Detecting Routines in Transaction Data: Applications to CRM in Ride-sharing

Ryan Dew, Eva Ascarza, Oded Netzer, Nachum Sicherman

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?

- Our main hypothesis: routine users may have a higher lifetime value

Why might routines matter?

- Our main hypothesis: routine users may have a higher lifetime value
  - Habits → automaticity
  - Greater share of routine = more purchasing
  - Spillovers out of the routine

Why might routines matter?

- Our main hypothesis: routine users may have a higher lifetime value
  - Habits → automaticity
  - Greater share of routine = more purchasing
  - Spillovers out of the routine
- Routine users may be better (or worse!) customers in other ways

Why might routines matter?

- Our main hypothesis: routine users may have a higher lifetime value
  - Habits → automaticity
  - Greater share of routine = more purchasing
  - Spillovers out of the routine
- Routine users may be better (or worse!) customers in other ways
- An additional tool for segmentation, targeting, and managing supply chains

Some Conceptual Distinctions

- Behavioral research:
  - Long history of research on habits
    (as far back as James,1890)
  - Key ideas: automaticity, often triggered by contexts
    (e.g., Wood et al., 2002; Verplanken et al., 2008; White et al., 2019)

Some Conceptual Distinctions

- Behavioral research:
  - Long history of research on habits
    (as far back as James,1890)
  - Key ideas: automaticity, often triggered by contexts
    (e.g., Wood et al., 2002; Verplanken et al., 2008; White et al., 2019)

- In CRM, habit is often equivalent to repeated behavior
  - “Buying habit” = repeat purchasing
    (e.g., Ehrenberg & Goodhardt, 1968; Ascarza et al., 2016)
  - Purchasing on promotion, making returns, buying low-margin items
    (Shah et al., 2004)

Some Conceptual Distinctions

- Behavioral research:
  - Long history of research on habits
    (as far back as James,1890)
  - Key ideas: automaticity, often triggered by contexts
    (e.g., Wood et al., 2002; Verplanken et al., 2008; White et al., 2019)

- In CRM, habit is often equivalent to repeated behavior
  - "Buying habit" = repeat purchasing
    (e.g., Ehrenberg & Goodhardt, 1968; Ascarza et al., 2016)
  - Purchasing on promotion, making returns, buying low-margin items
    (Shah et al., 2004)

- In economics, "habit stock" used to model overly smooth consumption
  (Dynan, 2000)

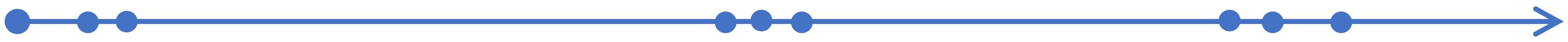
Characterizing Purchasing Behavior in CRM

- Buying habit: RF(M)
  (e.g., Fader et al., 2006)

Characterizing Purchasing Behavior in CRM

- Buying habit: RF(M)
  (e.g., Fader et al., 2006)

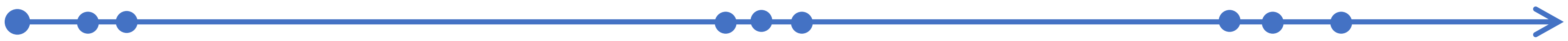
- Clumpiness
  (Zhang et al., 2015)



Characterizing Purchasing Behavior in CRM

- Buying habit: RF(M)
  (e.g., Fader et al., 2006)

- Clumpiness
  (Zhang et al., 2015)



- Regularity
  (Platzer and Reutterer, 2016)

This Project

1. Develop a novel model for detecting and quantifying individual-level routines from simple transaction data

This Project

1. Develop a novel model for detecting and quantifying individual-level routines from simple transaction data

2. Apply our model to a unique ride-sharing data set

This Project

1. Develop a novel model for detecting and quantifying individual-level routines from simple transaction data
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

This Project

1. Develop a novel model for detecting and quantifying individual-level routines from simple transaction data

2. Apply our model to a unique ride-sharing data set

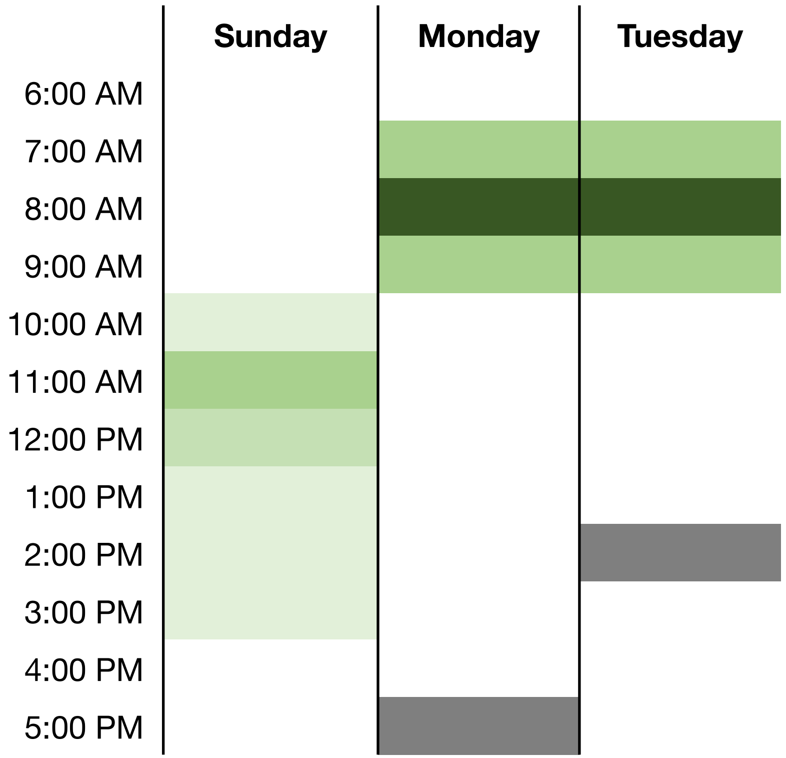
3. Show that customers with a high level of routine usage churn less, and spend more in the long run

4. Explore how (temporal) routineness predicts and moderates other important customer outcomes, over and above…
    - Mere habit
    - Routines in terms of “what”
    - Regularity and clumpiness

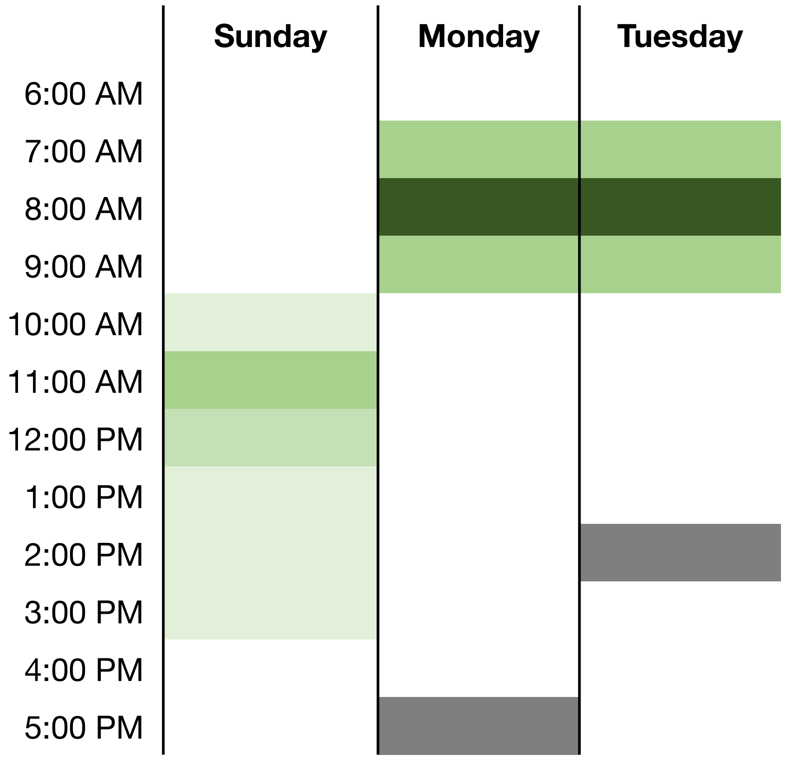
Model

A Statistical Framework for Measuring Routineness

From transaction data, we want to capture…

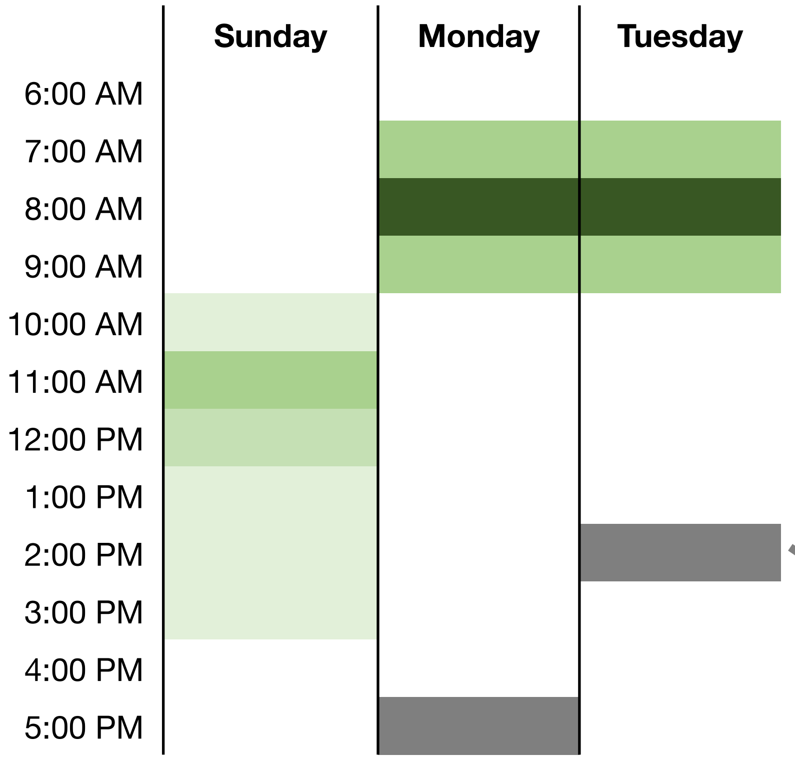


From transaction data, we want to capture…



And importantly:
How many uses were part of the routine?

From transaction data, we want to capture…



And importantly:
How many uses were part of the routine?

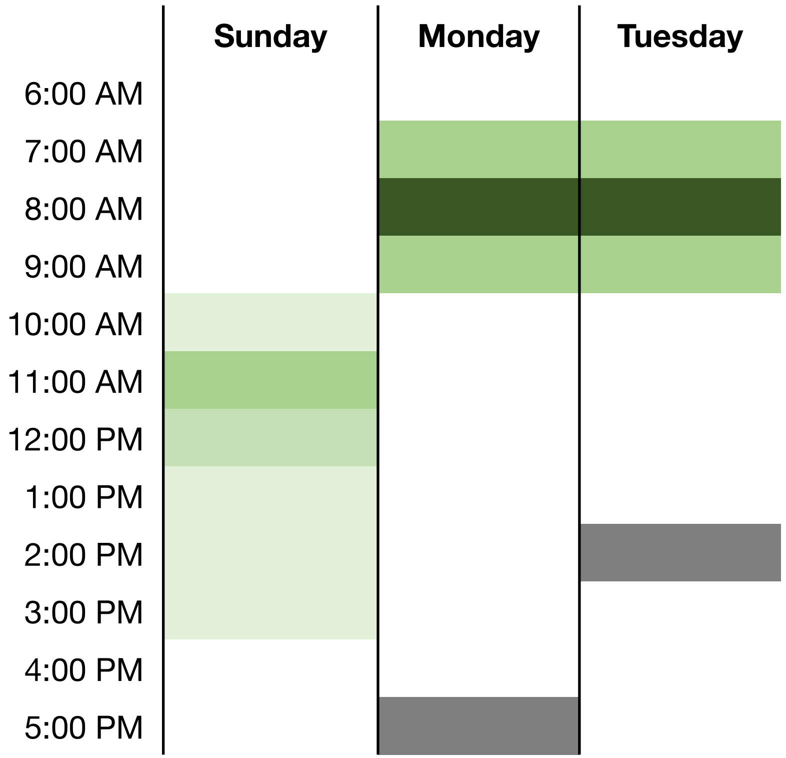
Routine Process

Requests arrive at with an individual-specific weekly rate, and an individual-specific day-hour rate

Random Process

Requests arrive at with an individual-specific weekly rate, and a common day-hour rate

From transaction data, we want to capture…



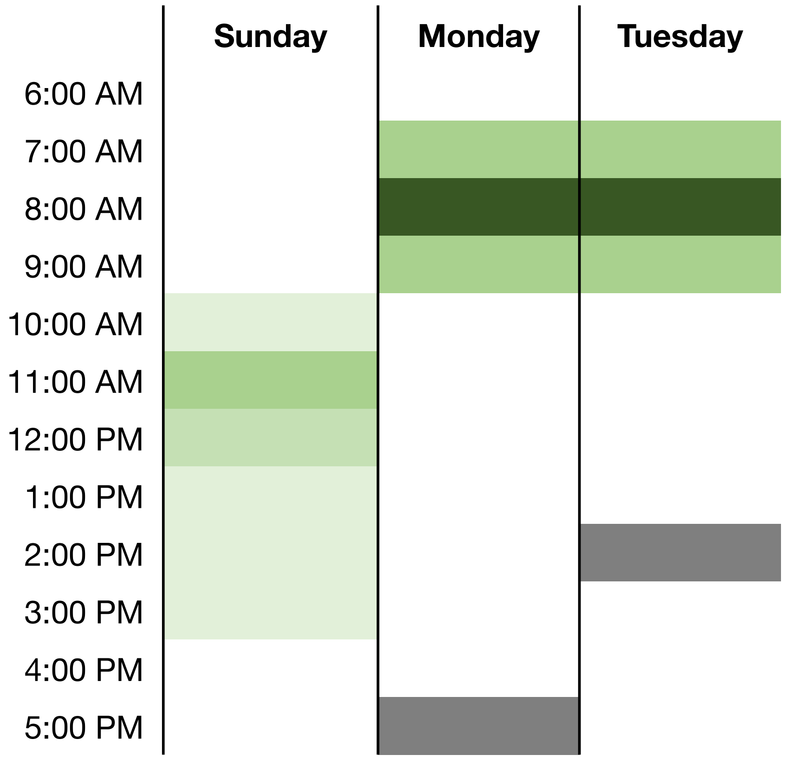
And importantly:
How many uses were part of the routine?

Routine Process
Requests arrive at with an individual-specific weekly rate, and an individual-specific day-hour rate

Joint Process
Superposition of the routine and random processes

Random Process
Requests arrive at with an individual-specific weekly rate, and a common day-hour rate

From transaction data, we want to capture…



And importantly:
How many uses were part of the routine?

Dependent variable: Usage (y)

Customer i, Week w, Day d, Hour h

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

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

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

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

Structured Decomposition:

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

Specifying the Rate(s)

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

Specifying the Rate(s)

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

"Scaling" – How active is this process in this week?

$$\alpha_i(w) \sim \mathcal{GP}(m_{\alpha_i}, k_e(w, w'; \phi_\alpha))$$

$$\gamma_i(w) \sim \mathcal{GP}(m_{\gamma_i}, k_e(w, w'; \phi_\gamma))$$

Specifying the Rate(s)

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

"Scaling" – How active is this process in this week?

$$\alpha_i(w) \sim \boxed{\mathcal{GP}}(m_{\alpha_i}, k_e(w, w'; \phi_\alpha))$$

$$\gamma_i(w) \sim \boxed{\mathcal{GP}}(m_{\gamma_i}, k_e(w, w'; \phi_\gamma))$$

Bayesian nonparametric priors over function spaces

Specifying the Rate(s)

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

“Scaling” – How active is this process in this week?

$$\alpha_i(w) \sim \boxed{\mathcal{GP}}(m_{\alpha_i}, k_e(w, w'; \phi_\alpha))$$

$$\gamma_i(w) \sim \boxed{\mathcal{GP}}(m_{\gamma_i}, k_e(w, w'; \phi_\gamma))$$

Bayesian nonparametric priors over function spaces

$$f(x) \sim \mathcal{GP}(m(x), k(x, x'))$$

$$\Updownarrow$$

$$f(x_1, \ldots, x_N) \sim \mathcal{N}(m(x_1, \ldots, x_N), K), \ \text{s.t.} \ K_{ij} = k(x_i, x_j)$$

Specifying the Rate(s)

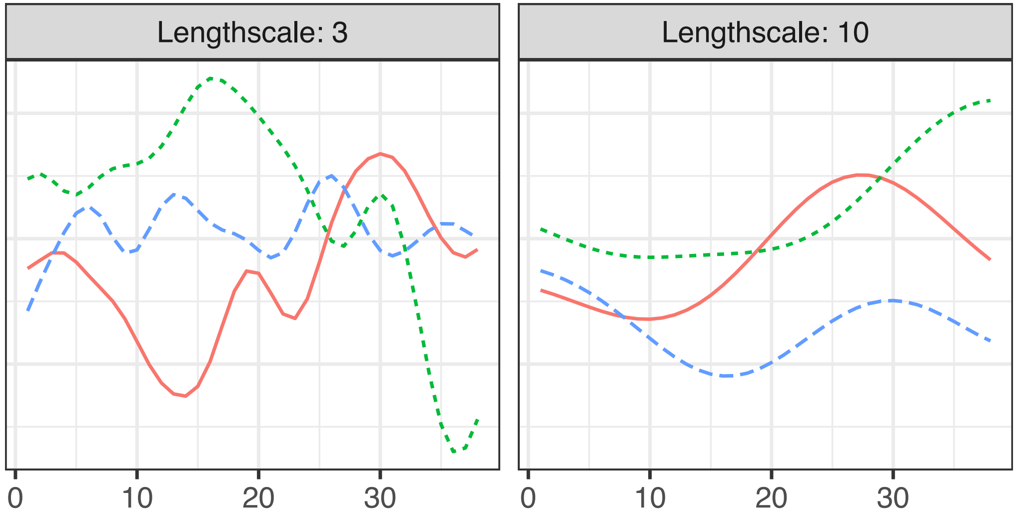
$$\lambda_{it} = \overbrace{\exp(\alpha_{iw} + \mu_j)}^{\text{Random usage}} + \overbrace{\exp(\gamma_{iw} + \eta_{ij})}^{\text{Routine usage}}$$

"Scaling" – How active is this process in this week?

$$\alpha_i(w) \sim \mathcal{GP}(m_{\alpha_i}, \boxed{k_e(w, w'; \phi_\alpha)})$$

$$\gamma_i(w) \sim \mathcal{GP}(m_{\gamma_i}, \boxed{k_e(w, w'; \phi_\gamma)})$$

$$\boxed{k_e(w, w') = \sigma^2 \exp\left(-\frac{|w - w'|}{\ell}\right)}$$



Specifying the Rate(s)

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

"Day-Hour Rate" – When do we expect usage to occur?

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

Specifying the Rate(s)

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

"Day-Hour Rate" – When do we expect usage to occur?

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

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

Novel "day-hour" kernel embeds assumptions about how routines works.

Specifying the Rate(s)

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

"Day-Hour Rate" – When do we expect usage to occur?

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

$$\boxed{k_{\mathrm{DH}}(\mathbf{d}, \mathbf{h}; \boldsymbol{\phi}) = \Omega_{d,d'} \times \exp\left[-\frac{2}{\rho^2}\sin^2\left(\frac{\pi|h-h'|}{24}\right)\right]}$$

Novel "day-hour" kernel embeds assumptions about how routines works.

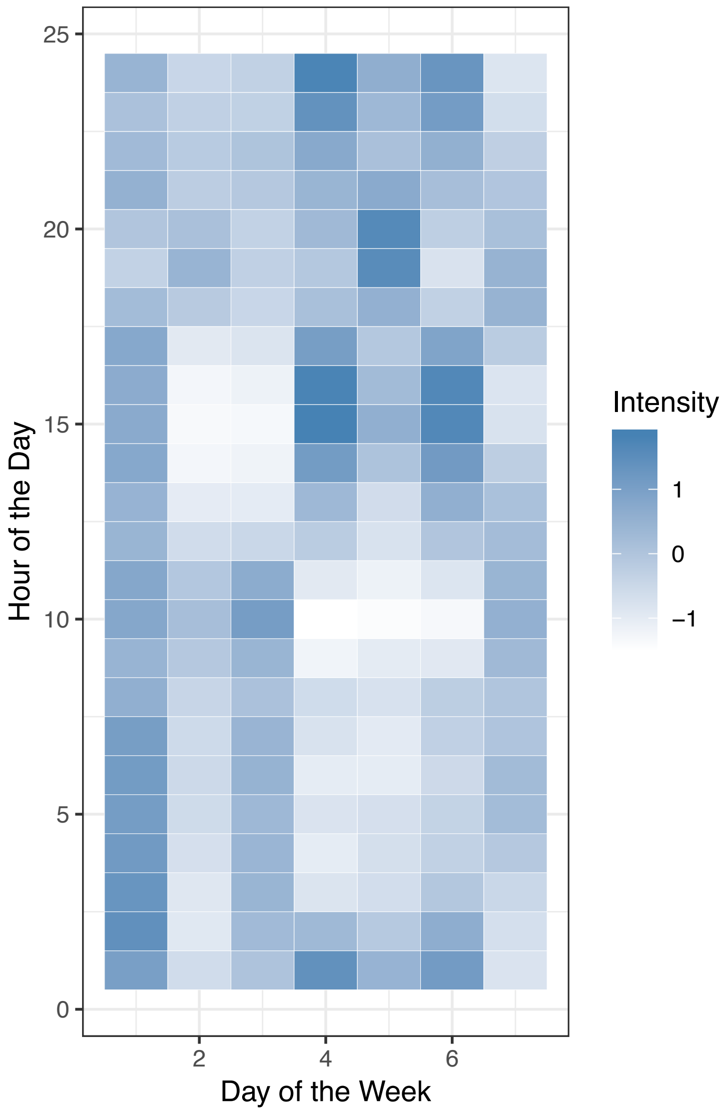
Specifying the Rate(s)

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

“Day-Hour Rate” – When do we expect usage to occur?

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

Novel “day-hour” kernel embeds assumptions about how routines works.



Specifying the Rate(s)

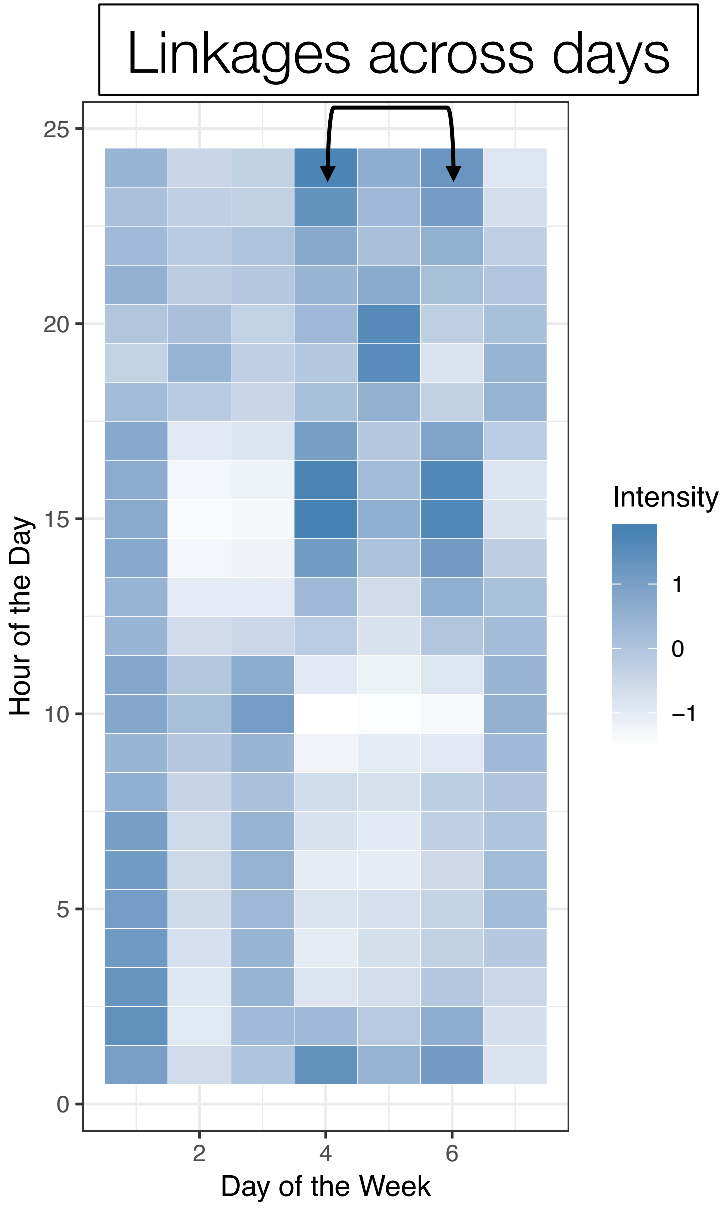
$$\lambda_{it} = \underbrace{\exp(\alpha_{iw} + \mu_j)}_{\text{Random usage}} + \underbrace{\exp(\gamma_{iw} + \eta_{ij})}_{\text{Routine usage}}$$

"Day-Hour Rate" – When do we expect usage to occur?

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

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

Novel "day-hour" kernel embeds assumptions about how routines works.



Specifying the Rate(s)

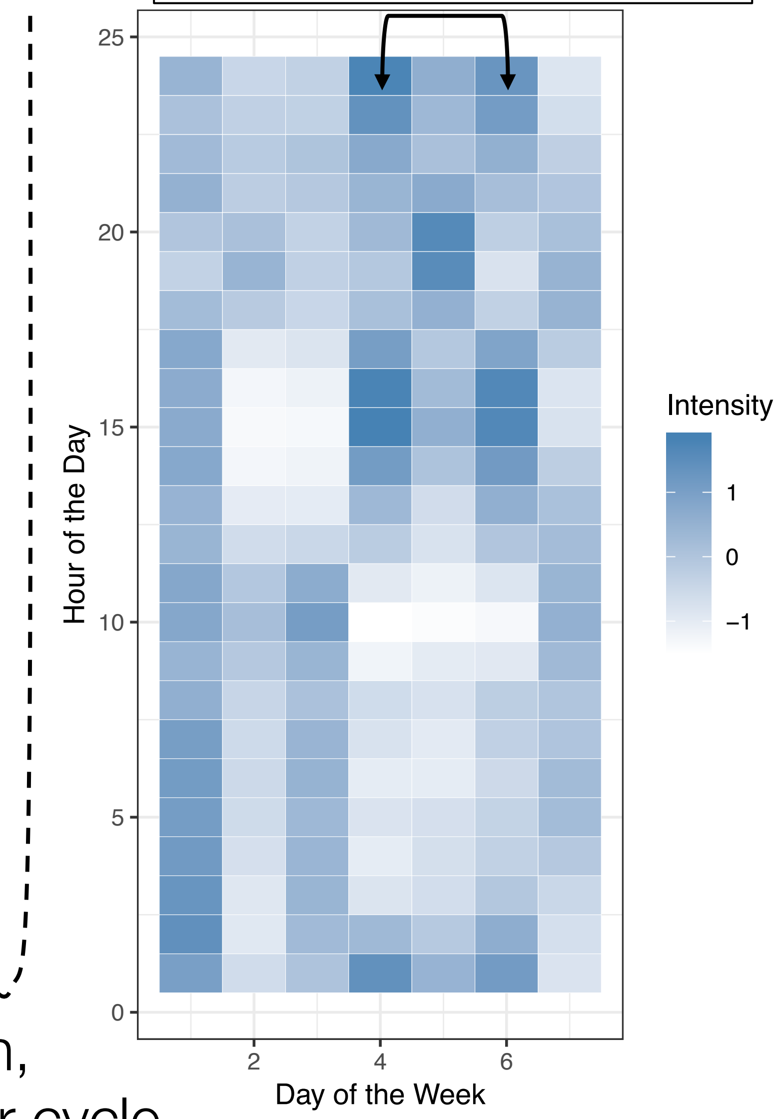
$$\lambda_{it} = \underbrace{\exp(\alpha_{iw} + \mu_j)}_{\text{Random usage}} + \underbrace{\exp(\gamma_{iw} + \eta_{ij})}_{\text{Routine usage}}$$

"Day-Hour Rate" – When do we expect usage to occur?

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

Novel "day-hour" kernel embeds assumptions about how routines works.

Linkages across days



Smooth, 24-hour cycle

Specifying the Rate(s)

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

“Day-Hour Rate” – When do we expect usage to occur?

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

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

Novel “day-hour” kernel embeds assumptions about how routines works.

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})

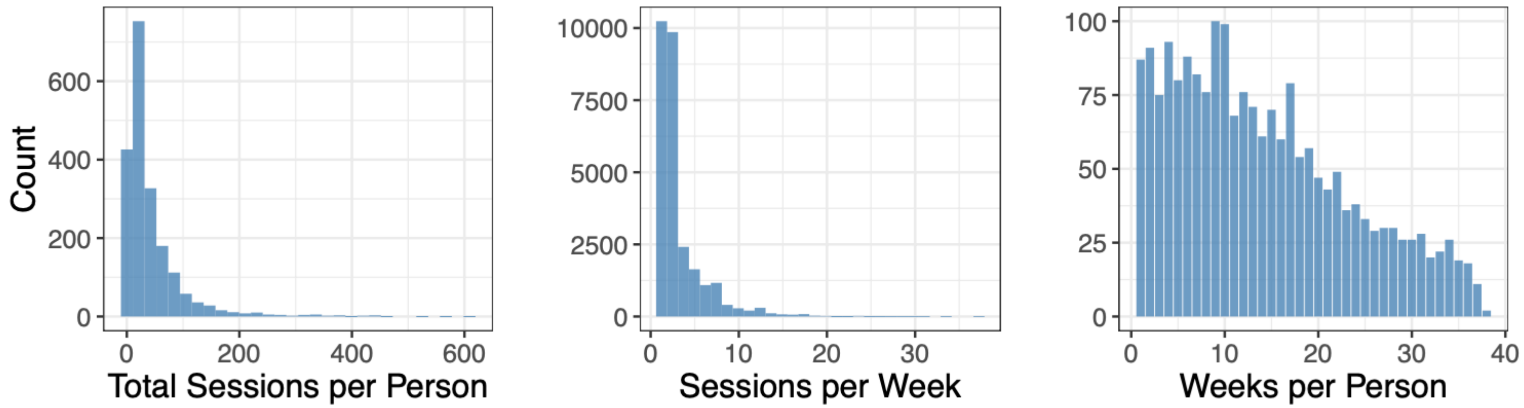
Results

Application to Ride-sharing Data

Ride-sharing Data

- Collaboration with a NYC-based ride-sharing company

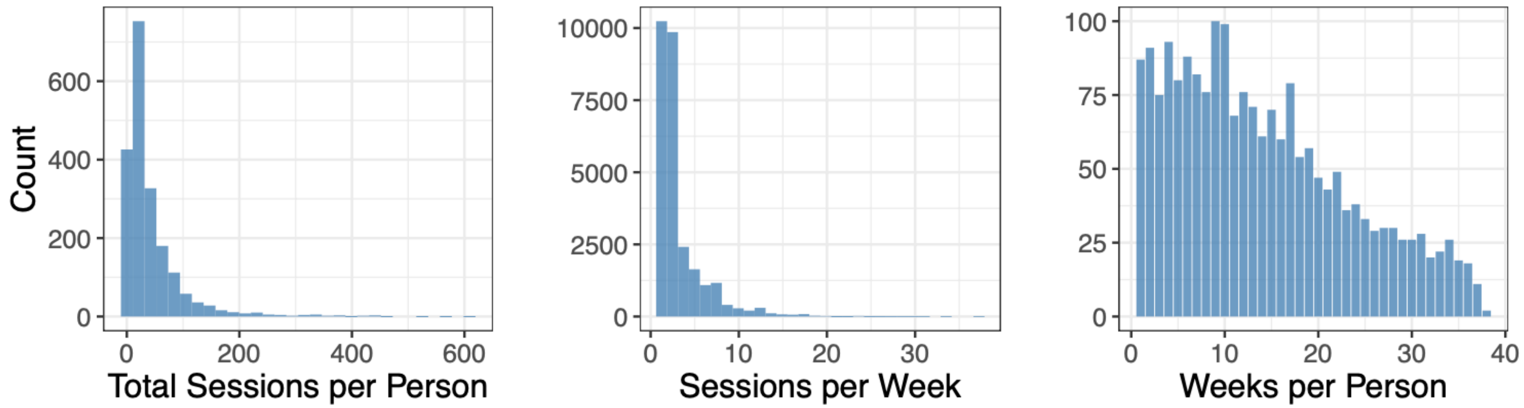
Total Customers	2,000
Total Weeks (Training)	38
Total Weeks (Holdout)	10
Number of Sessions	86,952
Sessions / Customer	43.48
Sessions / Customer / Week	3.10
Weeks in Data / Customer	14.02



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

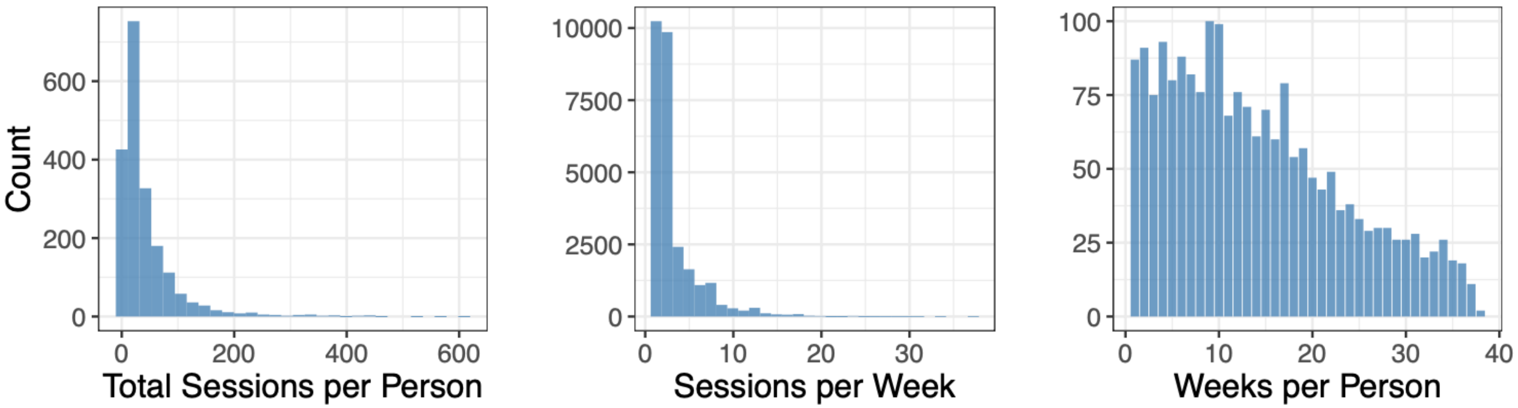


Ride-sharing Data

- 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 + fake customers with specific usage patterns

“Quasi-simulation”

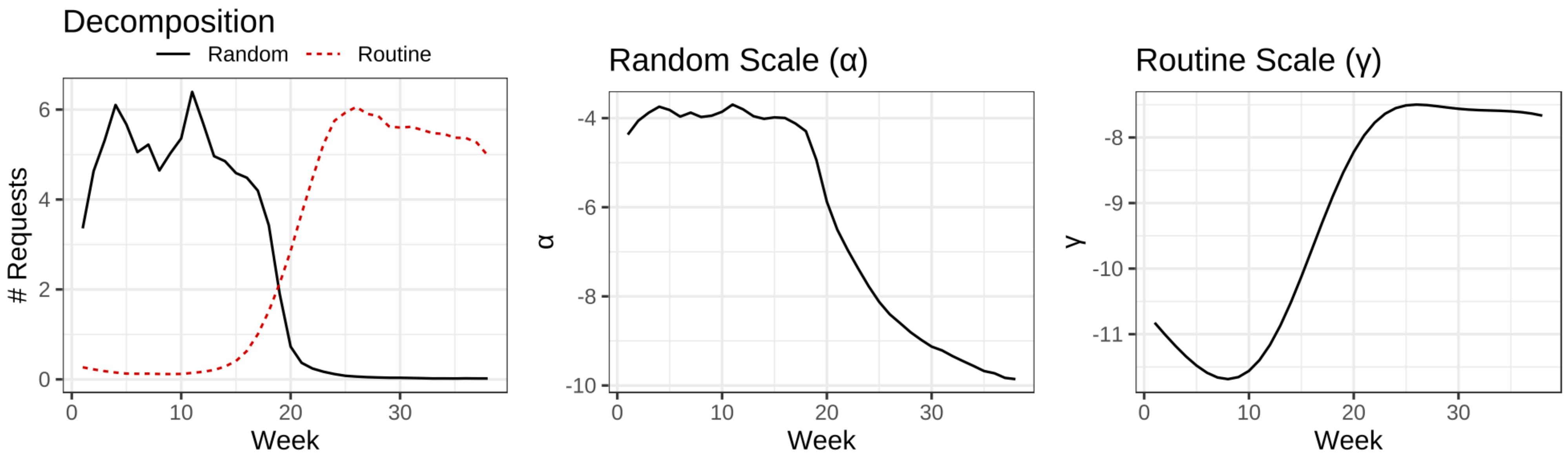
- 500 real customers + fake customers with specific usage patterns

“Fake” Case Study: Random then Routine

“Quasi-simulation”

- 500 real customers + fake customers with specific usage patterns

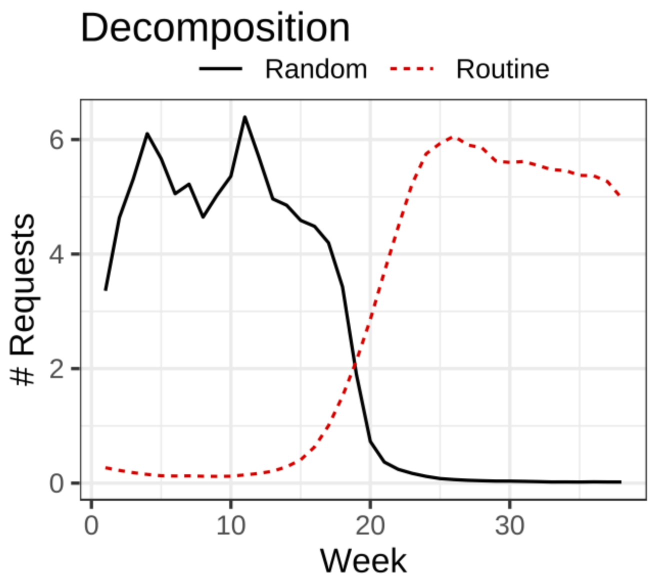
“Fake” Case Study: Random then Routine

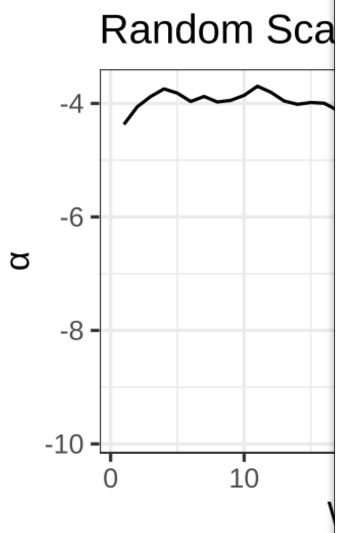


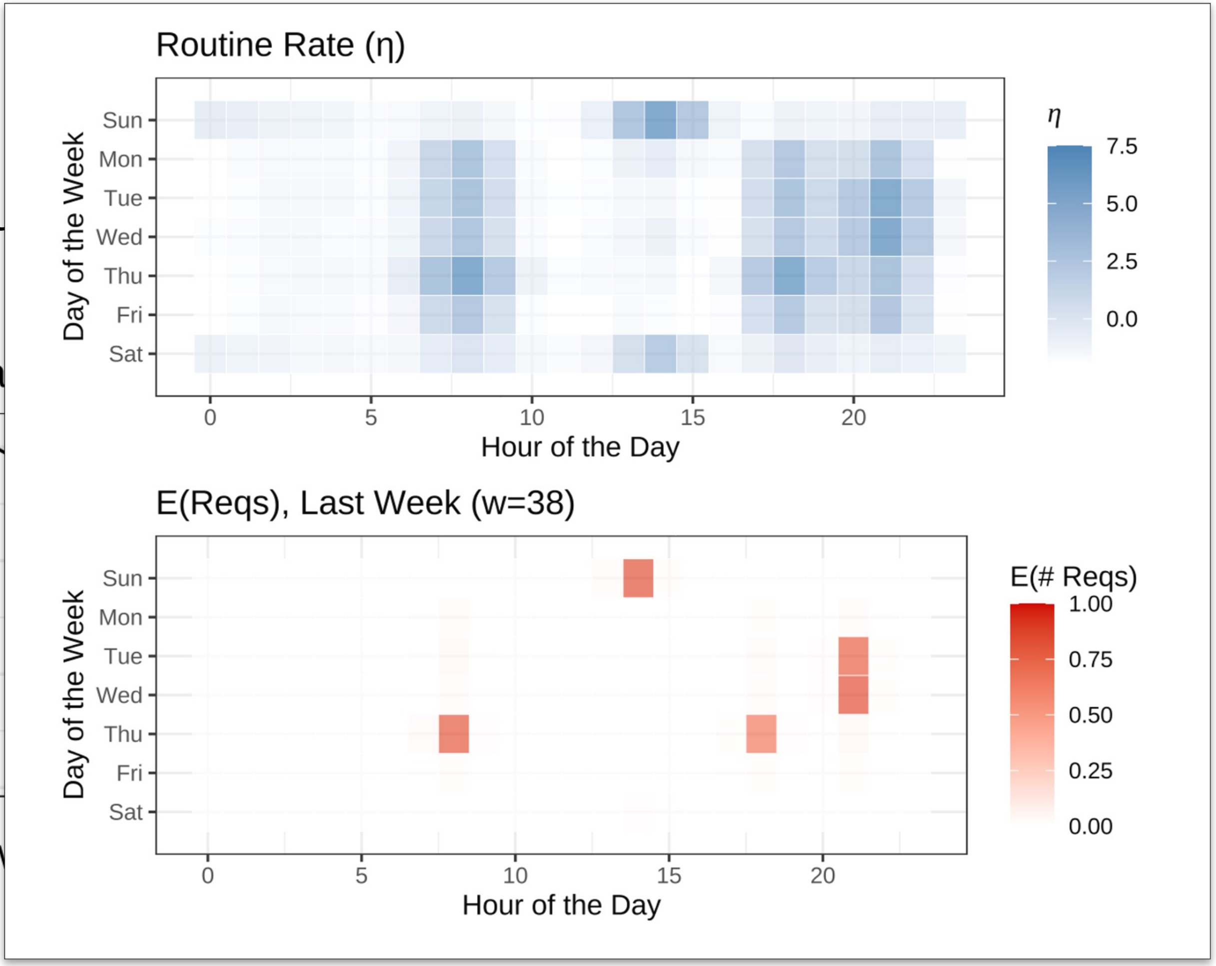
“Quasi-simulation”

- 500 real customers + fake customers with specific usage patterns

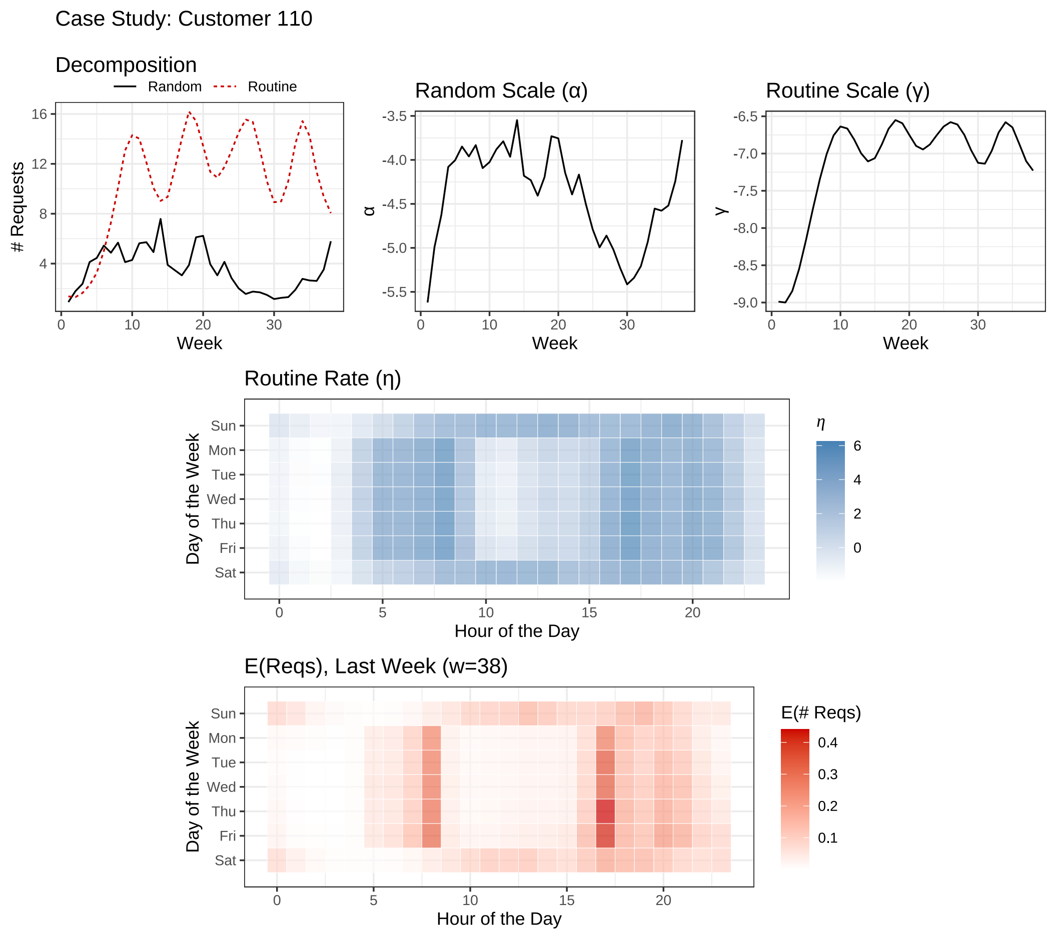
“Fake” Case Study: Random then Rou



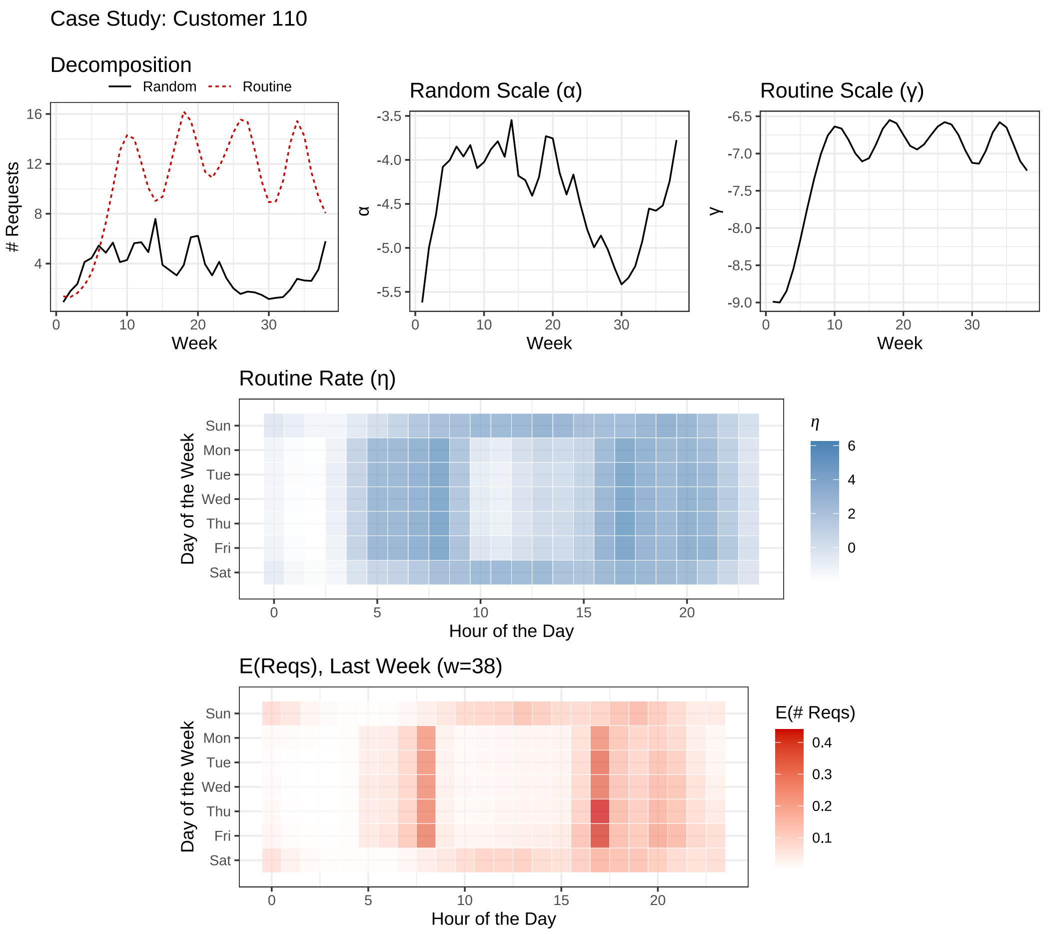


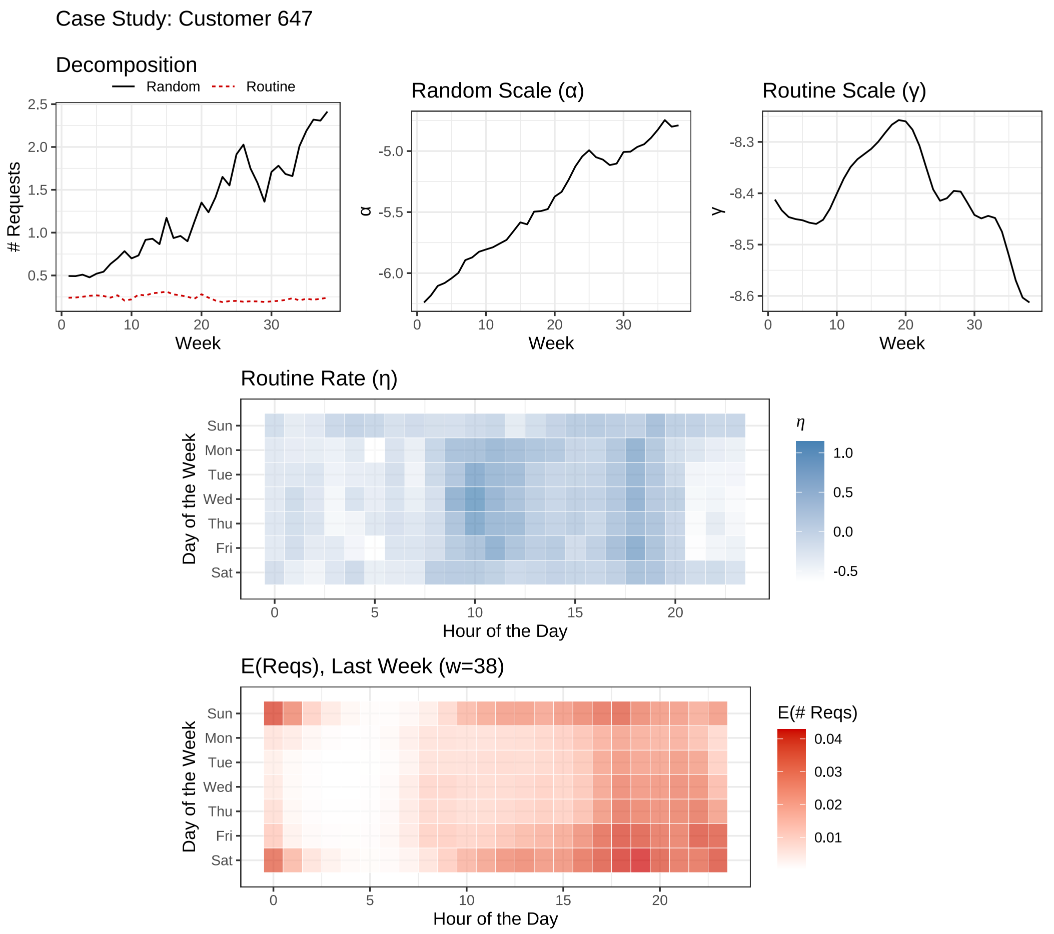


Real Case Studies

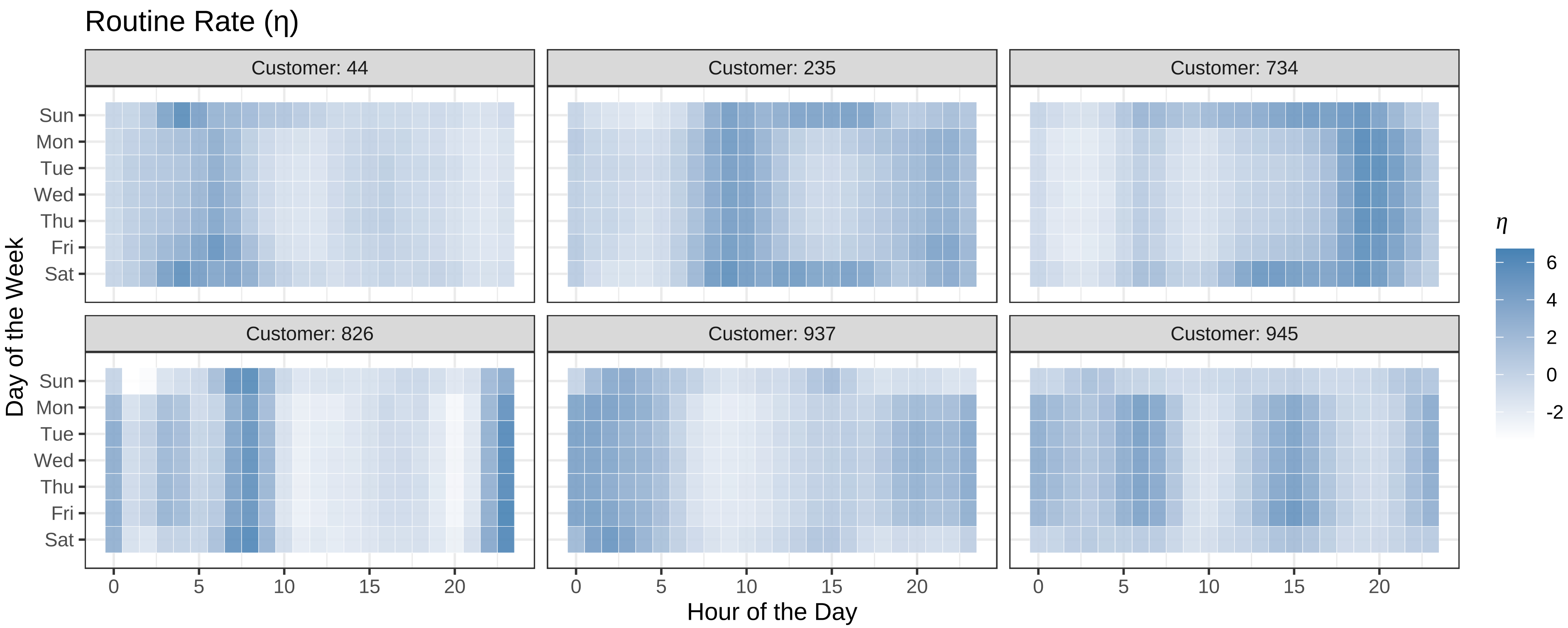


Real Case Studies

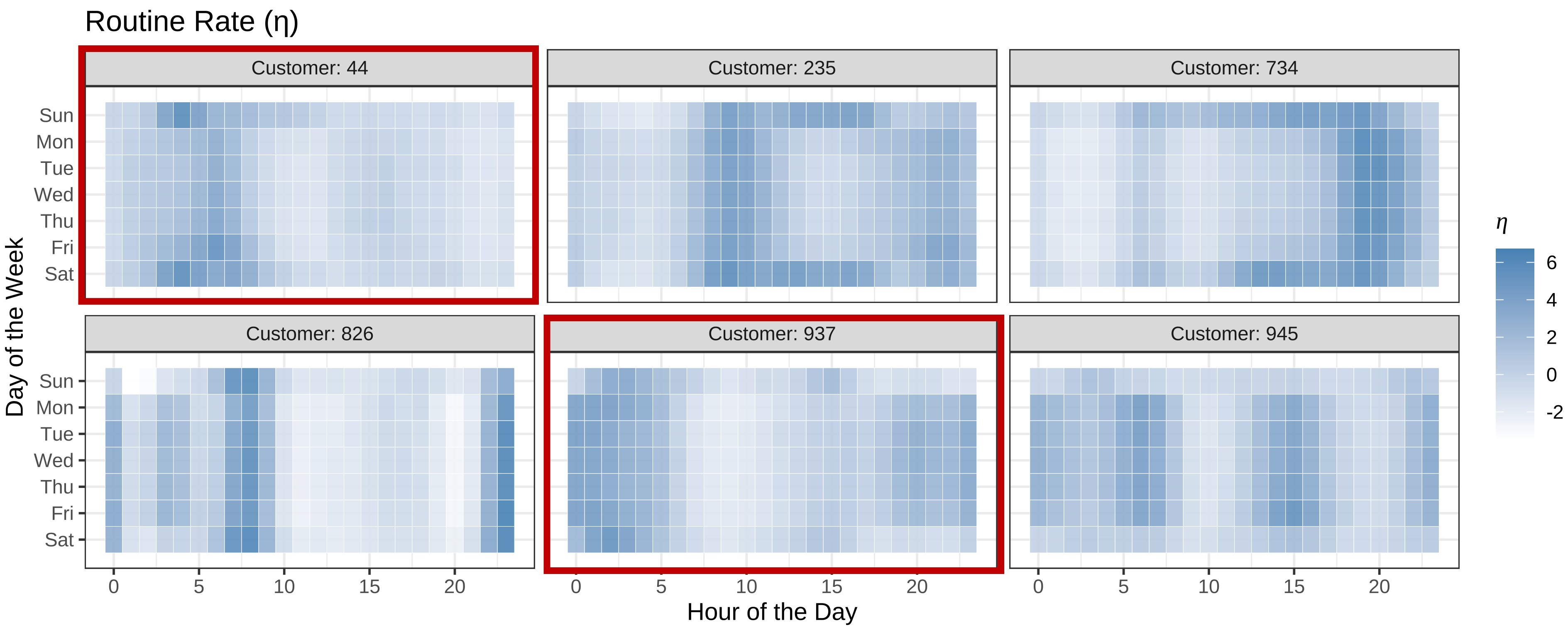




Many types of routines



Many types of routines



"Night owls"

Many types of routines

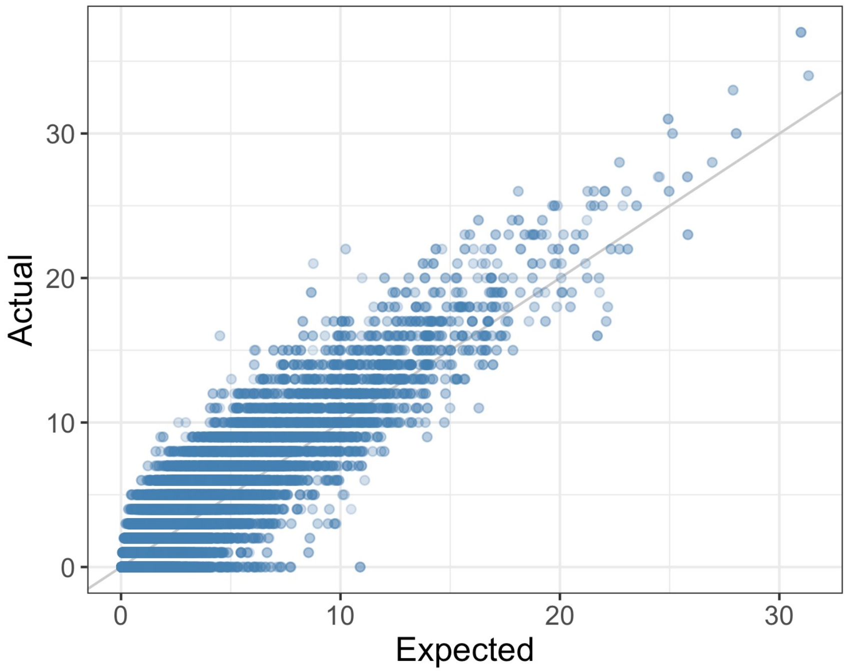
Morning and night

Validation

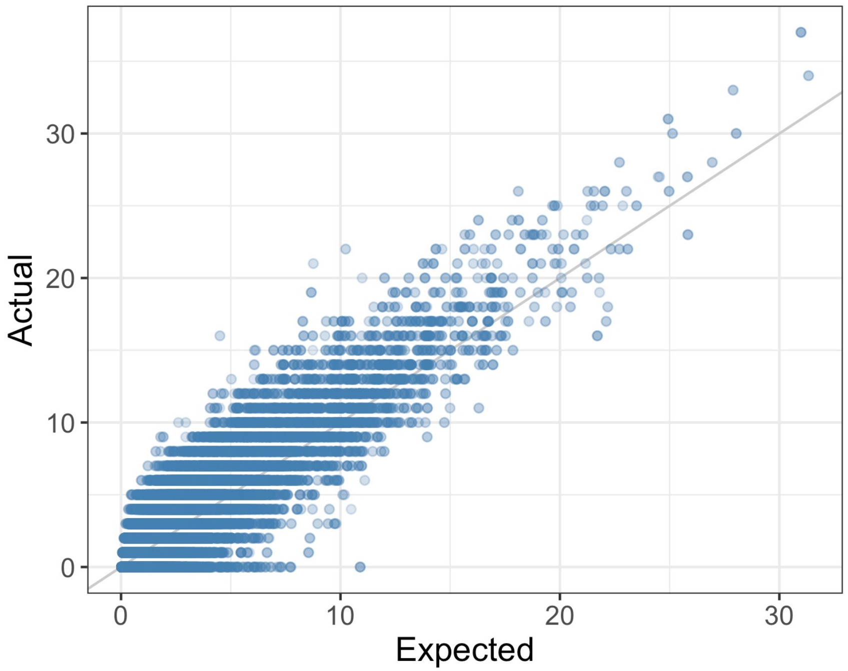
Fit and Forecasting Relative to Benchmarks

[Skip]

Predictive Performance

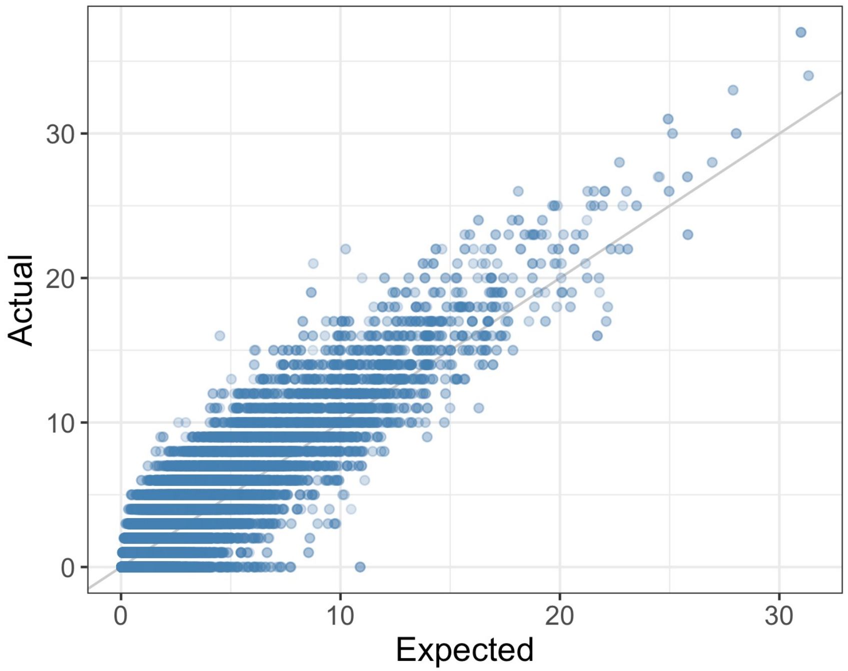


Predictive Performance



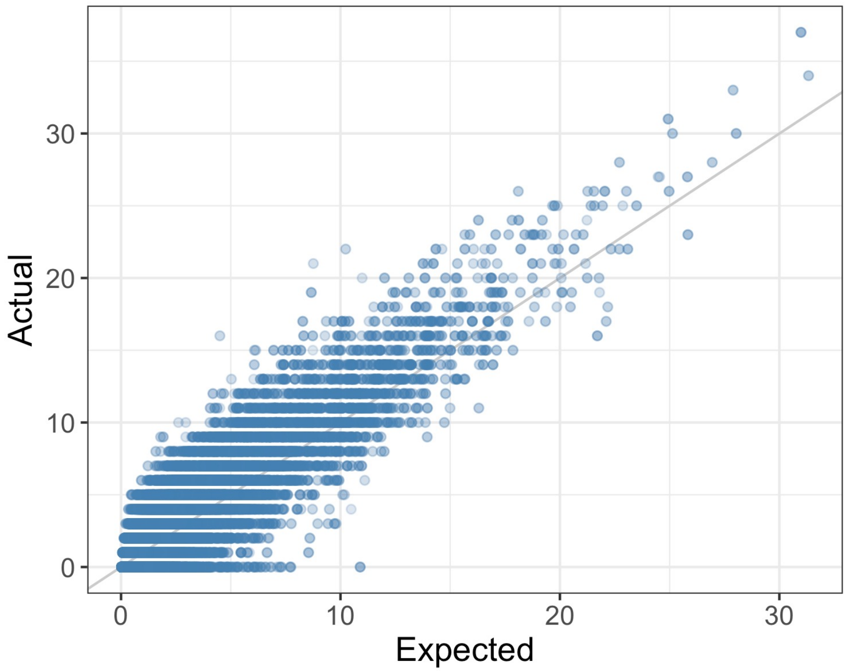
- Three benchmarks:
  - Current model with no routines
  - State-of-the-art Pareto-GGG probability model (captures regularity)
  - Customer-specific LSTM trained to predict ride times

Predictive Performance



- Three benchmarks:
  - Current model with no routines
  - State-of-the-art Pareto-GGG probability model (captures regularity)
  - Customer-specific LSTM trained to predict ride times
- Forecasting # sessions (MAE):
  - Equivalent performance to Pareto-GGG
  - Significantly outperforms LSTM

Predictive Performance



- Three benchmarks:
  - Current model with no routines
  - State-of-the-art Pareto-GGG probability model (captures regularity)
  - Customer-specific LSTM trained to predict ride times
- Forecasting # sessions (MAE):
  - Equivalent performance to Pareto-GGG
  - Significantly outperforms LSTM
- More interesting: predicting ride times
  - Measured by ranking expected request times
  - Significantly outperforms all benchmarks

[More Info]

Implications

Why should we care about routines?

Are routine customers more valuable?

	Dependent variable:			
	Full Holdout	Last 5 Weeks	Full Holdout	Last 5 Weeks
	OLS	*OLS*	*logistic*	*logistic*
	(1)	(2)	(3)	(4)
Requests ($w = 38$)	1.768***	0.467***	0.398***	0.178***
	(0.235)	(0.150)	(0.104)	(0.055)
Recency	−0.222***	−0.122***	−0.141***	−0.125***
	(0.042)	(0.026)	(0.010)	(0.010)
Frequency	0.090***	0.048***	−0.001	0.004**
	(0.007)	(0.004)	(0.002)	(0.002)
Routine ($w = 38$)	5.435***	2.802***	1.057***	0.288**
	(0.377)	(0.240)	(0.328)	(0.123)
Observations	2,000	2,000	2,000	2,000
R^2	0.574	0.424		
Note:				*p<0.1; **p<0.05; ***p<0.01 Intercept omitted for clarity.

Are routine customers more valuable?

Total Requests
(Week 38)

	Dependent variable:			
	Full Holdout	Last 5 Weeks	Full Holdout	Last 5 Weeks
	OLS	*OLS*	*logistic*	*logistic*
	(1)	(2)	(3)	(4)
$\boxed{\text{Requests } (w = 38)}$	1.768***	0.467***	0.398***	0.178***
	(0.235)	(0.150)	(0.104)	(0.055)
Recency	−0.222***	−0.122***	−0.141***	−0.125***
	(0.042)	(0.026)	(0.010)	(0.010)
Frequency	0.090***	0.048***	−0.001	0.004**
	(0.007)	(0.004)	(0.002)	(0.002)
Routine $(w = 38)$	5.435***	2.802***	1.057***	0.288**
	(0.377)	(0.240)	(0.328)	(0.123)
Observations	2,000	2,000	2,000	2,000
R^2	0.574	0.424		
Note:			*p<0.1; **p<0.05; ***p<0.01	
			Intercept omitted for clarity.	

Are routine customers more valuable?

Total Requests
(Week 38)

RF(M) Controls
(Week 38)

	Dependent variable:			
	Full Holdout	Last 5 Weeks	Full Holdout	Last 5 Weeks
	OLS	*OLS*	*logistic*	*logistic*
	(1)	(2)	(3)	(4)
Requests ($w = 38$)	1.768***	0.467***	0.398***	0.178***
	(0.235)	(0.150)	(0.104)	(0.055)
Recency	−0.222***	−0.122***	−0.141***	−0.125***
	(0.042)	(0.026)	(0.010)	(0.010)
Frequency	0.090***	0.048***	−0.001	0.004**
	(0.007)	(0.004)	(0.002)	(0.002)
Routine ($w = 38$)	5.435***	2.802***	1.057***	0.288**
	(0.377)	(0.240)	(0.328)	(0.123)
Observations	2,000	2,000	2,000	2,000
R^2	0.574	0.424		

Note: *p<0.1; **p<0.05; ***p<0.01
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:			
	Full Holdout	Last 5 Weeks	Full Holdout	Last 5 Weeks
	OLS	*OLS*	*logistic*	*logistic*
	(1)	(2)	(3)	(4)
Requests ($w = 38$)	1.768***	0.467***	0.398***	0.178***
	(0.235)	(0.150)	(0.104)	(0.055)
Recency	−0.222***	−0.122***	−0.141***	−0.125***
	(0.042)	(0.026)	(0.010)	(0.010)
Frequency	0.090***	0.048***	−0.001	0.004**
	(0.007)	(0.004)	(0.002)	(0.002)
Routine ($w = 38$)	5.435***	2.802***	1.057***	0.288**
	(0.377)	(0.240)	(0.328)	(0.123)
Observations	2,000	2,000	2,000	2,000
R^2	0.574	0.424		

Note: *p<0.1; **p<0.05; ***p<0.01

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:			
	Full Holdout	Last 5 Weeks	Full Holdout	Last 5 Weeks
	OLS	*OLS*	*logistic*	*logistic*
	(1)	(2)	(3)	(4)
Requests ($w = 38$)	1.768***	0.467***	0.398***	0.178***
	(0.235)	(0.150)	(0.104)	(0.055)
Recency	−0.222***	−0.122***	−0.141***	−0.125***
	(0.042)	(0.026)	(0.010)	(0.010)
Frequency	0.090***	0.048***	−0.001	0.004**
	(0.007)	(0.004)	(0.002)	(0.002)
Routine ($w = 38$)	5.435***	2.802***	1.057***	0.288**
	(0.377)	(0.240)	(0.328)	(0.123)
Observations	2,000	2,000	2,000	2,000
R^2	0.574	0.424		

Note: *p<0.1; **p<0.05; ***p<0.01
Intercept omitted for clarity.

Routineness vs. Regularity vs. Clumpiness

	Dependent variable:			
	# Requests		Activity (binary)	
	OLS		*logistic*	
	(1)	(2)	(3)	(4)
Requests ($w = 38$)	3.838***	1.614***	0.611***	0.461***
	(0.184)	(0.241)	(0.103)	(0.106)
Recency	−0.238***	−0.284***	−0.140***	−0.134***
	(0.045)	(0.043)	(0.010)	(0.010)
Frequency	0.130***	0.097***	0.0003	−0.002
	(0.008)	(0.008)	(0.002)	(0.002)
Regularity (k)	10.533***	4.511**	0.313	0.005
	(1.986)	(1.952)	(0.463)	(0.484)
Clumpiness (H)	9.269***	8.443***	−1.491***	−1.520***
	(1.918)	(1.836)	(0.380)	(0.381)
Routine ($w = 38$)		5.216***		1.035***
		(0.385)		(0.328)
Observations	2,000	2,000	2,000	2,000
R^2	0.540	0.579		

Note: *p<0.1; **p<0.05; ***p<0.01
Intercept omitted for clarity.

Routineness vs. Regularity vs. Clumpiness

	Dependent variable:			
	# Requests		Activity (binary)	
	OLS		*logistic*	
	(1)	(2)	(3)	(4)
Requests ($w = 38$)	3.838***	1.614***	0.611***	0.461***
	(0.184)	(0.241)	(0.103)	(0.106)
Recency	−0.238***	−0.284***	−0.140***	−0.134***
	(0.045)	(0.043)	(0.010)	(0.010)
Frequency	0.130***	0.097***	0.0003	−0.002
	(0.008)	(0.008)	(0.002)	(0.002)
Regularity (k)	10.533***	4.511**	0.313	0.005
	(1.986)	(1.952)	(0.463)	(0.484)
Clumpiness (H)	9.269***	8.443***	−1.491***	−1.520***
	(1.918)	(1.836)	(0.380)	(0.381)
Routine ($w = 38$)		5.216***		1.035***
		(0.385)		(0.328)
Observations	2,000	2,000	2,000	2,000
R^2	0.540	0.579		

Note: *p<0.1; **p<0.05; ***p<0.01
Intercept omitted for clarity.

Regularity's effect significantly dampened

Routineness remains positive

More to the story

- Routine trips are different: customers appear less sensitive to price during their routines, but much more sensitive to trip speed

More to the story

- Routine trips are different: customers appear less sensitive to price during their routines, but much more sensitive to trip speed

- Routine customers are also, on average, different:
  - More likely to accept proposals, in general
  - Less sensitive to trip characteristics (price, driver ETA), more sensitive to convenience factors (walking distance, pickup delay)

More to the story

- Routine trips are different: customers appear less sensitive to price during their routines, but much more sensitive to trip speed
- Routine customers are also, on average, different:
  - More likely to accept proposals, in general
  - Less sensitive to trip characteristics (price, driver ETA), more sensitive to convenience factors (walking distance, pickup delay)
- Location (i.e., *what*) doesn't matter as much as timing (i.e., *when*)

More to the story

- Routine trips are different: customers appear less sensitive to price during their routines, but much more sensitive to trip speed
- Routine customers are also, on average, different:
  - More likely to accept proposals, in general
  - Less sensitive to trip characteristics (price, driver ETA), more sensitive to convenience factors (walking distance, pickup delay)
- Location (i.e., *what*) doesn't matter as much as timing (i.e., *when*)
- **Heterogeneity:** Many of these effects vary by the routine type

More to the story

- Routine trips are different: customers appear less sensitive to price during their routines, but much more sensitive to trip speed
- Routine customers are also, on average, different:
  - More likely to accept proposals, in general
  - Less sensitive to trip characteristics (price, driver ETA), more sensitive to convenience factors (walking distance, pickup delay)
- Location (i.e., *what*) doesn't matter as much as timing (i.e., *when*)
- **Heterogeneity:** Many of these effects vary by the routine type
- **Spillovers:** Small but significant average correlation (0.095) between prior week's routineness and next week's randomness

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

Supplement

Forecast Metrics

- Benchmarks:
  - Pareto-GGG (as implemented in the BTYDPlus package)
  - Individual-specific LSTM, trained to predict ride times

Forecast Metrics

- Benchmarks:
  - Pareto-GGG (as implemented in the BTYDPlus package)
  - Individual-specific LSTM, trained to predict ride times

- Predict number of future requests: mean absolute error across customers

Model	MAE	CI
Ours	7.73	[7.171, 8.284]
P-GGG	7.513	[6.973, 8.051]
LSTM	26.13	[19.454, 32.796]

Forecast Metrics

- Benchmarks:
  - Pareto-GGG (as implemented in the BTYDPlus package)
  - Individual-specific LSTM, trained to predict ride times
- Predict number of future requests: mean absolute error across customers
- Predict *when* rides will take place: rank all possible times, compute precision metrics

Conditional Precision:

Model	Mean	CI
Ours	0.072	[0.066, 0.079]
P-GGG	0.016	[0.013, 0.018]
LSTM	0.089	[0.076, 0.102]

Average Precision:

Model	Mean	CI
Ours	0.131	[0.124, 0.139]
P-GGG	0.043	[0.041, 0.045]
LSTM	0.110	[0.076, 0.102]

[Back]

Routine Trips are Different

- DVs, measured throughout data:
  - Accept proposal, given characteristics
  - Request again within 7 days, given proposal and trip characteristics
- IVs / Controls:
  - Key moderator: the routineness of the day-hour
  - Customer fixed effects
  - Ride characteristics (car type, time of day, etc.)
- Key findings:
  - Less price sensitive during routine
  - More sensitive to speed

	Dependent variable:	
	Accept Proposal	Request Again
	(1)	(2)
Routineness	−0.007	−0.010
Price	−0.0001***	−0.00004***
Driver ETA	−0.023***	0.0003
Speed	0.007**	0.001
Pickup Walking Dist.	−0.0003***	0.00002
# Passengers	0.019***	0.018***
Pickup Delay		−0.0002
Dropoff Delay		−0.0002***
Dropoff Walking Dist.		−0.00003*
# On-board (Pickup)		−0.004**
# On-board (Dropoff)		0.001
Max On-board		−0.002
Routine x Price	0.0001***	0.0001***
Routine x Driver ETA	0.005	−0.0002
Routine x Speed	1.425***	−0.218
Routine x Pickup Walking Dist.	0.0002	−0.0001
Routine x # Passengers	−0.037	0.048*
Routine x Pickup Delay		0.001
Routine x Dropoff Delay		0.001
Routine x Dropoff Walking Dist.		0.0001
Routine x # On-board (Pickup)		0.012
Routine x # On-board (Dropoff)		0.0002
Routine x Max On-board		−0.004
Customer Fixed Effects	Yes	Yes
Observations	113,042	73,630
R^2	0.070	0.015

Note: *p<0.1; **p<0.05; ***p<0.01

[Back]

Routine Customers are Different

- Same DVs as before, but during *just the holdout*
- IVs / Controls:
  - Key IV/moderator: # routine requests at week 38 (end of training)
  - Customer random effects
  - Ride characteristics (car type, time of day, etc.)
- Key findings:
  - Routine customers are less sensitive to trip characteristics (price, driver ETA)
  - More sensitive to convenience factors (walking distance, pickup delay)

	Dependent variable:	
	Accept Proposal	Request Again
	(1)	(2)
# Requests (Week 38)	−0.004	0.021***
Routineness	0.015***	0.011**
Price	−0.00003***	−0.00003***
Driver ETA	−0.015***	−0.0003
ETA Destination	−0.001**	−0.0002
Speed	−0.030***	0.007
Pickup Walking Dist.	−0.0004***	−0.00002
# Passengers	−0.005	0.008
Pickup Delay		−0.004***
Dropoff Delay		−0.0003
Dropoff Walking Dist.		−0.0001
# On-board (Pickup)		−0.006
# On-board (Dropoff)		0.001
Max On-board		0.003
Routine x Price	−0.00000	0.00000***
Routine x Driver ETA	0.001***	0.0001
Routine x ETA Destination	−0.00000	−0.0001
Routine x Speed	0.063***	−0.009
Routine x Pickup Walking Dist.	−0.00003***	0.00000
Routine x # Passengers	−0.0003	0.002
Routine x Pickup Delay		0.001**
Routine x Dropoff Delay		0.00003
Routine x Dropoff Walking Dist.		0.00000
Routine x # On-board (Pickup)		0.001
Routine x # On-board (Dropoff)		−0.0001
Routine x Max On-board		−0.0005
Observations	38,166	14,704
R^2	0.051	0.093
Note:		*p<0.1; **p<0.05; ***p<0.01

[Back]

Routines of “what” vs. “when”

Consumers can have routines in both *when* and *what* they buy.

- In ride-sharing: are people routine in *where* people are going, in addition to *when* they’re going?
- Data: full set of locations for each customer, measured by “block”
- Two metrics of **location dispersion**:
  - Shannon entropy
  - “CRT dispersion”

Routines of “what” vs. “when”

Consumers can have routines in both *when* and *what* they buy.

- In ride-sharing: are people routine in *where* people are going, in addition to *when* they’re going?
- Data: full set of locations for each customer, measured by “block”
- Two metrics of **location dispersion**:
  - Shannon entropy
  - “CRT dispersion”

Entropy: given the empirical distribution $\boldsymbol{p}$ of locations, $\ell = 1, \dots, L$:

$$\text{Entropy} = -\sum_{\ell=1}^{L} p_k \log p_k$$

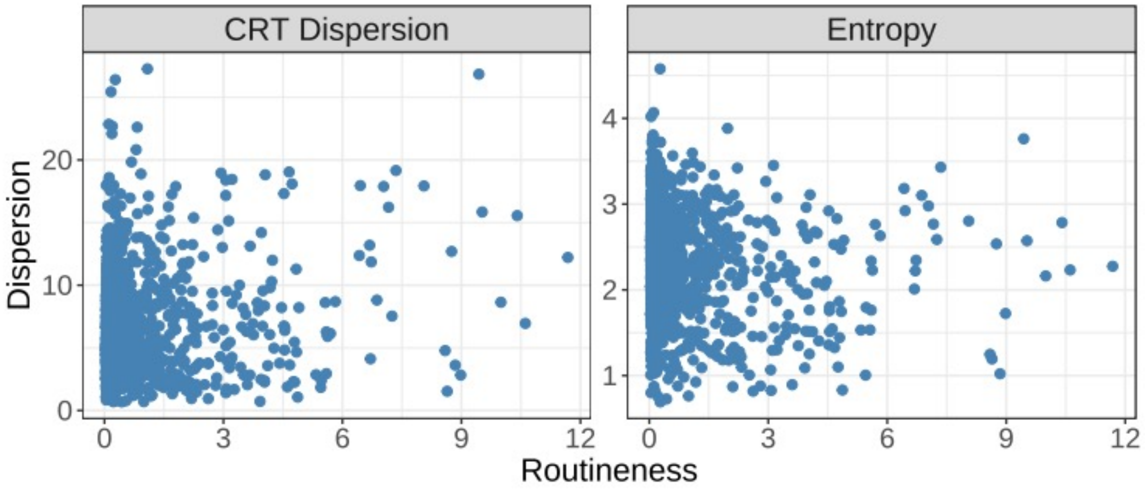
CRT Dispersion: given L unique locations in K total trips:

$$\text{CRT Disp.} = \frac{L}{\log K}$$

Routines of "what" vs. "when"

Consumers can have routines in both *when* and *what* they buy.

- In ride-sharing: are people routine in *where* people are going, in addition to *when* they're going?
- Data: full set of locations for each customer, measured by "block"
- Two metrics of **location dispersion**:
  - Shannon entropy
  - "CRT dispersion"



Routines of "what" vs. "when"

Consumers can have routines in both *when* and *what* they buy.

- In ride-sharing: are people routine in *where* people are going, in addition to *when* they're going?
- Data: full set of locations for each customer, measured by "block"
- Two metrics of **location dispersion**:
  - Shannon entropy
  - "CRT dispersion"

	Dependent variable:	
	# Requests	Activity
	OLS	*logistic*
	(1)	(2)
Requests ($w = 38$)	2.263***	0.381***
	(0.225)	(0.108)
Recency	−0.193***	−0.140***
	(0.042)	(0.010)
Frequency	0.093***	0.0002
	(0.009)	(0.003)
Routine ($w = 38$)	5.593***	1.132***
	(0.453)	(0.396)
Entropy	−1.753	0.080
	(1.092)	(0.240)
CRT Disp.	0.159	−0.011
	(0.199)	(0.046)
Observations	2,000	2,000
R^2	0.568	
Note:	*p<0.1; **p<0.05; ***p<0.01	
	Intercept omitted for clarity.	

Routines of “what” vs. “when”

Consumers can have routines in both *when* and *what* they buy.

- In ride-sharing: are people routine in *where* people are going, in addition to *when* they’re going?
- Data: full set of locations for each customer, measured by “block”
- Two metrics of **location dispersion**:
  - Shannon entropy
  - “CRT dispersion”

	Dependent variable:	
	# Requests	Activity
	OLS	*logistic*
	(1)	(2)
Requests ($w = 38$)	2.263***	0.381***
	(0.225)	(0.108)
Recency	−0.193***	−0.140***
	(0.042)	(0.010)
Frequency	0.093***	0.0002
	(0.009)	(0.003)
Routine ($w = 38$)	5.593***	1.132***
	(0.453)	(0.396)
Entropy	−1.753	0.080
	(1.092)	(0.240)
CRT Disp.	0.159	−0.011
	(0.199)	(0.046)
Observations	2,000	2,000
R^2	0.568	
Note:	*p<0.1; **p<0.05; ***p<0.01	
	Intercept omitted for clarity.	

[Back]