Detecting Routines in Ride-sharing: Implications For Customer Management

Ryan Dew, Eva Ascarza, Oded Netzer, Nachum Sicherman

Wharton UNIVERSITY of PENNSYLVANIA

HARVARD BUSINESS SCHOOL

Columbia Business School AT THE VERY CENTER OF BUSINESS™

Routine: behavior with a defined, recurring, temporal structure, such that the same behavior occurs at roughly the same time, period after period

Routine: behavior with a defined, recurring, temporal structure, such that the same behavior occurs at roughly the same time, period after period

Routine: behavior with a defined, recurring, temporal structure, such that the same behavior occurs at roughly the same time, period after period

Routine: behavior with a defined, recurring, temporal structure, such that the same behavior occurs at roughly the same time, period after period

Routine: behavior with a defined, recurring, temporal structure, such that the same behavior occurs at roughly the same time, period after period

Yet, there are **no existing models** for identifying routines from transaction data!

Why might routines matter?

1. Greater share of routine means an increase in *overall usage*

Why might routines matter?

Why might routines matter?

1. Greater share of routine means an increase in *overall usage*
2. Spillover effects: the more a company features in a customer's routine, the more it will be *top-of-mind* outside of the routine

Why might routines matter?

1. Greater share of routine means an increase in *overall usage*
2. Spillover effects: the more a company features in a customer’s routine, the more it will be *top-of-mind* outside of the routine
3. Segmentation and targeting based on *types of routines*

Why might routines matter?

1. Greater share of routine means an increase in *overall usage*
2. Spillover effects: the more a company features in a customer's routine, the more it will be *top-of-mind* outside of the routine
3. Segmentation and targeting based on *types of routines*
4. Long-run customer value: routine users have a *higher lifetime value*

Why might routines matter?

1. Greater share of routine means an increase in *overall usage*
2. Spillover effects: the more a company features in a customer's routine, the more it will be *top-of-mind* outside of the routine
3. Segmentation and targeting based on *types of routines*
4. Long-run customer value: routine users have a *higher lifetime value*

This Project

1. Develop a novel, all-purpose model for identifying individual-level routines
2. Apply our model to a unique ride-sharing data set
3. Show that customers with a high level of routine usage **churn less**, and **spend more** in the long run
4. Explore how temporal routineness predicts and moderates other important customer outcomes, over and above: mere habit, routines in terms of “what,” and other regularity metrics

This Project

1. Develop a novel, all-purpose model for identifying individual-level routines
2. Apply our model to a unique ride-sharing data set
3. Show that customers with a high level of routine usage **churn less**, and **spend more** in the long run
4. Explore how temporal routineness predicts and moderates other important customer outcomes, over and above: mere habit, routines in terms of "what," and other regularity metrics

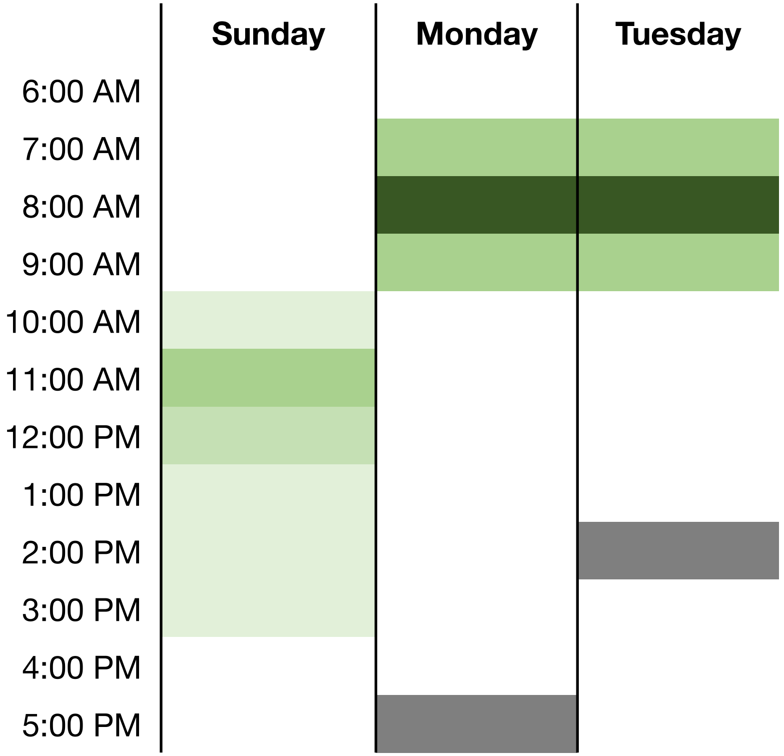
In short: we show the "shape" of customers' interactions matters!

Model

A Statistical Framework for Measuring Routineness

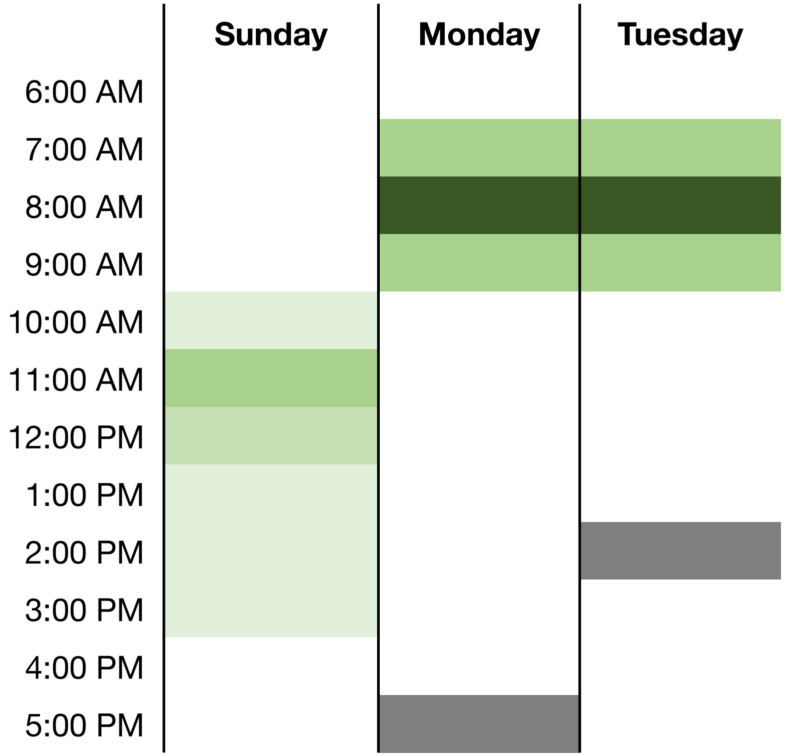
Model of Usage

From transaction data, we want to capture…



Model of Usage

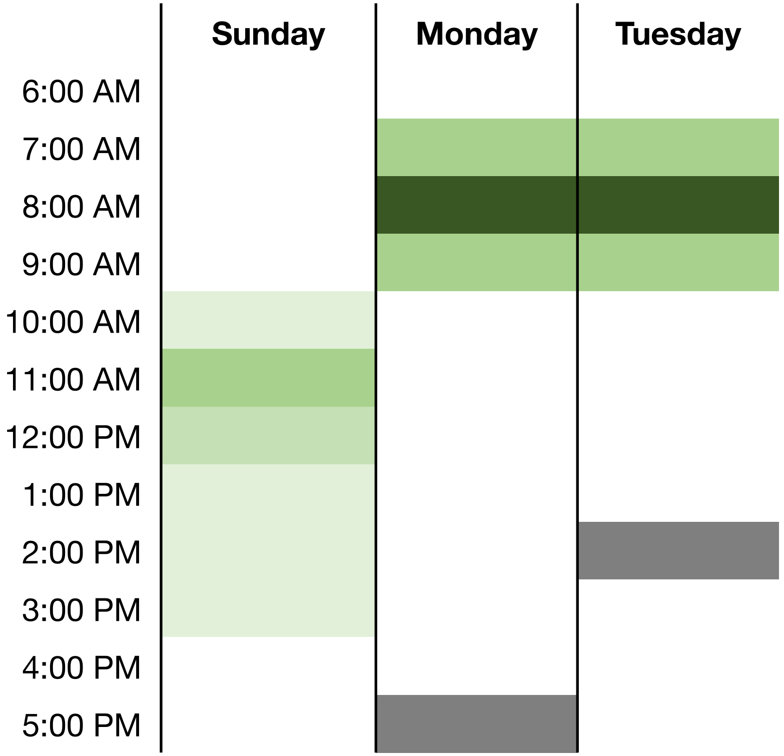
From transaction data, we want to capture…



Big question: How much of a person's activity is part of a routine?

Model of Usage

From transaction data, we want to capture…



Big question: How much of a person's activity is part of a routine?

Dependent variable: Usage (y)

Customer i, Week w, Day d, Hour h

Time $t = (w, d, h)$, Day-hour $j = (d, h)$

$$y_{it} \sim \mathrm{Poisson}(\lambda_{it})$$

$$\lambda_{it} = \underbrace{\exp(\alpha_{iw} + \mu_j)}_{\text{Random usage}} + \underbrace{\exp(\gamma_{iw} + \eta_{ij})}_{\text{Routine usage}}$$

- α_{iw} and γ_{iw} – Individual- and week-specific scaling terms
- μ_j – Common day-hour rate
- η_{ij} – Individual-specific day-hour rate

These random and routine usage terms give us a **structured decomposition** of overall usage.

Model-based Decomposition

$$y_{it} \sim \text{Poisson}(\lambda_{it})$$

$$\lambda_{it} = \underbrace{\exp(\alpha_{iw} + \mu_j)}_{\text{Random usage}} + \underbrace{\exp(\gamma_{iw} + \eta_{ij})}_{\text{Routine usage}}$$

Model-based Decomposition

$$y_{it} \sim \text{Poisson}(\lambda_{it})$$

$$\lambda_{it} = \underbrace{\exp(\alpha_{iw} + \mu_j)}_{\text{Random usage}} + \underbrace{\exp(\gamma_{iw} + \eta_{ij})}_{\text{Routine usage}}$$

$\updownarrow$

Requests follow an individual-level

inhomogeneous Poisson point process

Model-based Decomposition

$$y_{it} \sim \text{Poisson}(\lambda_{it})$$

$$\lambda_{it} = \underbrace{\exp(\alpha_{iw} + \mu_j)}_{\text{Random usage}} + \underbrace{\exp(\gamma_{iw} + \eta_{ij})}_{\text{Routine usage}}$$

$\updownarrow$

Requests follow an individual-level
inhomogeneous Poisson point process

Superposition of point processes:
If $Y_A \sim \text{PP}(\lambda_A)$ and $Y_B \sim \text{PP}(\lambda_B)$,
then $Y_A + Y_B \sim \text{PP}(\lambda_A + \lambda_B)$.

Random requests:

$$E_{iw}^{\text{Random}} = \sum_j \exp(\alpha_{iw} + \mu_j)$$

Routine requests:

$$E_{iw}^{\text{Routine}} = \sum_j \exp(\gamma_{iw} + \eta_{ij})$$

Model-based Decomposition

$$y_{it} \sim \text{Poisson}(\lambda_{it})$$

$$\lambda_{it} = \underbrace{\exp(\alpha_{iw} + \mu_j)}_{\text{Random usage}} + \underbrace{\exp(\gamma_{iw} + \eta_{ij})}_{\text{Routine usage}}$$

↕

Requests follow an individual-level **inhomogeneous Poisson point process**

Superposition of point processes:
If $Y_A \sim \text{PP}(\lambda_A)$ and $Y_B \sim \text{PP}(\lambda_B)$,
then $Y_A + Y_B \sim \text{PP}(\lambda_A + \lambda_B)$.

Random requests:

$$E_{iw}^{\text{Random}} = \sum_j \exp(\alpha_{iw} + \mu_j)$$

Routine requests:

$$\boxed{E_{iw}^{\text{Routine}}} = \sum_j \exp(\gamma_{iw} + \eta_{ij})$$

“Routineness”

How do we specify those intensity functions?

$$\lambda_{it} = \exp(\alpha_{iw} + \mu_j) + \exp(\gamma_{iw} + \eta_{ij})$$

How do we specify those intensity functions?

$$\lambda_{it} = \exp(\alpha_{iw} + \mu_j) + \exp(\gamma_{iw} + \eta_{ij})$$

Scaling terms for weekly variation:

$$\alpha_{iw} \sim \mathcal{N}(\alpha_{iw-1}, \tau)$$
$$\gamma_i(w) \sim \mathcal{GP}(\gamma_0, k_{\mathrm{SE}}(w, w'; \phi_\gamma))$$

Rate terms for day-hour variation:

$$\mu(d, h) \sim \mathcal{GP}\left(0, k_{\mathrm{DH}}(d, h; \phi_\mu)\right)$$
$$\eta_i(d, h) \sim \mathcal{GP}\left(0, k_{\mathrm{DH}}(d, h; \phi_\eta)\right)$$

How do we specify those intensity functions?

$$\lambda_{it} = \exp(\alpha_{iw} + \mu_j) + \exp(\gamma_{iw} + \eta_{ij})$$

Scaling terms for weekly variation:

$$\alpha_{iw} \sim \mathcal{N}(\alpha_{iw-1}, \tau)$$

$$\boxed{\gamma_i(w) \sim \mathcal{GP}(\gamma_0, k_{\mathrm{SE}}(w, w'; \phi_\gamma))}$$

Gaussian process: a Bayesian nonparametric prior over a function space

Rate terms for day-hour variation:

$$\boxed{\begin{aligned} \mu(d,h) &\sim \mathcal{GP}\left(0, k_{\mathrm{DH}}(d,h;\phi_\mu)\right) \\ \eta_i(d,h) &\sim \mathcal{GP}\left(0, k_{\mathrm{DH}}(d,h;\phi_\eta)\right) \end{aligned}}$$

How do we specify those intensity functions?

$$\lambda_{it} = \exp(\alpha_{iw} + \mu_j) + \exp(\gamma_{iw} + \eta_{ij})$$

Scaling terms for weekly variation:

$$\alpha_{iw} \sim \mathcal{N}(\alpha_{iw-1}, \tau)$$

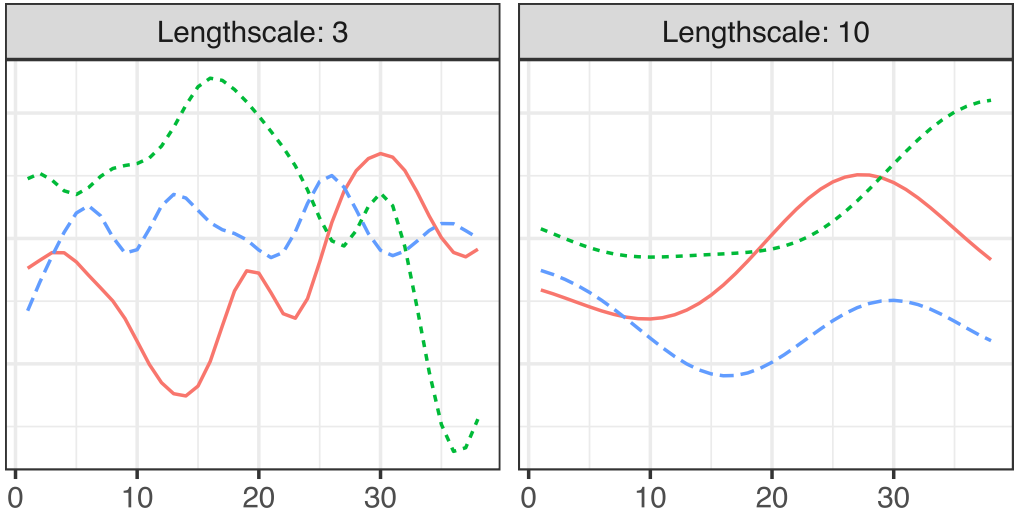
$$\gamma_i(w) \sim \mathcal{GP}(\gamma_0, \boxed{k_{\mathrm{SE}}(w, w'; \phi_\gamma)})$$

Squared exponential kernel with fixed smoothness

Rate terms for day-hour variation:

$$\mu(d,h) \sim \mathcal{GP}\left(0, k_{\mathrm{DH}}(d,h;\phi_\mu)\right)$$

$$\eta_i(d,h) \sim \mathcal{GP}\left(0, k_{\mathrm{DH}}(d,h;\phi_\eta)\right)$$



Key Assumption

The routine scale, $\gamma_i(w)$ evolves <u>slowly</u> over weeks

How do we specify those intensity functions?

$$\lambda_{it} = \exp(\alpha_{iw} + \mu_j) + \exp(\gamma_{iw} + \eta_{ij})$$

Scaling terms for weekly variation:

$$\alpha_{iw} \sim \mathcal{N}(\alpha_{iw-1}, \tau)$$

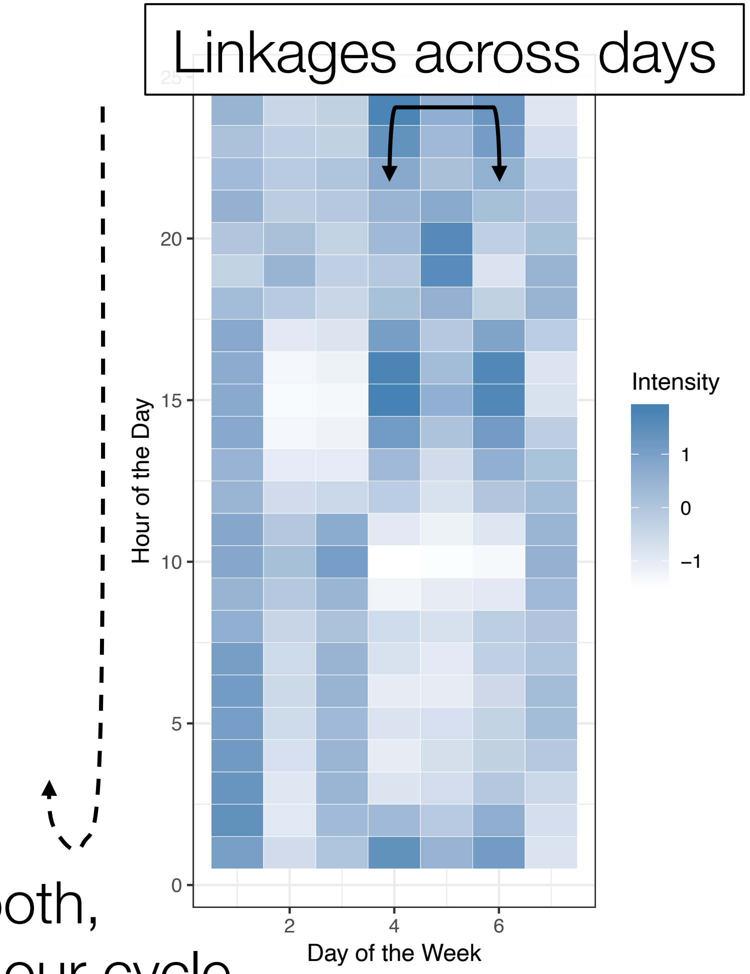
$$\gamma_i(w) \sim \mathcal{GP}(\gamma_0, k_{\text{SE}}(w, w'; \phi_\gamma))$$

Rate terms for day-hour variation:

$$\mu(d, h) \sim \mathcal{GP}\left(0, \boxed{k_{\text{DH}}(d, h; \phi_\mu)}\right)$$

$$\eta_i(d, h) \sim \mathcal{GP}\left(0, \boxed{k_{\text{DH}}(d, h; \phi_\eta)}\right)$$

Novel day-hour kernel



Smooth,
24-hour cycle

How do we specify those intensity functions?

$$\lambda_{it} = \exp(\alpha_{iw} + \mu_j) + \exp(\gamma_{iw} + \eta_{ij})$$

Scaling terms for weekly variation:

$$\alpha_{iw} \sim \mathcal{N}(\alpha_{iw-1}, \tau)$$
$$\gamma_i(w) \sim \mathcal{GP}(\gamma_0, k_{\mathrm{SE}}(w, w'; \phi_\gamma))$$

Rate terms for day-hour variation:

$$\mu(d, h) \sim \mathcal{GP}\left(0, k_{\mathrm{DH}}(d, h; \phi_\mu)\right)$$
$$\eta_i(d, h) \sim \mathcal{GP}\left(0, k_{\mathrm{DH}}(d, h; \phi_\eta)\right)$$

This structure basically buys us two things:

1. A decomposition of total usage into “random” and “routine”
2. An individual-level estimate of what that routine is (η_{ij})

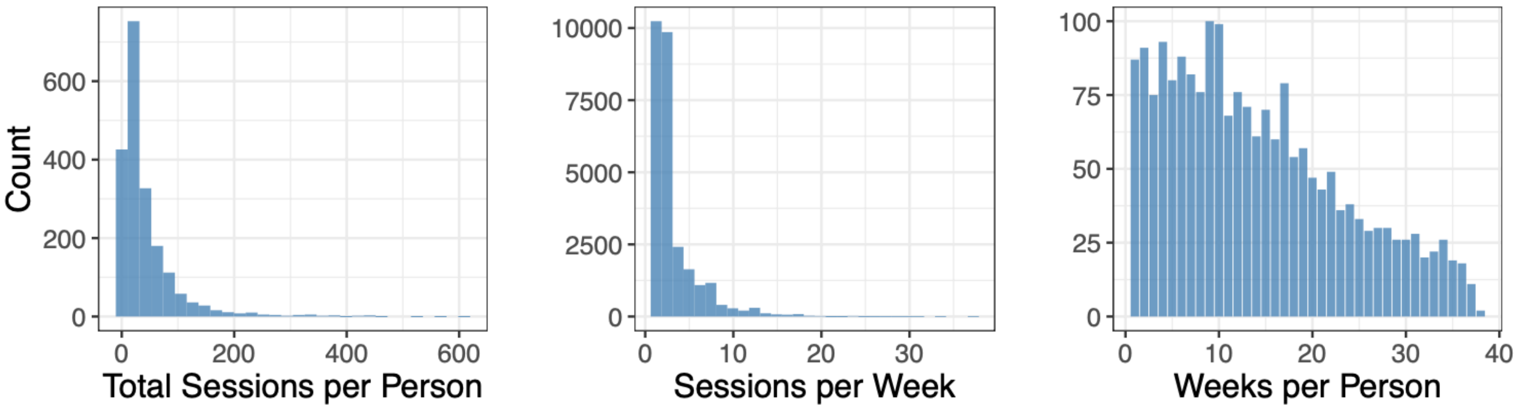
Results

Application to Ride-sharing Data

Ride-sharing Data

- Collaboration with a NYC-based ride-sharing company

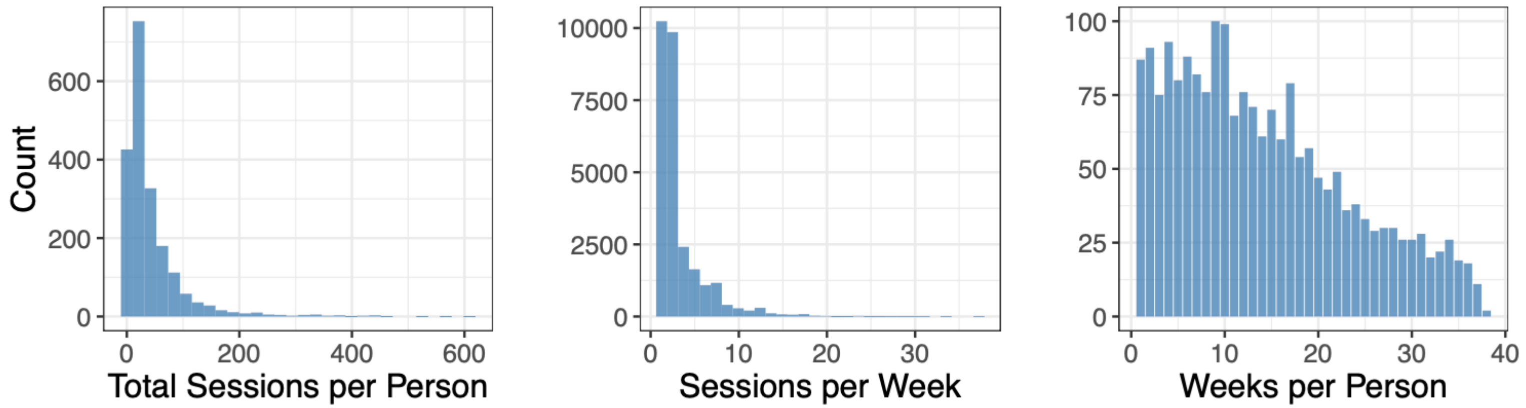
Total Customers	2,000
Total Weeks (Training)	38
Total Weeks (Holdout)	10
Number of Sessions	86,952
Sessions / Customer	43.48
Sessions / Customer / Week	3.10
Weeks in Data / Customer	14.02



Ride-sharing Data

- Collaboration with a NYC-based ride-sharing company

Total Customers	2,000
Total Weeks (Training)	38
Total Weeks (Holdout)	10
Number of Sessions	86,952
Sessions / Customer	43.48
Sessions / Customer / Week	3.10
Weeks in Data / Customer	14.02

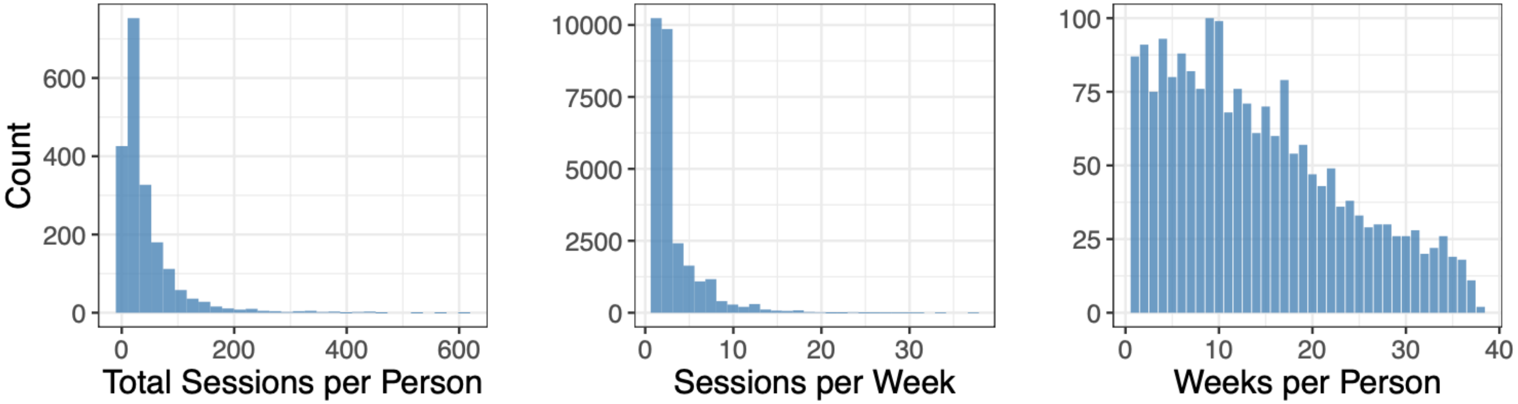


Ride-sharing Data

- Collaboration with a NYC-based ride-sharing company

Total Customers	2,000
Total Weeks (Training)	38
Total Weeks (Holdout)	10
Number of Sessions	86,952
Sessions / Customer	43.48
Sessions / Customer / Week	3.10
Weeks in Data / Customer	14.02

Basic unit of analysis:
a "session"



“Quasi-simulation”

- 500 real customers + 15 fake customers with specific usage patterns

“Quasi-simulation”

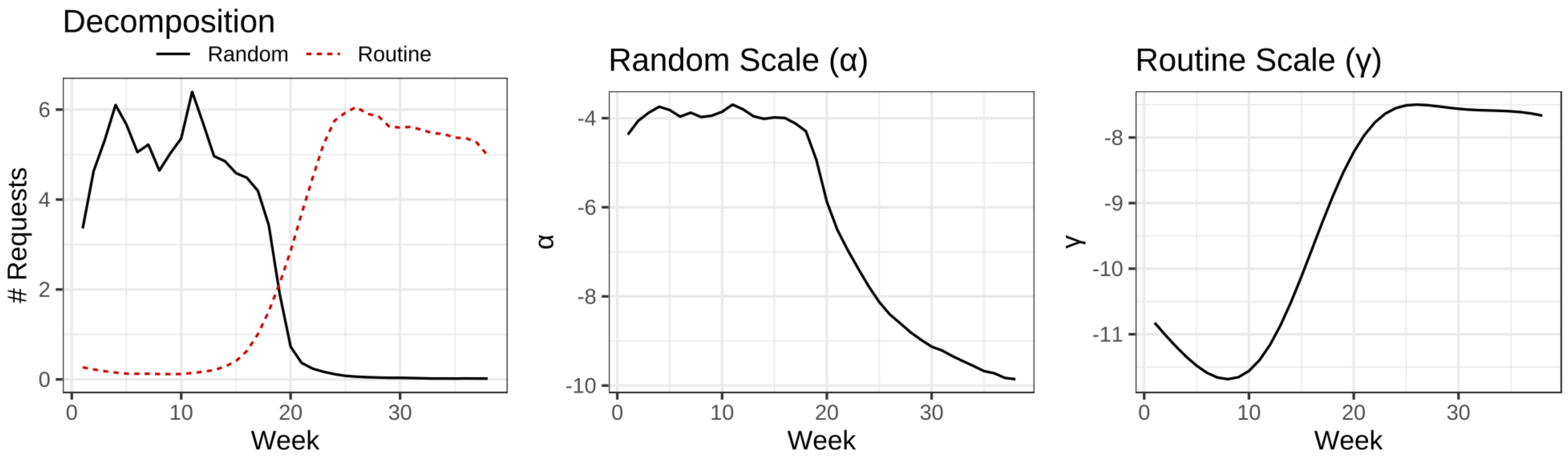
- 500 real customers + 15 fake customers with specific usage patterns

Case 1: Random then Routine

“Quasi-simulation”

- 500 real customers + 15 fake customers with specific usage patterns

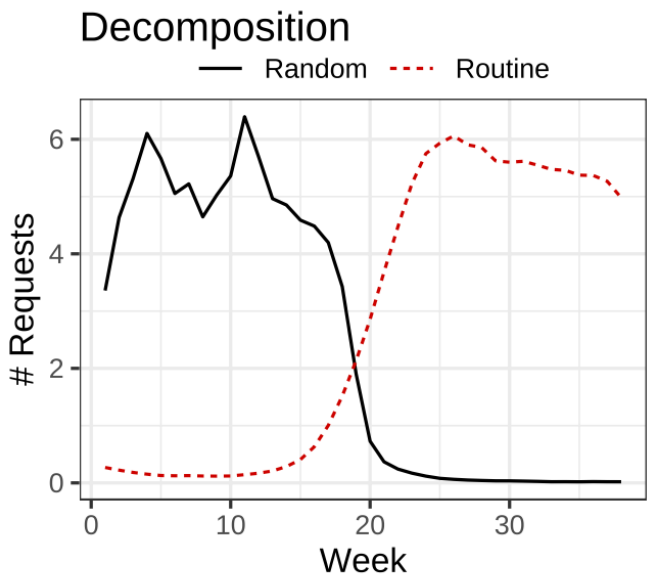
Case 1: Random then Routine

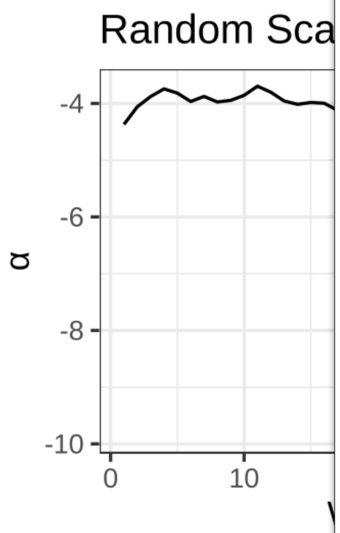


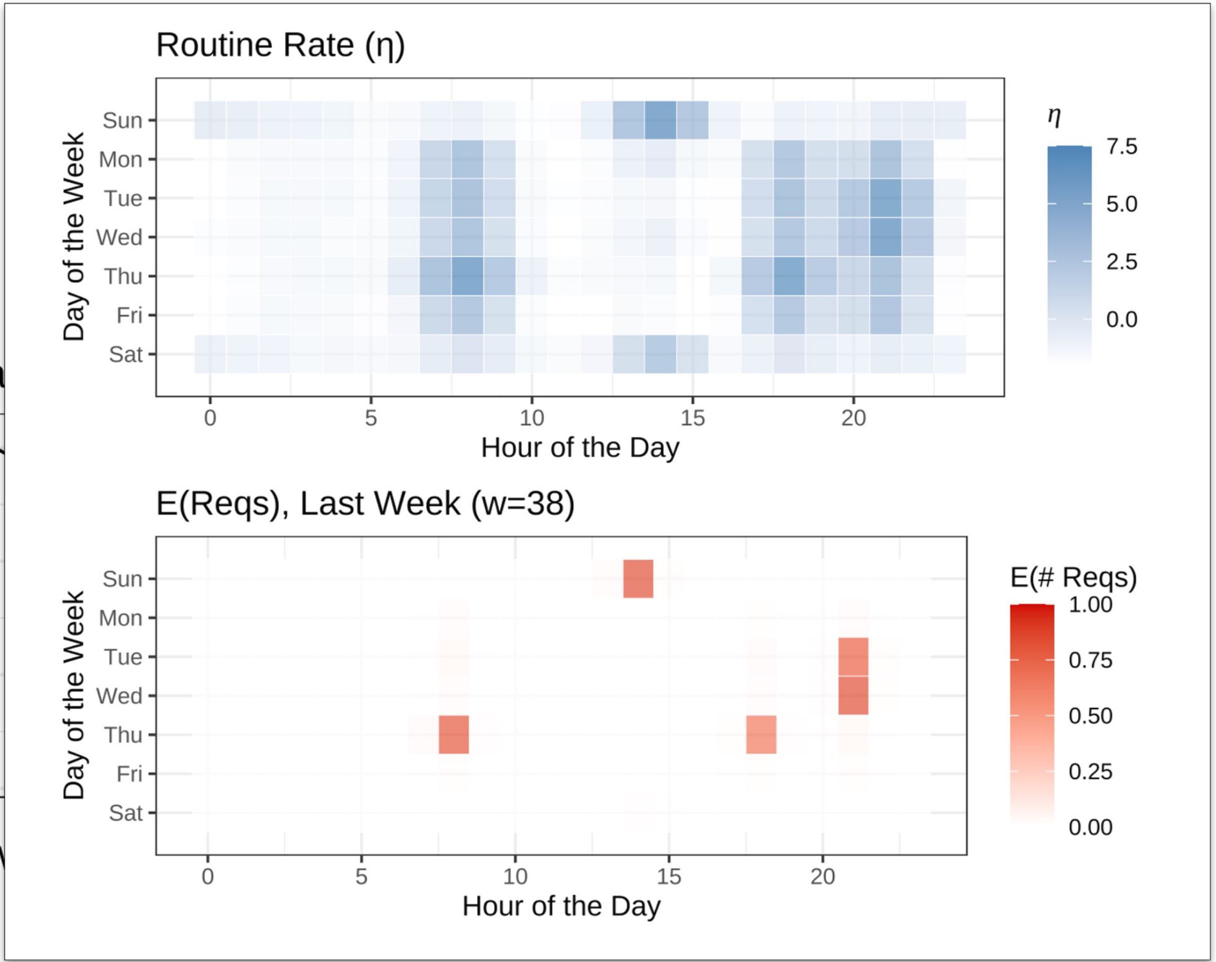
“Quasi-simulation”

- 500 real customers + 15 fake customers with specific usage patterns

Case 1: Random then Routine







“Quasi-simulation”

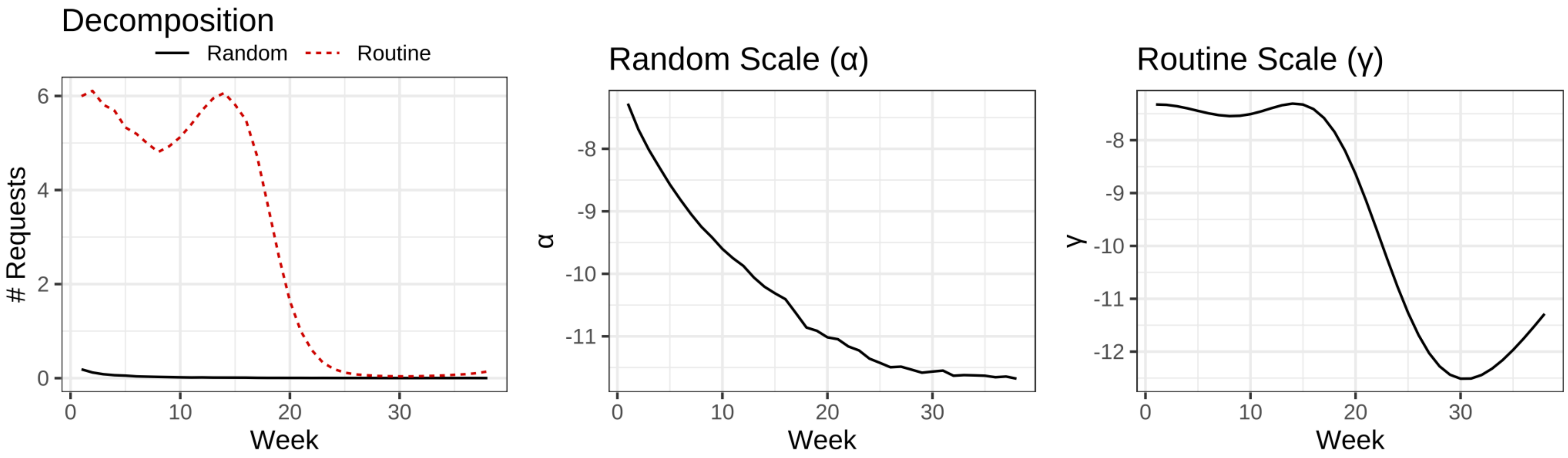
- 500 real customers + 15 fake customers with specific usage patterns

Case 2: Routine then Churn

“Quasi-simulation”

- 500 real customers + 15 fake customers with specific usage patterns

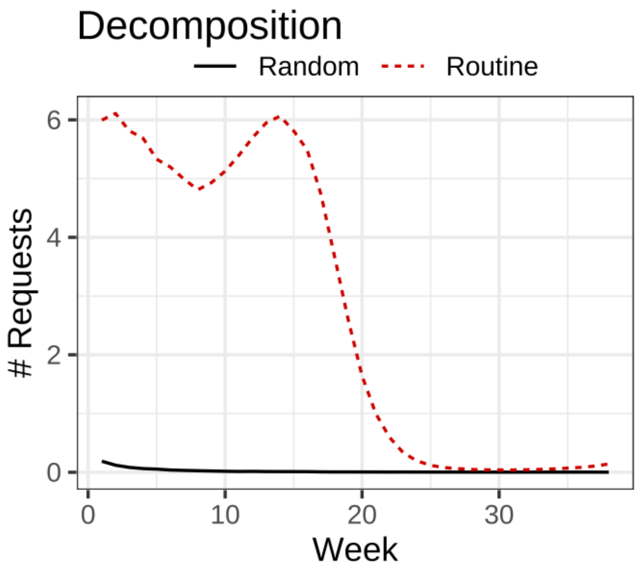
Case 2: Routine then Churn

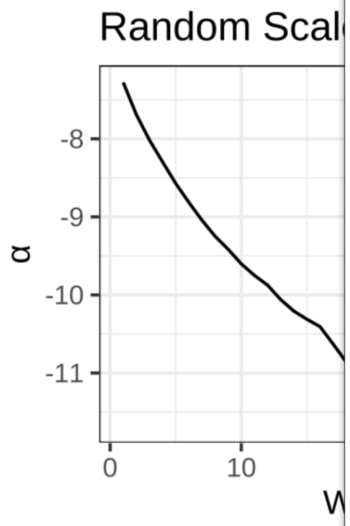


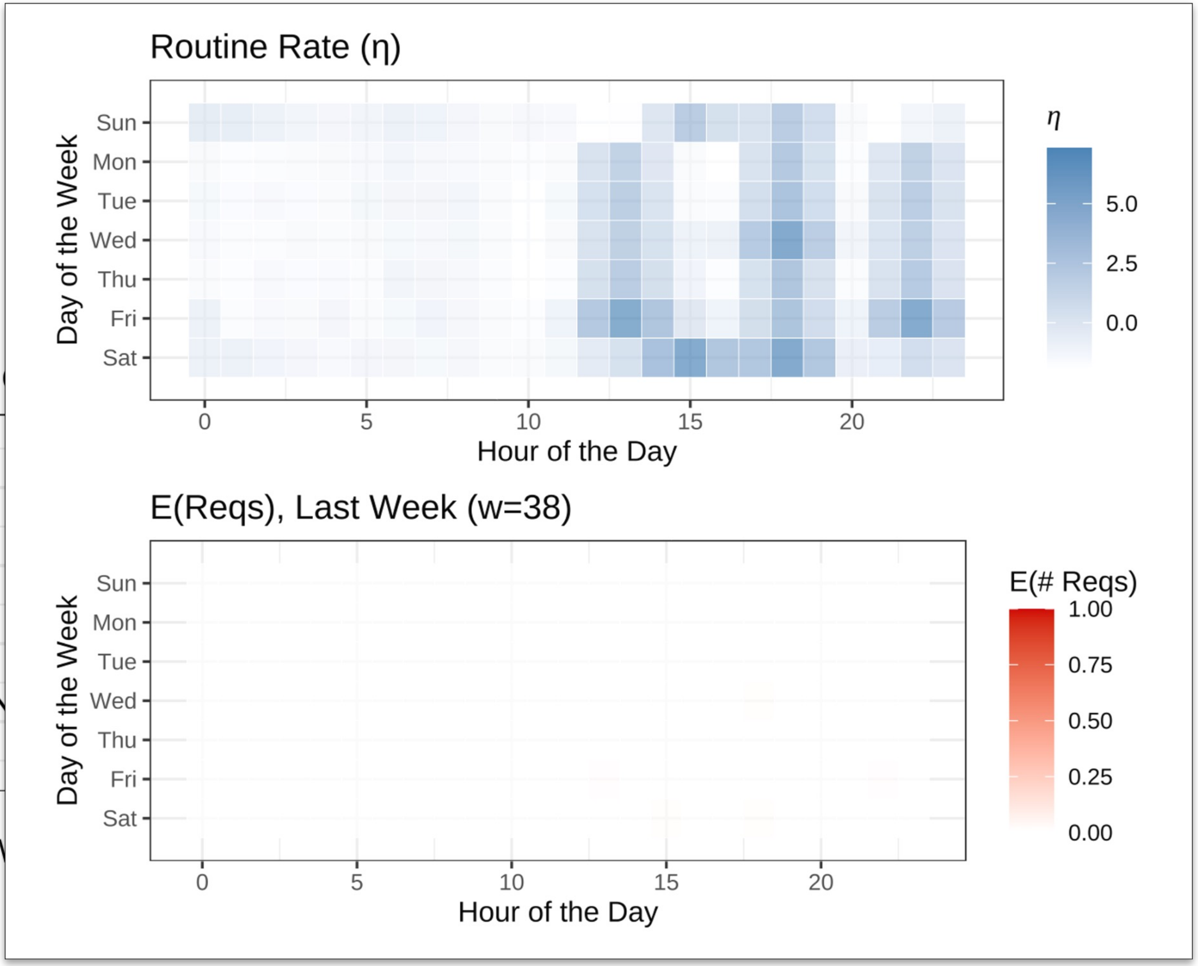
“Quasi-simulation”

- 500 real customers + 15 fake customers with specific usage patterns

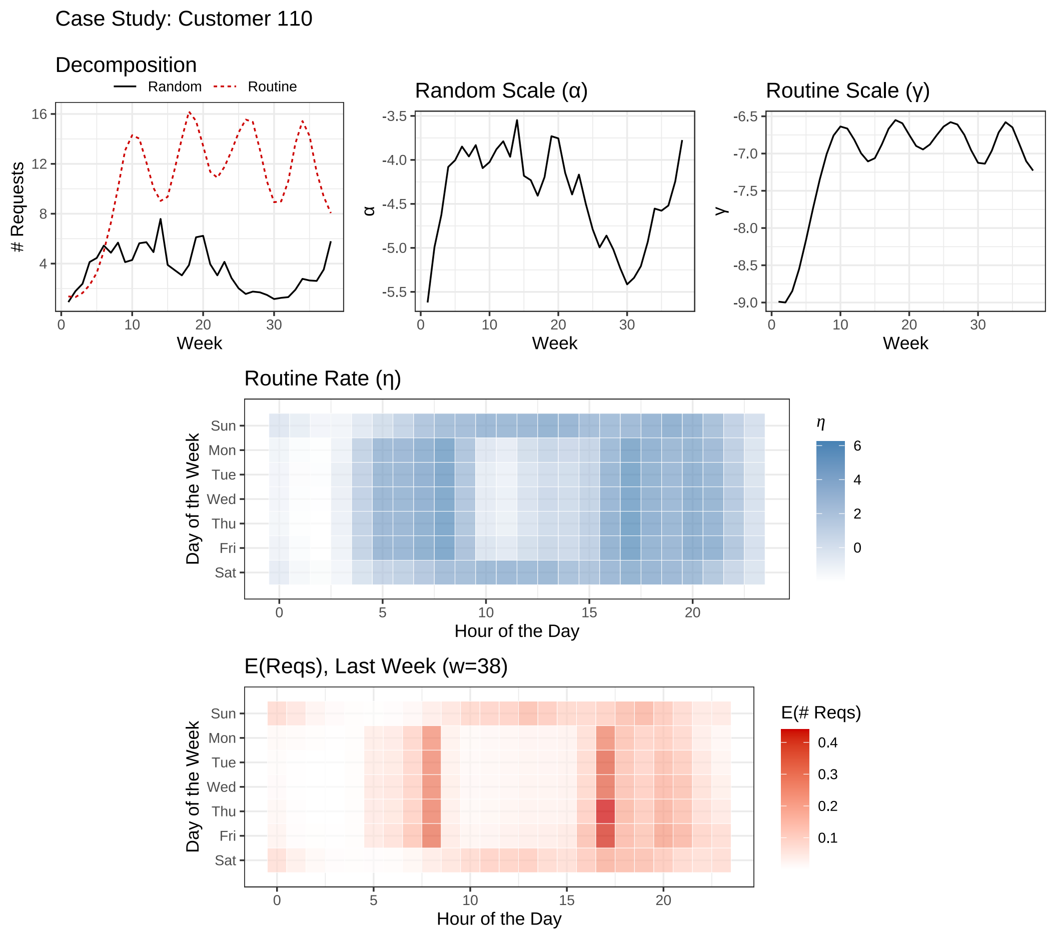
Case 2: Routine then Churn



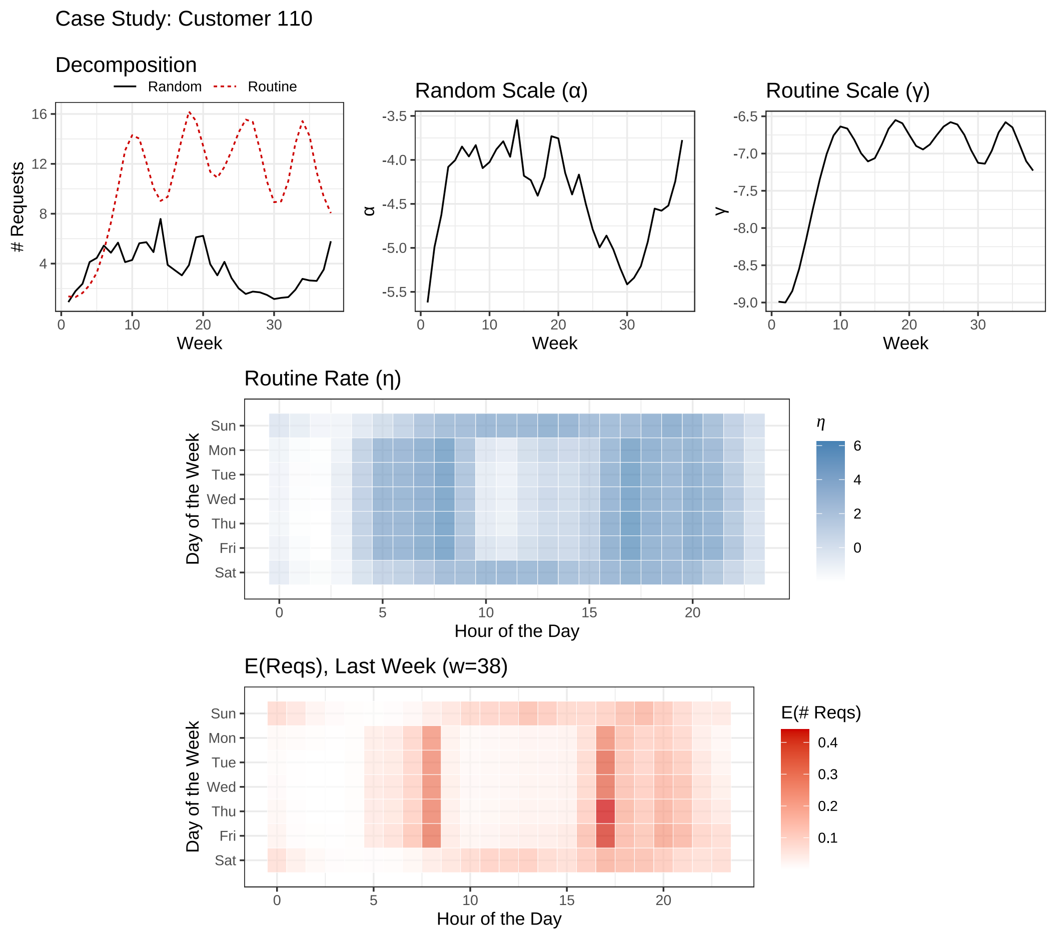


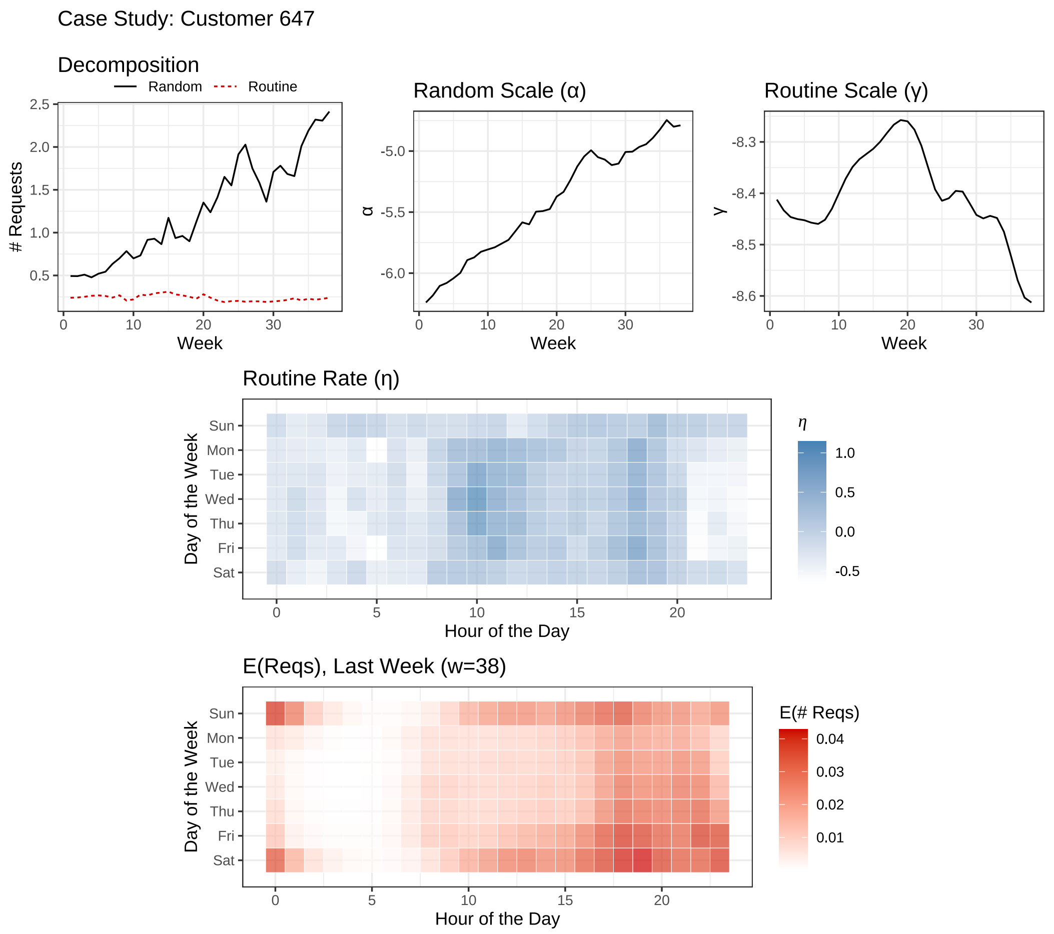


Real Case Studies

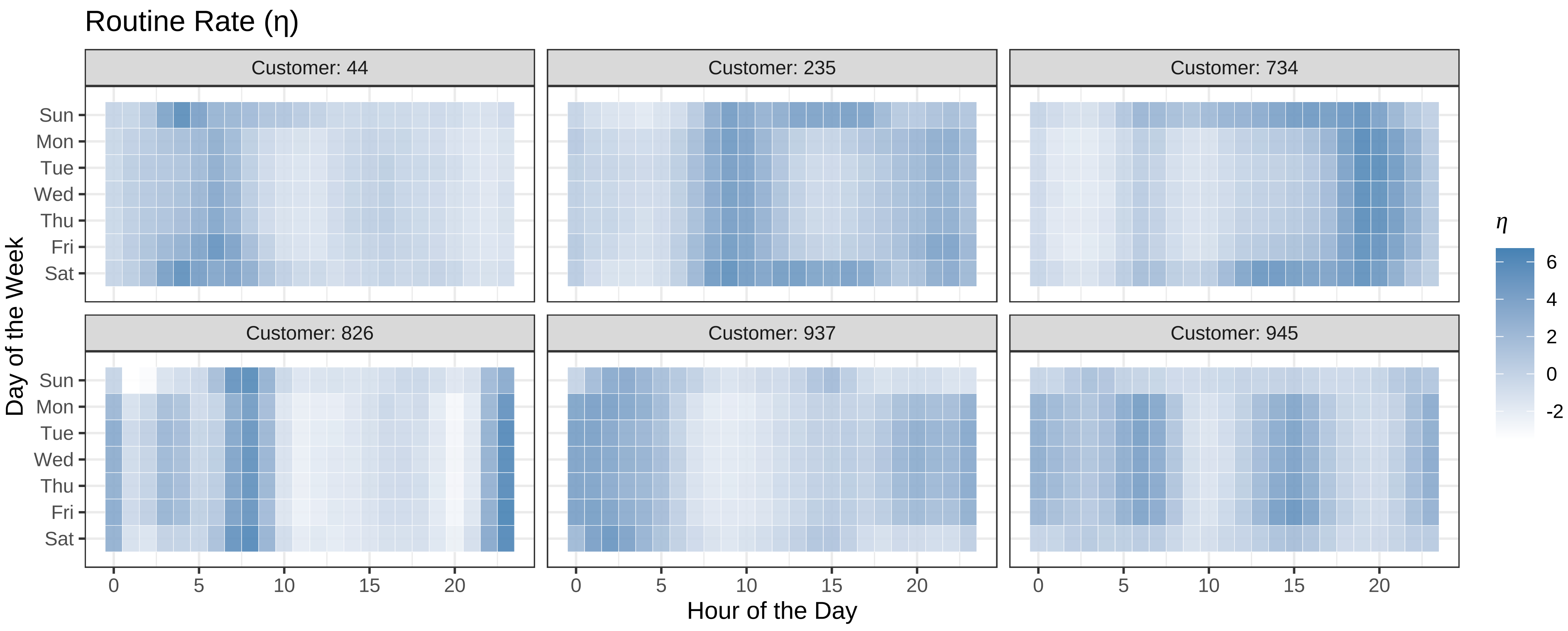


Real Case Studies

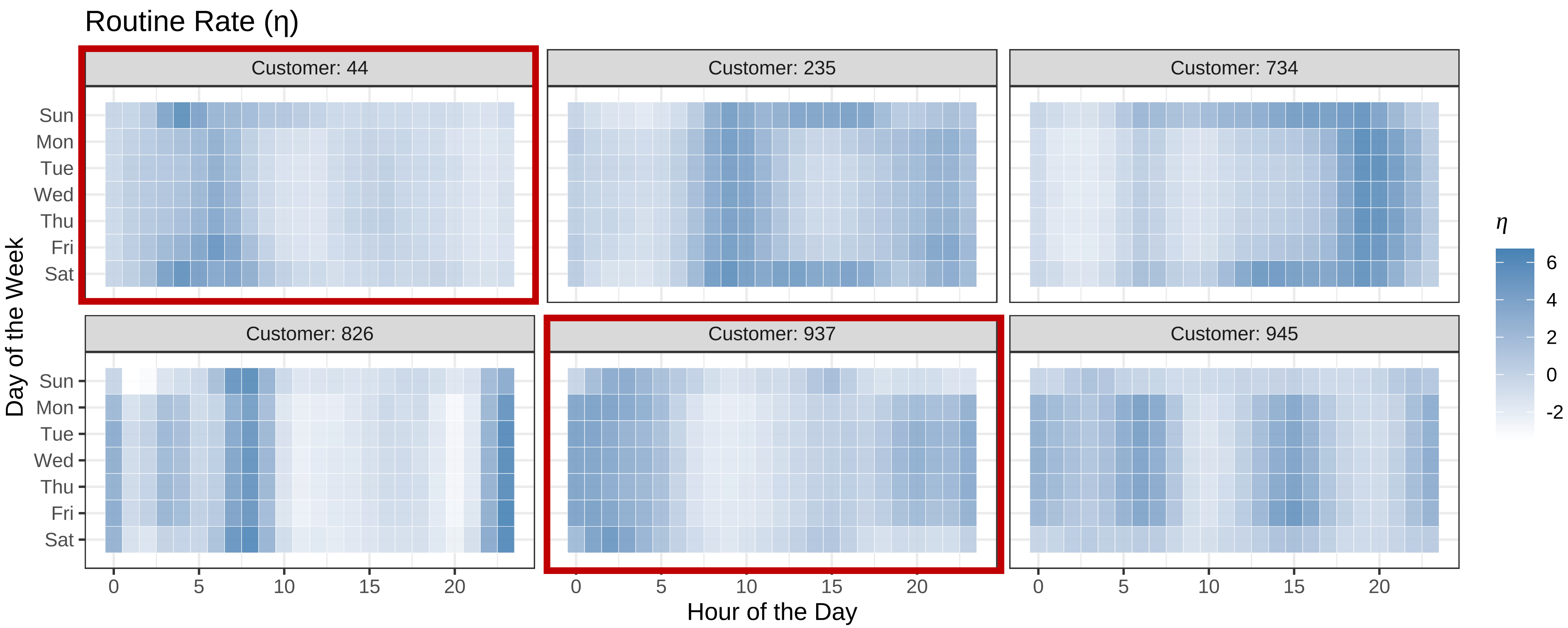




Many types of routines



Many types of routines



“Night owls”

Many types of routines

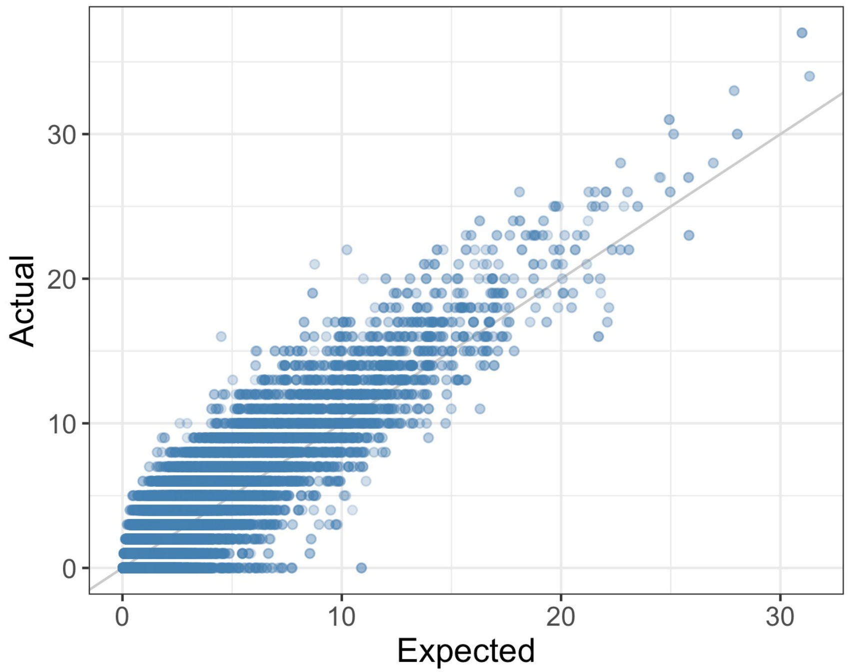
“Morning and night”

Validation

Can we trust these results?

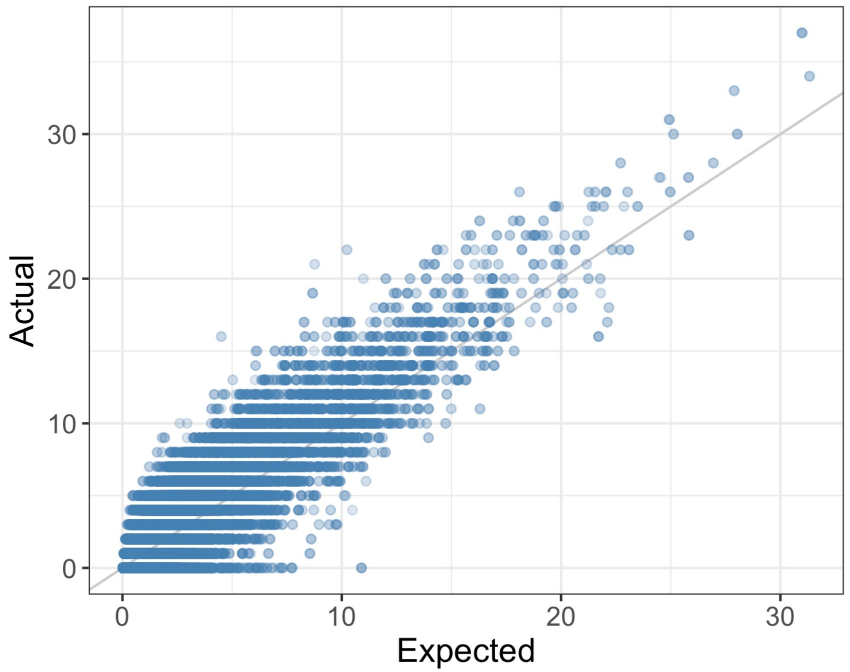
Validation

Can we trust these results?



Validation

Can we trust these results?



Prediction of transaction times:

- Predict timing of <u>routine</u> rides
- 100+% improvement over "universal" routine
- Also outperforms sophisticated LSTM-based benchmarks

Implications

Why should we care about routines?

Are routine customers more valuable?

	Dependent variable:			
	# Sessions		Activity	
	OLS		*Logistic*	
	Full Holdout	Last 5 Weeks	Full Holdout	Last 5 Weeks
	(1)	(2)	(3)	(4)
Requests ($w = 38$)	2.224^{***}	0.597^{***}	0.383^{***}	0.180^{**}
	(0.223)	(0.141)	(0.108)	(0.057)
Recency	-0.189^{***}	-0.106^{***}	-0.140^{***}	-0.124^{***}
	(0.042)	(0.026)	(0.010)	(0.010)
Frequency	0.095^{***}	0.049^{***}	-0.00002	0.004^{*}
	(0.007)	(0.004)	(0.002)	(0.002)
Routine ($w = 38$)	5.750^{***}	3.284^{***}	1.110^{**}	0.307^{*}
	(0.436)	(0.275)	(0.385)	(0.147)
Observations	2,000	2,000	2,000	2,000
R^2	0.567	0.425		
Note:			$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$	
			Intercept omitted for clarity.	

Are routine customers more valuable?

	Dependent variable:			
	# Sessions		Activity	
	OLS		*Logistic*	
	Full Holdout	Last 5 Weeks	Full Holdout	Last 5 Weeks
	(1)	(2)	(3)	(4)
Requests ($w = 38$)	2.224^{***}	0.597^{***}	0.383^{***}	0.180^{**}
	(0.223)	(0.141)	(0.108)	(0.057)
Recency	-0.189^{***}	-0.106^{***}	-0.140^{***}	-0.124^{***}
	(0.042)	(0.026)	(0.010)	(0.010)
Frequency	0.095^{***}	0.049^{***}	-0.00002	0.004^{*}
	(0.007)	(0.004)	(0.002)	(0.002)
Routine ($w = 38$)	5.750^{***}	3.284^{***}	1.110^{**}	0.307^{*}
	(0.436)	(0.275)	(0.385)	(0.147)
Observations	2,000	2,000	2,000	2,000
R^2	0.567	0.425		
Note:				$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$ Intercept omitted for clarity.

Model-based Routineness

(Week 38)

Are routine customers more valuable?

Total Requests
(Week 38)

Model-based Routineness
(Week 38)

	Dependent variable:			
	# Sessions		Activity	
	OLS		*Logistic*	
	Full Holdout	Last 5 Weeks	Full Holdout	Last 5 Weeks
	(1)	(2)	(3)	(4)
Requests ($w = 38$)	2.224***	0.597***	0.383***	0.180**
	(0.223)	(0.141)	(0.108)	(0.057)
Recency	−0.189***	−0.106***	−0.140***	−0.124***
	(0.042)	(0.026)	(0.010)	(0.010)
Frequency	0.095***	0.049***	−0.00002	0.004*
	(0.007)	(0.004)	(0.002)	(0.002)
Routine ($w = 38$)	5.750***	3.284***	1.110**	0.307*
	(0.436)	(0.275)	(0.385)	(0.147)
Observations	2,000	2,000	2,000	2,000
R^2	0.567	0.425		

Note: *p<0.05; **p<0.01; ***p<0.001
Intercept omitted for clarity.

Are routine customers more valuable?

Total Requests (Week 38)

RF(M) Controls (Week 38)

Model-based Routineness (Week 38)

	Dependent variable:			
	# Sessions		Activity	
	OLS		*Logistic*	
	Full Holdout	Last 5 Weeks	Full Holdout	Last 5 Weeks
	(1)	(2)	(3)	(4)
Requests ($w = 38$)	2.224***	0.597***	0.383***	0.180**
	(0.223)	(0.141)	(0.108)	(0.057)
Recency	−0.189***	−0.106***	−0.140***	−0.124***
	(0.042)	(0.026)	(0.010)	(0.010)
Frequency	0.095***	0.049***	−0.00002	0.004*
	(0.007)	(0.004)	(0.002)	(0.002)
Routine ($w = 38$)	5.750***	3.284***	1.110**	0.307*
	(0.436)	(0.275)	(0.385)	(0.147)
Observations	2,000	2,000	2,000	2,000
R^2	0.567	0.425		

Note: *p<0.05; **p<0.01; ***p<0.001

Intercept omitted for clarity.

Are routine customers more valuable?

Total Requests (Week 38)

RF(M) Controls (Week 38)

Model-based Routineness (Week 38)

	Dependent variable:			
	# Sessions		Activity	
	OLS		*Logistic*	
	Full Holdout	Last 5 Weeks	Full Holdout	Last 5 Weeks
	(1)	(2)	(3)	(4)
Requests ($w = 38$)	2.224***	0.597***	0.383***	0.180**
	(0.223)	(0.141)	(0.108)	(0.057)
Recency	−0.189***	−0.106***	−0.140***	−0.124***
	(0.042)	(0.026)	(0.010)	(0.010)
Frequency	0.095***	0.049***	−0.00002	0.004*
	(0.007)	(0.004)	(0.002)	(0.002)
Routine ($w = 38$)	5.750***	3.284***	1.110**	0.307*
	(0.436)	(0.275)	(0.385)	(0.147)
Observations	2,000	2,000	2,000	2,000
R^2	0.567	0.425		

Note: *p<0.05; **p<0.01; ***p<0.001
Intercept omitted for clarity.

Key result: The number of rides *coming from a routine* is a significant predictor of **short- and long-run usage and retention.**

More to the story…

Are routine customers better in other ways?

More to the story…

Are routine customers better in other ways?

In ride-sharing, we find highly routine customers are…

- More likely to accept ride proposals and request again
- Less price sensitive
- Somewhat more picky about the rides they take
- Somewhat less sensitive to bad service (pickup delay)

More to the story…

Are routine customers better in other ways?

In ride-sharing, we find highly routine customers are…

- More likely to accept ride proposals and request again
- Less price sensitive
- Somewhat more picky about the rides they take
- Somewhat less sensitive to bad service (pickup delay)

In short: **yes!**

More to the story…

Are routine customers better in other ways?

In ride-sharing, we find highly routine customers are…

- More likely to accept ride proposals and request again
- Less price sensitive
- Somewhat more picky about the rides they take
- Somewhat less sensitive to bad service (pickup delay)

In short: **yes!**

What's more important: when someone is calling rides, or where they're going?

More to the story…

Are routine customers better in other ways?

In ride-sharing, we find highly routine customers are…

- More likely to accept ride proposals and request again
- Less price sensitive
- Somewhat more picky about the rides they take
- Somewhat less sensitive to bad service (pickup delay)

In short: **yes!**

What's more important: when someone is calling rides, or where they're going?

- Data: full set of locations for each customer, measured by "block"
- Two metrics of location dispersion:
    - Shannon entropy
    - "CRT dispersion"

More to the story…

Are routine customers better in other ways?

In ride-sharing, we find highly routine customers are…

- More likely to accept ride proposals and request again
- Less price sensitive
- Somewhat more picky about the rides they take
- Somewhat less sensitive to bad service (pickup delay)

In short: **yes!**

What's more important: when someone is calling rides, or where they're going?

- Data: full set of locations for each customer, measured by "block"
- Two metrics of location dispersion:
    - Shannon entropy
    - "CRT dispersion"

After controlling for timing routineness, location dispersion is **not significantly linked** to future activity.

Summary

- **Methodological:** Our model decomposes transaction histories into routine and random components
  - Gaussian process with novel day-hour kernel allows for precise individual-level routine estimates
  - Nesting GP in an inhomogeneous Poisson process → structured decomposition of usage
  - The result: a novel routineness metric

- **Substantive:** The shape of a customer's transaction history matters!
  - Additional evidence for the role of habit, and specifically routines, for CRM
  - A new "KPI" for predicting customer value: higher routineness = higher value
  - Routine customers are also better in other ways: price sensitivity, resilience to disruptions
  - Temporal routines are distinct from "what" (or "where") routines

Thank you!

Questions / comments?

ryandew@wharton.upenn.edu

Working paper available at www.rtdew.com