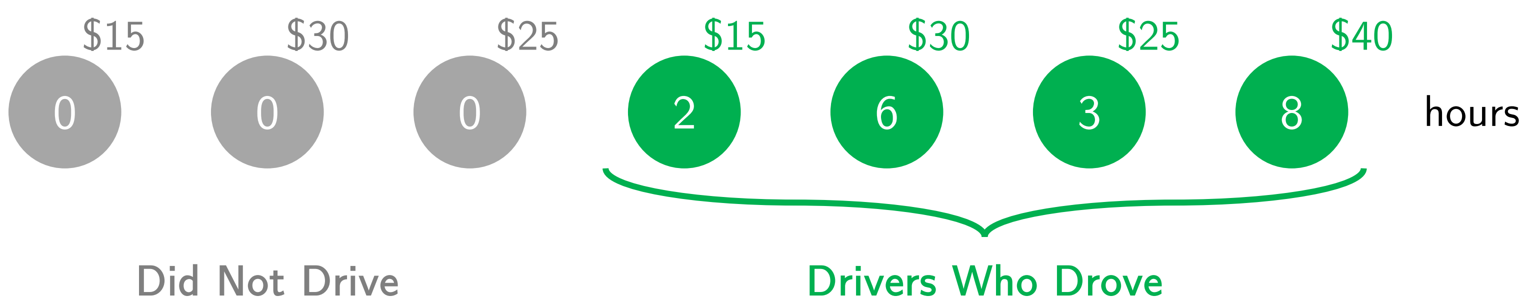
Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Decision to work is **not random**

Hours ~ Wage?



Empirical Strategy Challenges

Simultaneity

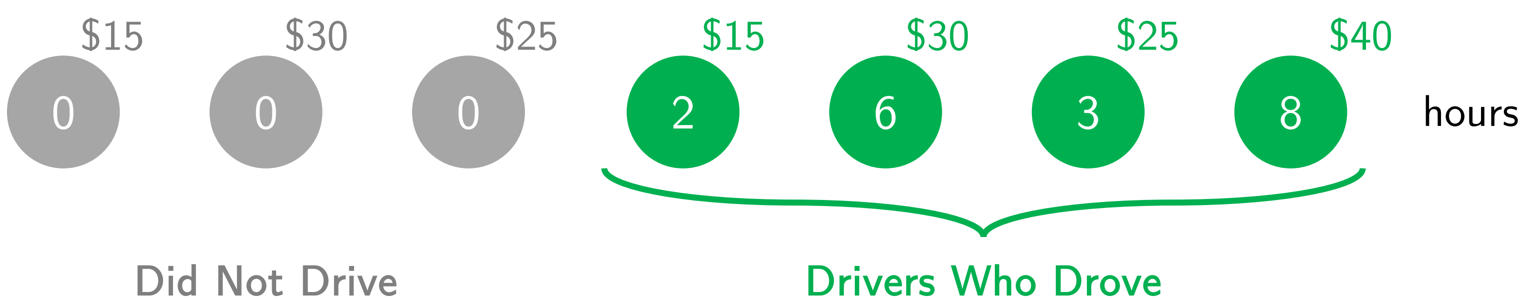
Solution: Instrumental Variables

Selection Bias

Solution: Modified Heckman Two-Stage Method

("Heckit" - Heckman 1979)

Hours ~ Wage



Empirical Strategy Heckman + IV

1. Work or not?

Control Function Probit:

P(drive) on ***Offer***

Empirical Strategy Heckman + IV

1 **Work or not?**

Control Function Probit:

P(drive) on ***Offer*** + Controls_1

Demand

Short-term Habits

Long-term Habits

Empirical Strategy Heckman + IV

1 **Work or not?**

Control Function Probit:

P(drive) on ***Offer*** + Controls$_1$

Demand
- **# Ride-Hailing Pickups** right before
- **Hourly Weather** Humidity, Temp, PrecipProb
- **Holiday, Day of Week**
- **Month-Year FE**

Short-term Habits
- **Past Work Habits**
  Hours worked last week, income last week

Long-term Habits
- **Driver's Experience** New?
- **Driver's FE***

*Bias corrected with panel jackknife (Hahn & Newey 2004)

Empirical Strategy Heckman + IV

1. **Work or not?**

Control Function Probit:

P(drive) on ***Offer*** + **ISF** + **Controls**$_1$

Income So Far
= cumulative income since beginning of day

Empirical Strategy Heckman + IV

1 **Work or not?**

Control Function Probit:

P(drive) on ***Offer*** + ISF + HSF + Controls$_1$

Income So Far

Hours So Far
= cumulative active hours
since beginning of day

Empirical Strategy Heckman + IV

1 **Work or not?**

Control Function Probit:

P(drive) on ***Offer*** + **ISF** + **HSF** + **Controls**$_1$

Income So Far **Hours So Far**

Conditional on working

2 **How long to work?**

2SLS with Fixed Effects

 # Hours on ***Earning*** + **ISF** + **HSF** + + **Controls**$_2$

Empirical Strategy Heckman + IV

1 Work or not?

Control Function Probit:

P(drive) on ***Offer*** + ISF + HSF + Controls_1

Income So Far Hours So Far

Conditional on working

2 How long to work?

2SLS with Fixed Effects

Hours on ***Earning*** + ISF + HSF + IMR + Controls_2

Inverse Mills Ratio
= correct for selection bias

Empirical Strategy Heckman + IV

1 Work or not?

CF: Regress **hourly offer** on IVs. Keep **residuals**

Probit: Estimate **P(drive)**

$$P(Drive_{i,t} = 1 | \boldsymbol{X_{i,t}}) = \Phi(\alpha_{0,t} + \alpha_w w_{i,t} + \boldsymbol{\alpha X_{i,t}} + \alpha_e \hat{e}_{i,t})$$

Bias corrected with panel jackknife (Hahn & Newey 2004)

Inverse Mills Ratio (IMR)

$$\lambda(c_z) = \frac{\phi(c_z)}{1 - \Phi(c_z)}$$

Conditional on working

2 How long to work?

IV: Estimate **hourly earning** from IVs

OLS: Estimate **hours**

$$f(Hour_{i,t}) = \beta_{0,i} + \beta_\omega \omega_{i,t} + \boldsymbol{\beta Z_{i,t}} + \theta \lambda_{i,t} + u_{i,t}$$

Adjust standard errors to account for the fact that IMR is an estimate (and hence random) covariate in the above model.

Results

"Income So Far" **"Hours So Far"**

Financial Offer → **1 Work or not?**

ISF → **2 # Hours**

HSF → **2 # Hours**

Within-Day

Midday → Late Night

Across-Days

Tuesday → Sunday

Results

"Income So Far" **"Hours So Far"**

Financial Offer ISF HSF

1 **Work or not?** 2 **# Hours**

Within-Day

Midday

↓

Late Night

Across-Days

Tuesday

↓

Sunday

+	-
Positive	Negative

Results Within Day

1

Work or not?

SUV	Mean	IV-F	N
Midday	0.343	372.9	124,769
PM-Peak	0.277	345.1	131,910
PM-OPeak	0.182	320.6	130,651
Late Night	0.117	379.0	125,382

Results Within Day

1

Work or not?

SUV	Mean	IV-F	Offer	N
Midday	0.343	372.9	+	124,769
PM-Peak	0.277	345.1	-	131,910
PM-OPeak	0.182	320.6	+	130,651
Late Night	0.117	379.0	+	125,382

Financial incentives have a generally **positive** impact.

Results Within Day

1

Work or not?

SUV	Mean	IV-F	Offer	ISF	N
Midday	0.343	372.9	+	-	124,769
PM-Peak	0.277	345.1	-	-	131,910
PM-OPeak	0.182	320.6	+	-	130,651
Late Night	0.117	379.0	+	-	125,382

Income Targeting:
The more you earned, the less likely you'll work a new shift.

Income targeting has a **consistently negative** impact.

Results Within Day

1

Work or not?

SUV	Mean	IV-F	Offer	ISF	HSF	N
Midday	0.343	372.9	+	-	+	124,769
PM-Peak	0.277	345.1	-	-	+	131,910
PM-OPeak	0.182	320.6	+	-	+	130,651
Late Night	0.117	379.0	+	-	+	125,382

Inertia:
The longer you've been active, the more likely you'll work another shift.

Inertia has a **consistently positive** impact.

Results Within Day

	1						2		
	Work or not?						# Hours		
SUV	Mean	IV-F	Offer	ISF	HSF	N	Mean	IV-F	N
Midday	0.343	372.9	+	-	+	124,769	4.987	180.2	45,329
PM-Peak	0.277	345.1	-	-	+	131,910	2.421	58.5	39,592
PM-OPeak	0.182	320.6	+	-	+	130,651	0.731	50.5	26,699
Late Night	0.117	379.0	+	-	+	125,382	1.996	39.91	17,137

Results Within Day

			1 Work or not?				2 # Hours						
SUV	Mean	IV-F	Offer	ISF	HSF	N	Mean	IV-F	Earn	ISF	HSF	R^2	N
Midday	0.343	372.9	+	-	+	124,769	4.987	180.2	+	-	+	0.552	45,329
PM-Peak	0.277	345.1	-	-	+	131,910	2.421	58.5	+	-	+	0.244	39,592
PM-OPeak	0.182	320.6	+	-	+	130,651	0.731	50.5	+	-	+	0.281	26,699
Late Night	0.117	379.0	+	-	+	125,382	1.996	39.91	+	-	+	0.296	17,137

The three effects are consistent in both stages

Results Across Days

Results Across Days

1. Work or not?

SUV	Mean	IV-F	N
Tuesday	0.409	43.6	28,883
Wednesday	0.418	55.9	21,965
Thursday	0.426	73.4	29,233
Friday	0.412	74.0	20,294
Saturday	0.203	98.1	15,788
Sunday	0.162	82.2	13,025

Results Across Days

1. Work or not?

SUV	Mean	IV-F	Offer	ISF	HSF	N
Tuesday	0.409	43.6	+	+	+	28,883
Wednesday	0.418	55.9	+	+	+	21,965
Thursday	0.426	73.4	+	+	+	29,233
Friday	0.412	74.0	+	+	+	20,294
Saturday	0.203	98.1	-	-	+	15,788
Sunday	0.162	82.2	-	-	+	13,025

Inertia has consistent positive effect.

Financial incentive and cumulative income switch from positive to negative later on.

Results Across Days

1 Work or not?

SUV	Mean	IV-F	Offer	ISF	HSF	N
Tuesday	0.409	43.6	+	+	+	28,883
Wednesday	0.418	55.9	+	+	+	21,965
Thursday	0.426	73.4	+	+	+	29,233
Friday	0.412	74.0	+	+	+	20,294
Saturday	0.203	98.1	-	-	+	15,788
Sunday	0.162	82.2	-	-	+	13,025

2 # Hours

	IV-F	Earn	ISF	HSF	R^2	N
Tuesday	18.3	-	-	+	0.422	9,482
Wednesday	26.2	-	-	+	0.422	10,120
Thursday	34.6	-	-	+	0.412	9,894
Friday	33.7	+	-	+	0.436	9,283
Saturday	19.1	-	+	-	0.398	4,372
Sunday	15.1	+	+	-	0.390	3,240

Hours not decided at the day level

Results Summary

As day/week proceeds…

Neoclassical Financial Incentive	encourages working

Results Summary

As day/week proceeds...

Neoclassical Financial Incentive	encourages working
Behavioral Income Target	discourages working later on

Results Summary

	As day/week proceeds...
Neoclassical Financial Incentive	encourages working
Behavioral Income Target	discourages working later on
New Inertia	encourages working

Robustness Tests

- Isolating ISF and HSF effect
  - Positive HSF (inertia) effect dominates ISF (targeting) effect.
- Nonlinearity in ISF and HSF
- Instrumenting for ISF and HSF
- Controlling on types of promotions
- Other approaches to sample selection
  - Two-part models: insights stay the same in both parts
  - Dahl's correction: using B-splines instead of IMR
- Looking at granular log-on/log-off/break data
  - Smaller sample of sedan and SUV drivers
  - The longer the previous active session is,
    the longer the current active session is.

Outline

- What has been done

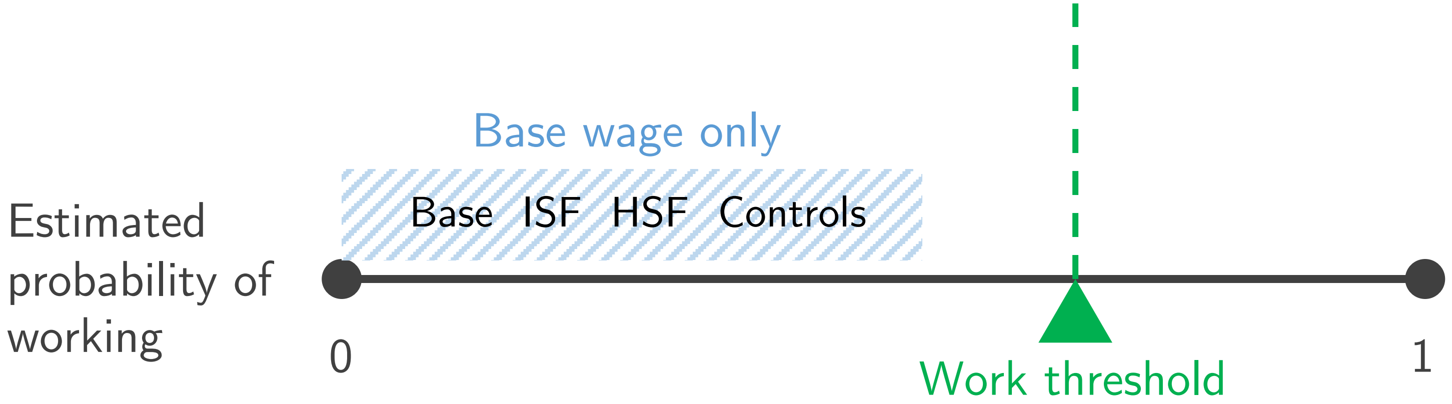
  Practice / labor economics / OM / hypotheses
- Data and empirical strategy

  Dealing with endogeneity and selection bias
- Results

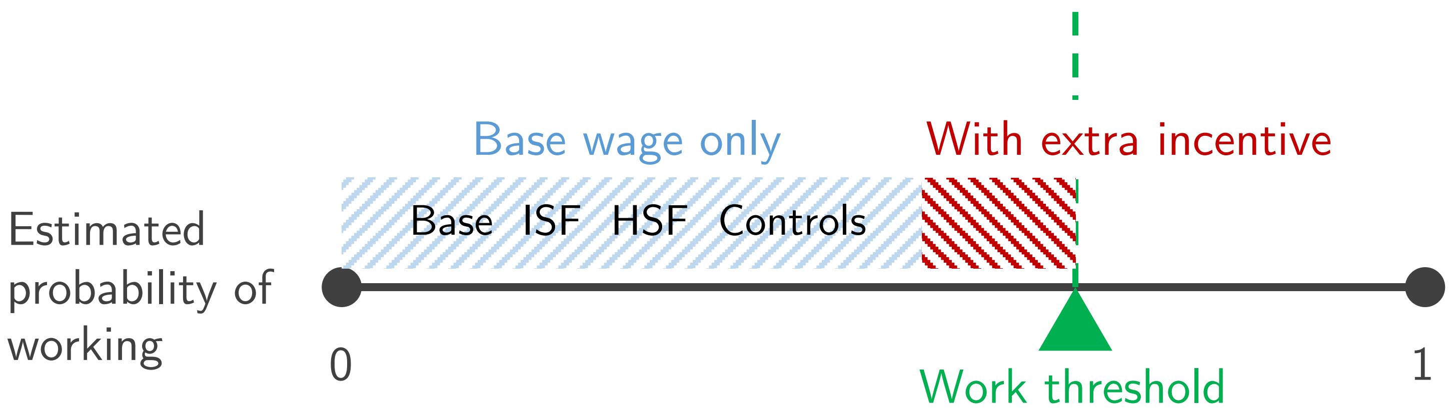
  Impact of incentive and behavioral elements on labor decisions
- Implications

  Simulation of optimal incentive re-allocation

Optimal Targeted Incentive

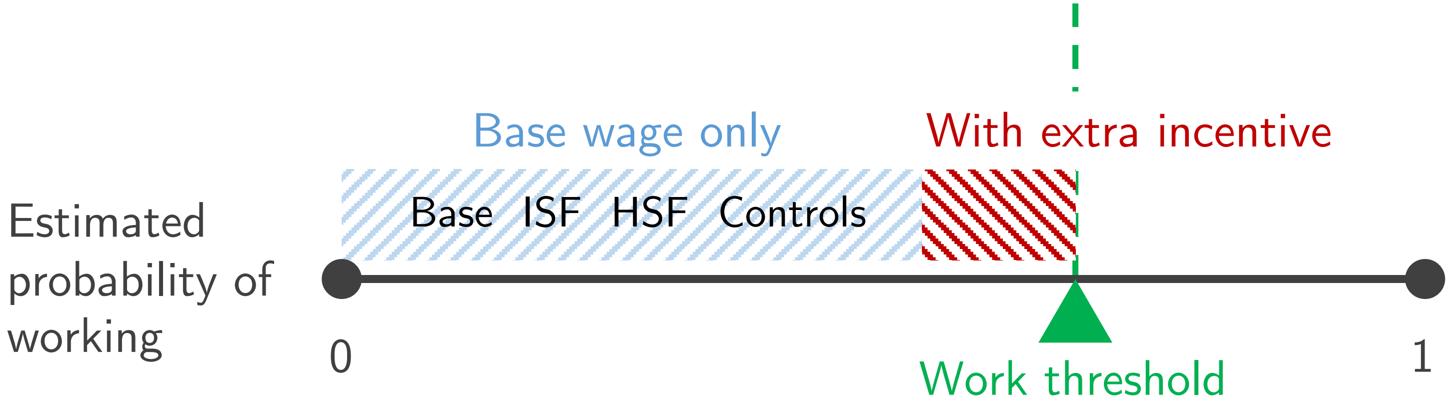


Optimal Targeted Incentive



Optimal Targeted Incentive

Ranking each driver by her
minimum work-inducing incentive
= how much to trigger working decision



Reallocating Incentives

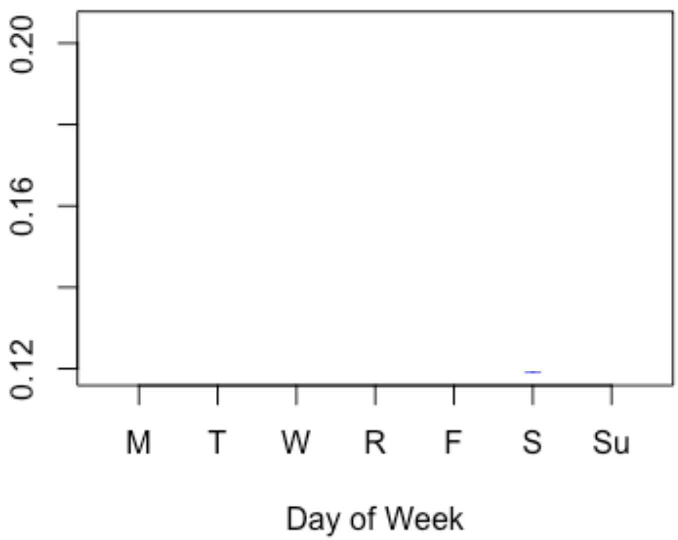
Compared to current practice from Jan to Sep 2017 out-of-sample
(Using data from Oct 2016 to right before the focal date as training)

Reallocating Incentives

Compared to current practice from Jan to Sep 2017 out-of-sample

Given the same budget

Extra drivers

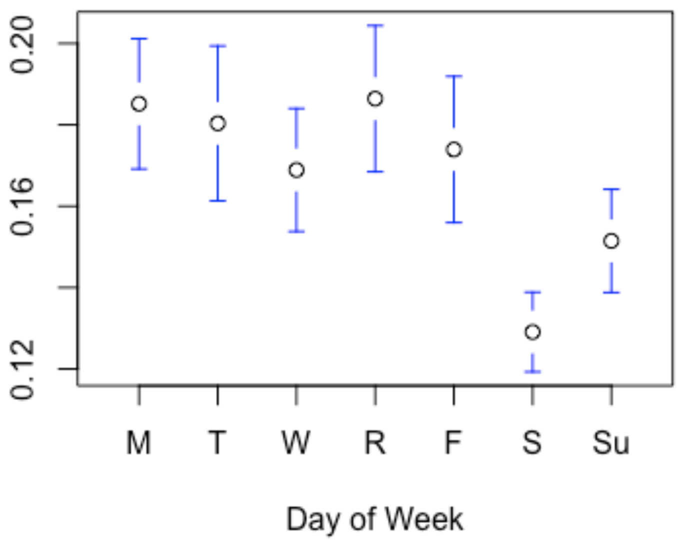


Reallocating Incentives

Compared to current practice from Jan to Sep 2017 out-of-sample

Given the same budget

Extra drivers



Can recruit **17% more drivers**

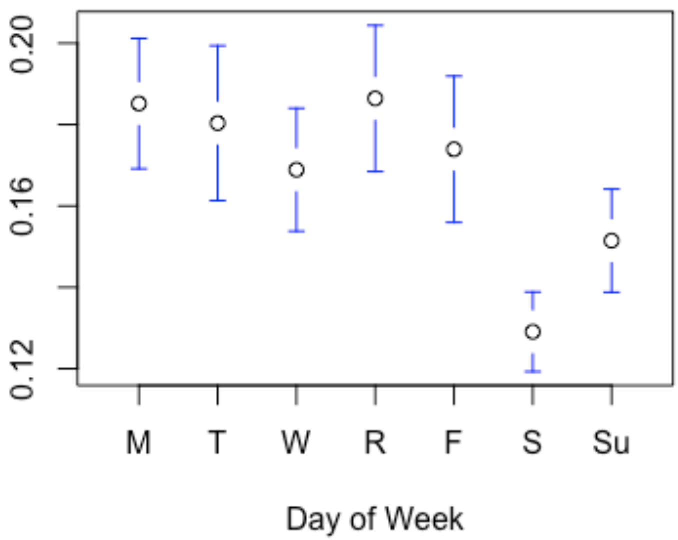
Average promo: 1.61x

Reallocating Incentives

Compared to current practice from Jan to Sep 2017 out-of-sample

Given the same budget

Extra drivers

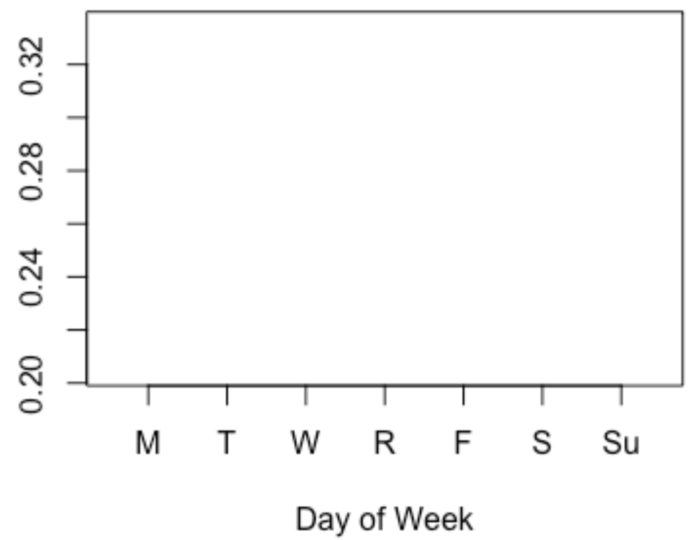


Can recruit **17% more drivers**

Average promo: 1.61x

Given the same capacity

Cost saved

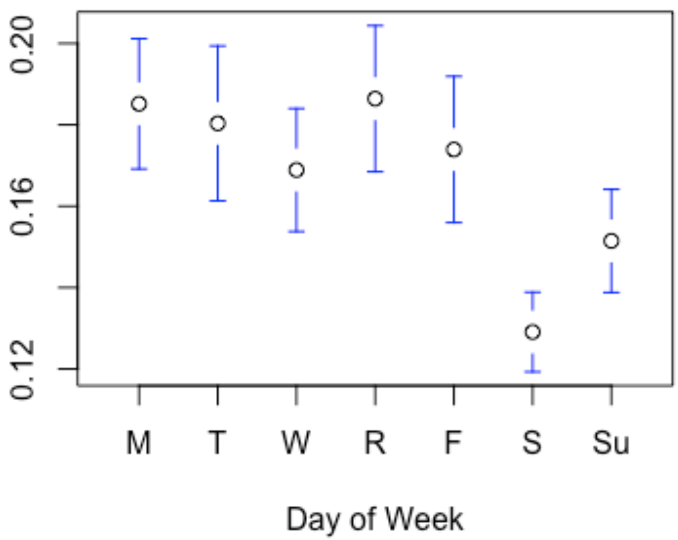


Reallocating Incentives

Compared to current practice from Jan to Sep 2017 out-of-sample

Given the same budget

Extra drivers

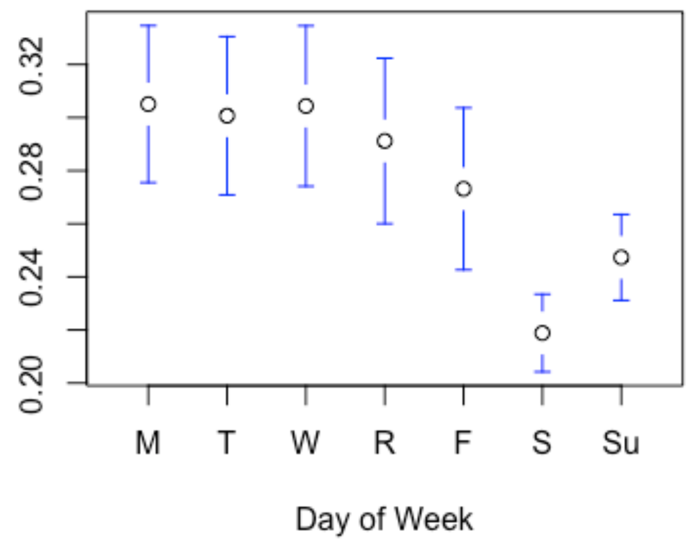


Can recruit **17% more drivers**

Average promo: 1.61x

Given the same capacity

Cost saved



Costs 28% less to maintain capacity

Summary

How do gig economy workers make labor decisions?

Approach

- Shift-level data from ride-hailing company
- Modified Heckman estimation w/ IVs and fixed effects

Findings

As day/week proceeds...

Neoclassical	Financial Incentive	encourages working
Behavioral	Income Target	discourages working later on
New phenomenon	Inertia	encourages working

Implications

- Compared to current practice, our approach can improve service capacity without incurring extra cost or maintain the same capacity at a lower cost

Gad Allon, Maxime Cohen, **Park Sinchaisri** (swich@wharton.upenn.edu | @wichinpong)

Appendix

Results Within Day/Effect Size

1

% Change in P(Work)

+$10 Offer	+$10 ISF	+1h HSF

2

Change in Minutes Worked

+$10 Earn	+$10 ISF	+1h HSF

Results Within Day/Effect Size

SUV	1: % Change in P(Work)			2: Change in Minutes Worked		
	+$10 Offer	+$10 ISF	+1h HSF	+$10 Earn	+$10 ISF	+1h HSF
Midday	0.82	-5.73	57.21	0.51	-4.87	109.53
PM-Peak	-4.39	-0.57	15.27	13.64	-0.24	18.96
PM-OPeak	0.27	-0.36	6.43	1.72	-0.08	1.18
Late Night	0.34	-0.22	3.32	14.87	-0.11	1.35

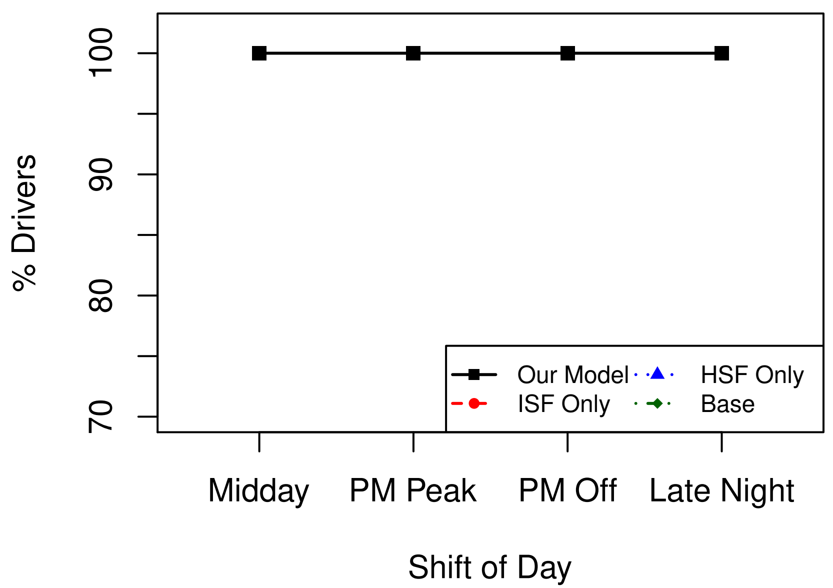
Results Within Day/Effect Size

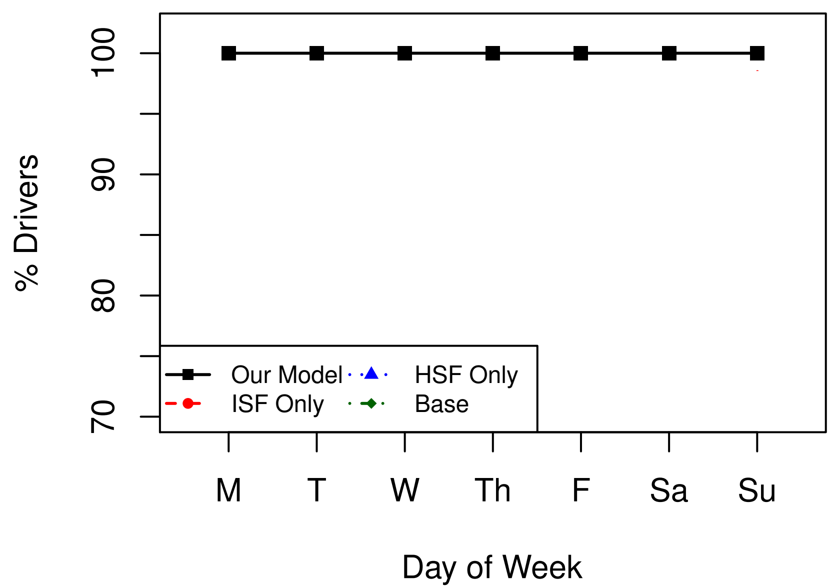
SUV	1 % Change in P(Work)			2 Change in Minutes Worked		
	+$10 Offer	+$10 ISF	+1h HSF	+$10 Earn	+$10 ISF	+1h HSF
Midday	0.82	-5.73	57.21	0.51	-4.87	109.53
PM-Peak	-4.39	-0.57	15.27	13.64	-0.24	18.96
PM-OPeak	0.27	-0.36	6.43	1.72	-0.08	1.18
Late Night	0.34	-0.22	3.32	14.87	-0.11	1.35

For an average SUV driver during Midday,
+$10 hourly wage → P(work) increases by **1%** / extends duration by **30s**
+$10 cumulative income → P(work) drops **6%** / shortens duration by **5m**
+1 cumulative hours → P(work) increases by **57%** / extends duration by **2h**

Ignoring Behavioral Factors

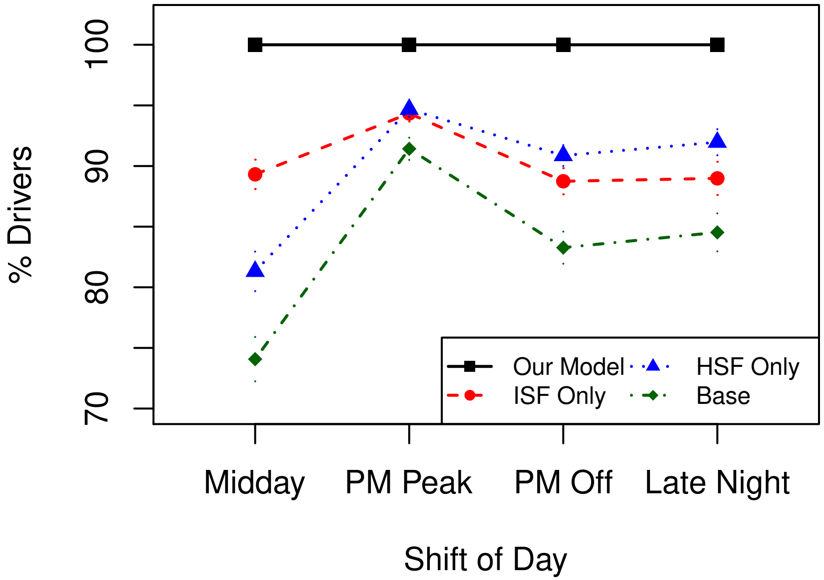
Assuming our model is correct, how many workers the firm would fail to attract if it did not incorporate income targeting and inertia?

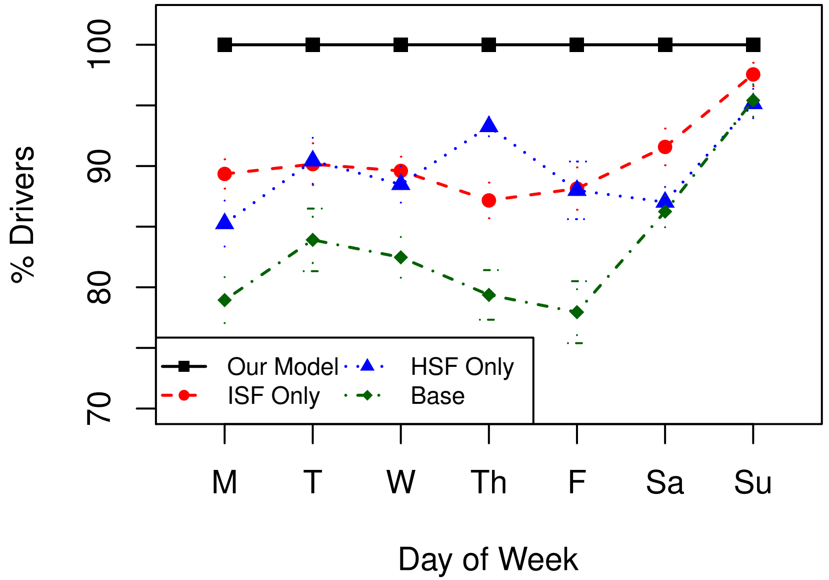




Ignoring Behavioral Factors

Assuming our model is correct, how many workers the firm would fail to attract if it did not incorporate income targeting and inertia?

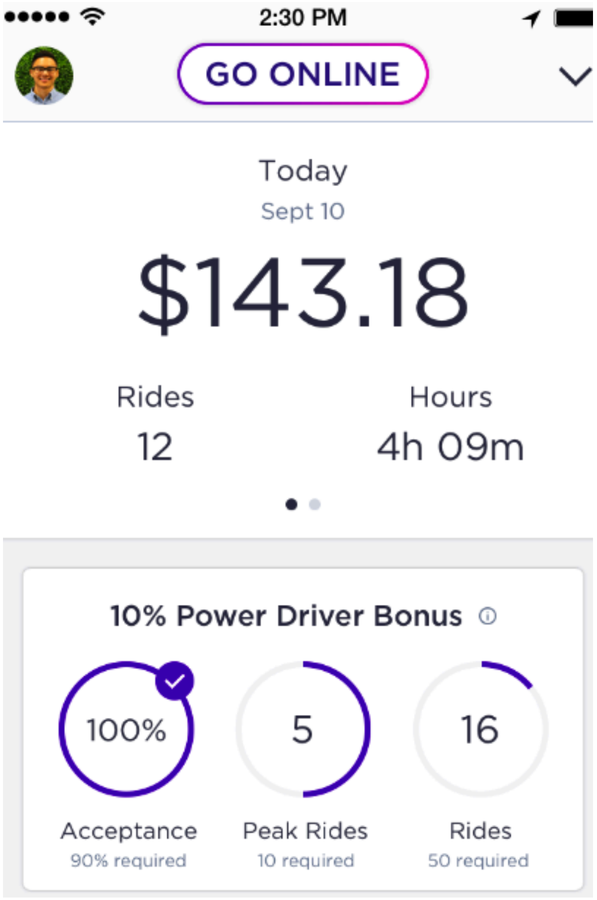


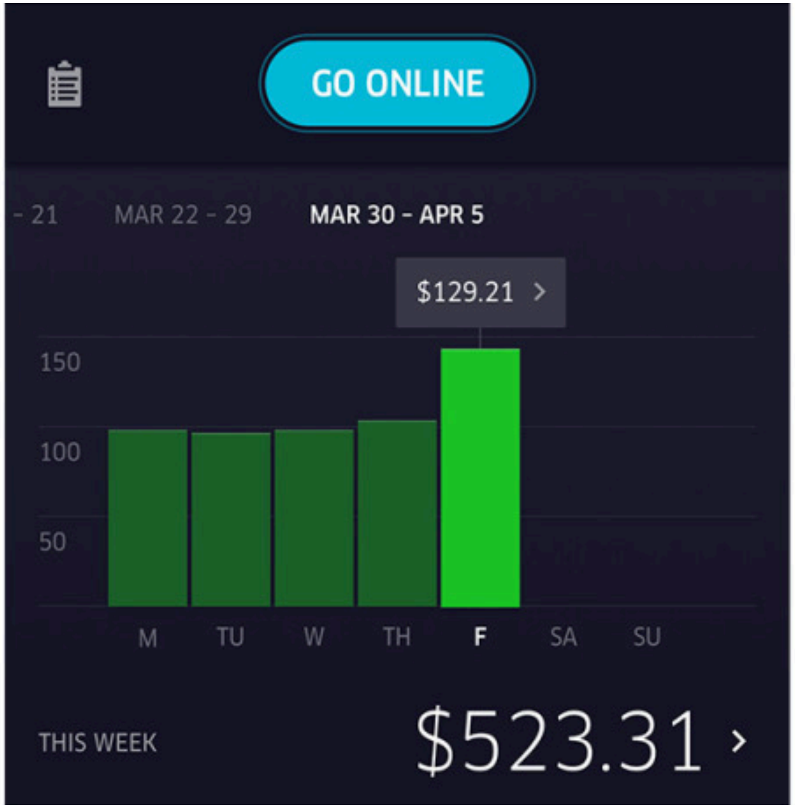


Average loss in capacity:

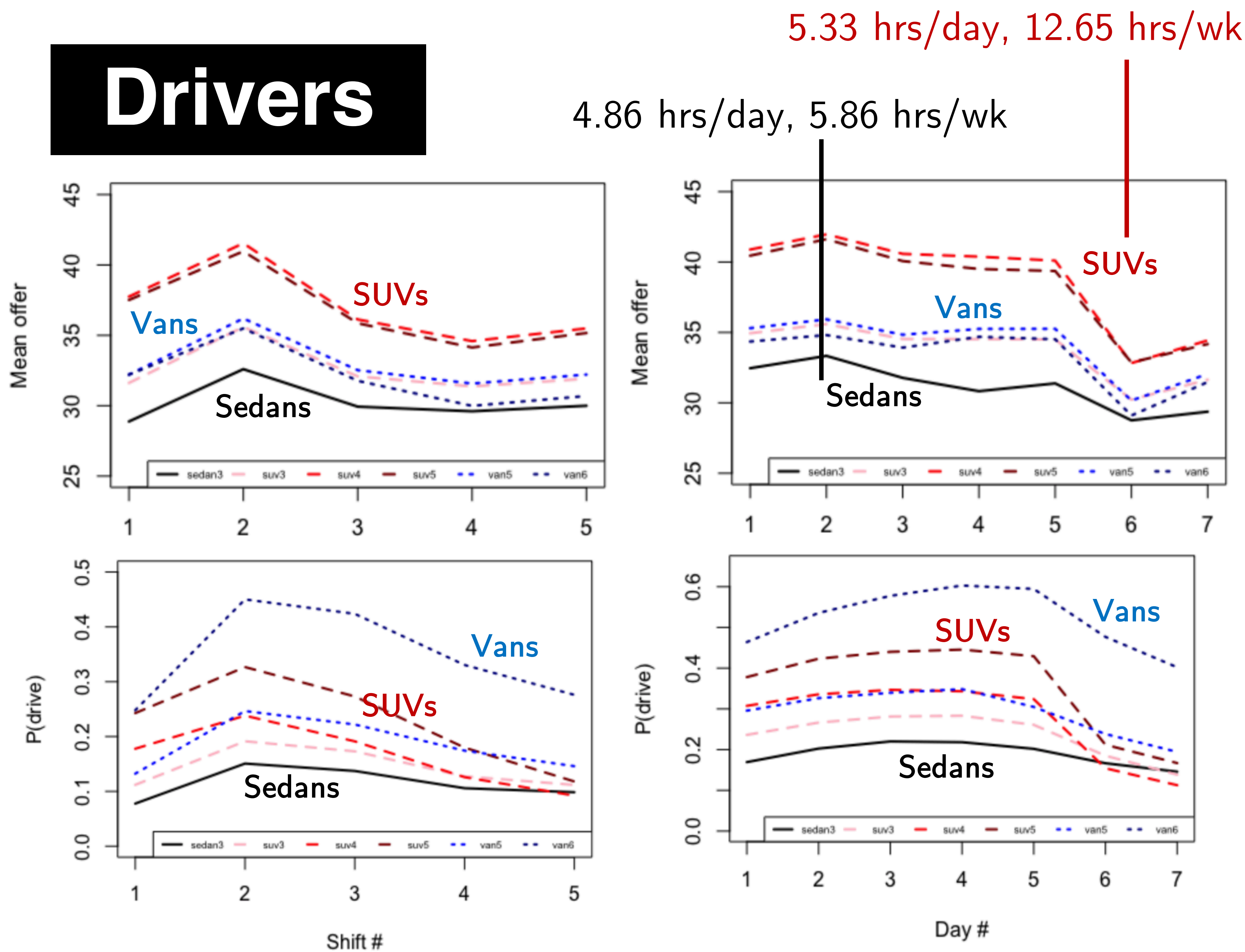
Ignoring: Income targeting 10.32%, inertia 9.63%, or both 16.70%

Driver's View





Drivers



5.33 hrs/day, 12.65 hrs/wk
4.86 hrs/day, 5.86 hrs/wk
Mean offer
P(drive)
SUVs
Vans
Sedans
sedan3
suv3
suv4
suv5
van5
van6
Shift #
Day #

Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Endogenous Variable	Instrument
Hourly offer	Average offers of “co-skippers”

Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Endogenous Variable	Instrument
Hourly offer	Average offers of "co-skippers"

Level of inactivity for drivers eligible to work today

last worked 6+ days ago	last worked 3-5 days ago	last worked 2 days ago	last worked yesterday

Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Endogenous Variable	Instrument
Hourly offer	Average offers of "co-skippers"

Level of inactivity for drivers eligible to work today

last worked 6+ days ago	last worked 3-5 days ago	last worked 2 days ago	last worked yesterday

Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Endogenous Variable	Instrument
Hourly offer	Average offers of “co-skippers”

Level of inactivity for drivers eligible to work today

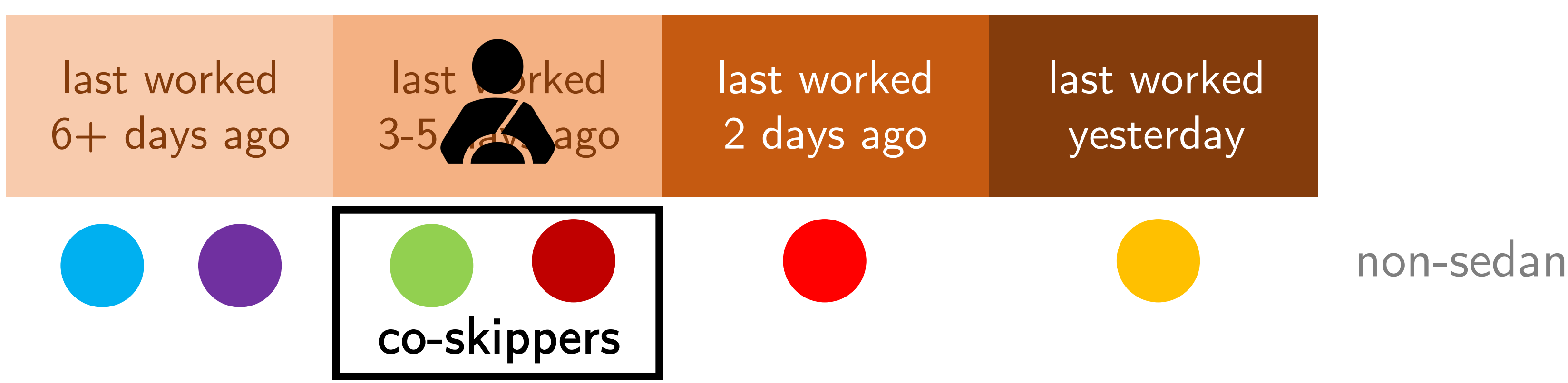
Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Endogenous Variable	Instrument
Hourly offer	Average offers of "co-skippers"

Level of inactivity for drivers eligible to work today



Gig Economy x Retail

Baker Retailing Center

- **Retail candidates value flexible shift patterns** and shorter work weeks over compensation and benefits.

- Benefits of adopting flexible workforce: Seasonality, resolving high turnover, matching consumer trend, high quality/fresh perspective

- Many have already adopted/worked closely with gig companies
  - Delivery business: Walmart x Uber/Lyft, GM x Lyft, Apple x Didi
  - Flexible staffing: IKEA x TaskRabbit, Samsung x Upwork

snag.work on-demand

Wonolo

coople

(allwork)

Heckman Sample Selection

Suppose that the pattern of missingness (I'll refer to this as censored hereafter) is related to the latent (unobserved) process

$$\mathbf{z}^* = \mathbf{w}\gamma + \mathbf{u}$$

From this process, the researcher can observe

$$\begin{aligned} z_i =& 1 \text{ if } z_i^* > 0 \\ =& 0 \text{ if } z_i^* \leq 0 \end{aligned}$$

or $z_i = 1$ (y_i not censored) when

$$u_i \geq -\mathbf{w}_i\gamma$$

The probability of y_i not censored is

$$\begin{aligned} Pr(u_i \geq -\mathbf{w}_i\gamma) =& 1 - \Phi(-\mathbf{w}_i\gamma) \\ =& \Phi(\mathbf{w}_i\gamma) \end{aligned}$$

if we are willing to assume that $\mathbf{u} \sim N(\mathbf{0}, \mathbf{I})$. Note for identification purposes in the Heckman Model we restrict $Var(u_i) = 1$. Also note that $1 - \Phi(-\mathbf{w}_i\gamma) = \Phi(\mathbf{w}_i\gamma)$ by symmetry of the standard normal distribution.

Heckman Sample Selection

Having constructed a model a model for censoring, we can construct "amounts" equation as follows. Denoting $\mathbf{y}$ as the not censored (observed) dependent variable, the censoring model defines what is in the estimation sample as

$$y_i = y_i^* = \mathbf{x}_i\beta + \epsilon_i \text{ observed, if } z_i = 1$$

Finally, the joint distribution of the errors in the selection (u_i) and amounts equation (ϵ) is distributed iid as

$$\begin{bmatrix} u_i \\ \epsilon_i \end{bmatrix} \sim Normal\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & \sigma_\epsilon^2 \end{bmatrix}\right)$$

To see how the selection and amounts model are related, consider

$$\begin{aligned} E(y_i|y_i \text{ observed}) = & E(y_i|z^* > 0) \\ = & E(y_i|u_i > -\mathbf{w}_i\gamma) \\ = & \mathbf{x}_i\beta + E(\epsilon_i|u_i > -\mathbf{w}_i\gamma) \\ = & \mathbf{x}_i\beta + \rho\sigma_\epsilon \frac{\phi(\mathbf{w}_i\gamma)}{\Phi(\mathbf{w}_i\gamma)} \end{aligned}$$

What is immediately apparent is that the conditional mean ($E(y_i|y_i \text{ observed})$) differs from the unconditional mean ($\mathbf{x}_i\beta$) only if $\rho \neq 0$ since all the other elements in the far right hand term (i.e., the variance of the error in the amounts equation, σ_ϵ, and the Inverse Mills Ratio, $\frac{\phi(\mathbf{w}_i\gamma)}{\Phi(\mathbf{w}_i\gamma)}$) in the preceding equation are strictly positive. So if the errors in the amounts and selection equations are uncorrelated ($\rho = 0$) we can safely apply ordinary least squares to uncover unbiased estimates for β and can ignore endogenous selection effects and the selection equation portion of the model.

Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Endogenous Variable	Instrument
Hourly offer	Average offers of "co-workers"

Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Endogenous Variable

Hourly offer

Instrument

Average offers of “co-workers”

Empirical Strategy Challenges

Simultaneity

Solution: Instrumental Variables

Endogenous Variable

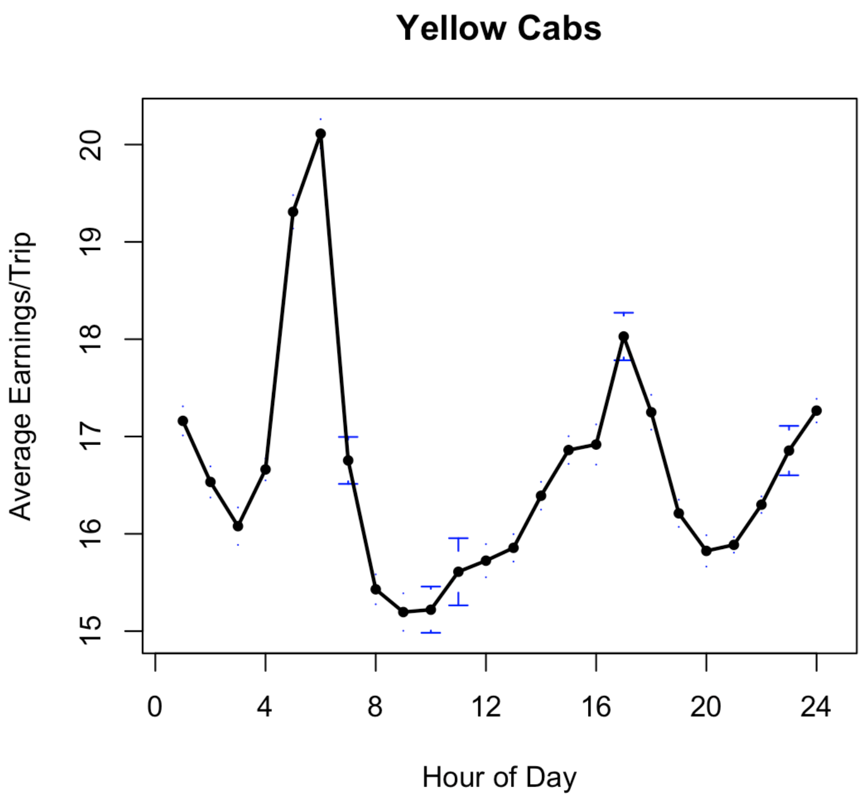
Hourly offer

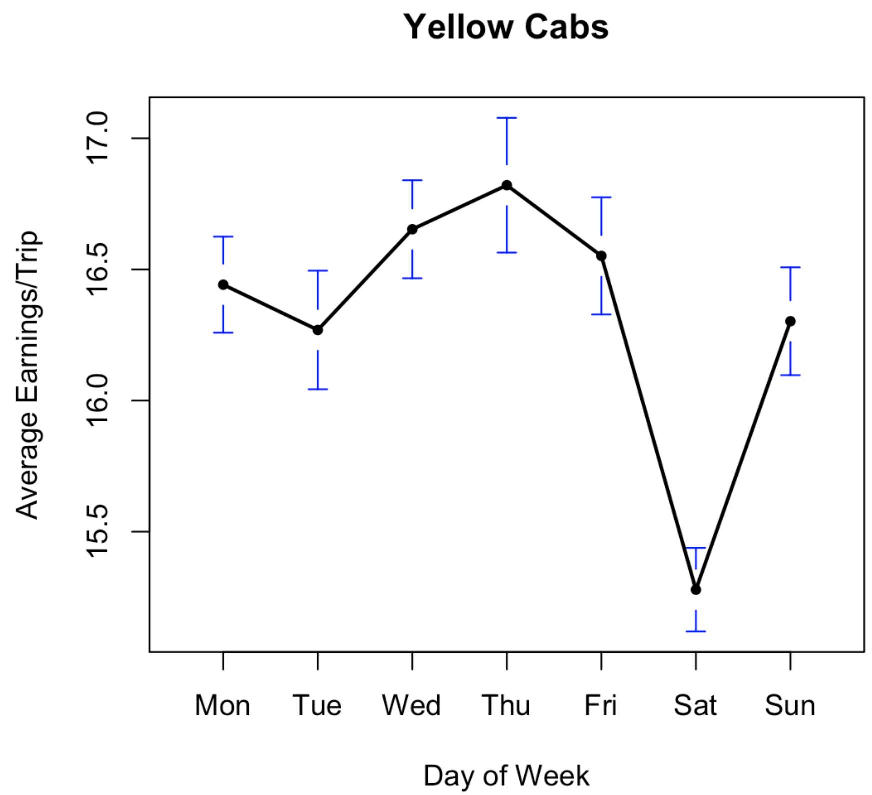
Instrument

Average offers of “co-workers”

TLC Data

Fares/earnings for all yellow cab trips





101M yellow cab trips