At what time?

Lucas Javaudin[1], André de Palma[1], Nathalie Picard[2]

[1]THEMA, CY Cergy Paris Université
[2]BETA, Strasbourg University

June 14, 2023
ITEA 2023

Departure-time choice

- When traveling, individuals face a **trade-off** between travel cost and schedule cost.
- **Travel cost**: function of the travel time and value of time.
- **Schedule cost**: function of the departure and / or arrival time.
- Example trade-off:
    1. Leave at 8:00, have a 30-minute trip and arrive on-time at the appointment.
    2. Leave at 7:50, have a 20-minute trip and arrive 15 minutes early at the appointment.

Schedule cost

Since Vickrey (1969) and Arnott et al. (1990), schedule cost is often represented with a linear penalty for early and late arrivals (**α-β-γ model**):

$$SC(t_d, t_a) = \beta \cdot [t^* - t_a]^+ + \gamma \cdot [t_a - t^*]^+,$$

- t^*: **desired arrival time** at destination
- β: penalty for early arrivals (\$ / h)
- γ: penalty for late arrivals (\$ / h)

Desired arrival time = time at which the individual would choose to arrive if travel time was null (can be different from work start time).

Why departure-time choice matters?

- Road opening → Decreased congestion during peak period → Shift of some individuals from off-peak to peak period → Increased congestion during peak period (rebound effect)
- Knowledge of the schedule-cost function is required to predict the extend of the shift from off-peak to peak period (e.g., when running transport simulations).

Literature review

Studies relying on a **travel survey** which includes the **work start time**:

- Small (1982): Multinomial Logit model; α-β-γ model.
- Thorauge et al. (2021): Latent Class Choice model; α-β-γ model with travel-time variability.

Studies using **time-specific constants** (with cyclical functions):

- Zeid et al. (2006); Popuri et al. (2008): Multinomial Logit model.
- Lemp and Kockelman (2010); Lemp et al. (2010): Continuous Logit model estimated with Bayesian estimations.

This paper

Goal: Estimate a departure-time choice model without knowledge of the t^* (desired arrival-time) distribution.

Part I:

- We estimate the $\mathbf{t^*}$ **distribution** from the arrival-time distribution of individuals with a constant travel time.
- We identify the demographic variables which explain t^* for the **home-work commute**.

Part II:

- **Continuous Logit model** to estimate α, β and γ, using the t^* distribution estimated in Part I.

Summary of results

Part I:

- The t^* distribution depends on **profession category** and **destination area**.
- On average, t^* is earlier for *blue-collar workers* than for other profession categories.

Part II:

- β is similar for all profession categories.
- γ is larger for high-qualification jobs (upper and intermediate category) than for low-qualification jobs (blue-collar workers and employees).

Enquête Globale Transport (EGT)

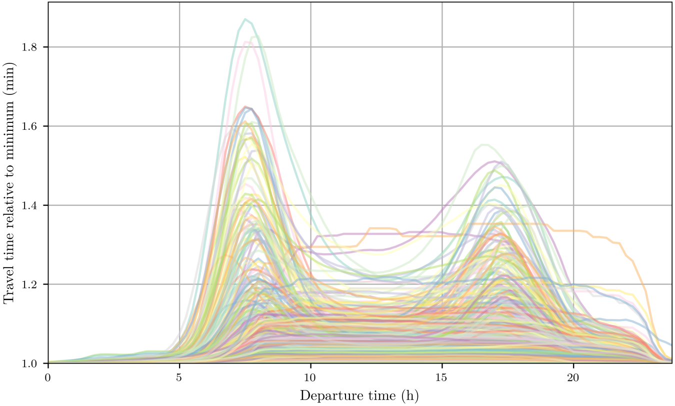
- 2010 **transport survey** for Île-de-France (Paris' region, with 12 millions inhabitants)
- 14 855 households, 35 175 individuals surveyed
- Observations: households characteristics, individual characteristics, trips of the previous day (including, mode, departure time, purpose)

Travel-Time Data

- Source: HERE, Q1 2016
- Historical link-level speed for 15-minute intervals (typical day)
- 977 618 links in the Île-de-France area (18.51 % with a non-constant travel time)
- OD-level travel-time functions computed using a routing algorithm (Time-dependent Contraction Hierarchies)
- Link-level and OD-level travel time functions are piecewise linear functions

Travel-Time Data

Random sample of 500 OD pairs.

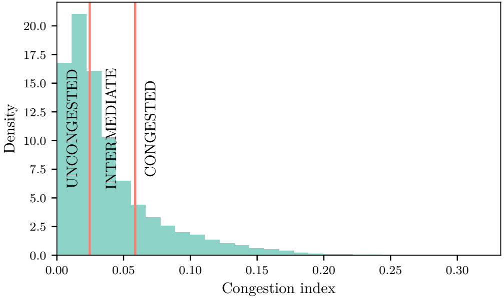


Congestion Index

For each OD pair, we have a travel-time function defined by breakpoints $\{(td_i, tt_i)\}_i$. We compute a congestion index as

$$c = \sigma_{tt}/tt_0,$$

where $\sigma_{tt} = \sqrt{(1/n)\sum_i (tt_i - \bar{tt})^2}$ is the standard-deviation of the travel times and $tt_0 = \min_i tt_i$ is the minimum travel time.



Car trips are split in three categories of equal size based on the congestion index (uncongested, intermediate and congested).

Scope and goal

Scope:

- Home-work trips
- Trips by car (as a driver alone) or walk
- Trips contained in the time window 5AM to 11AM
- Sample size: 4212 trips (3540 by car and 672 by walk)

Goal: Estimate the t^* distribution for trips to work with any mode.

Basic Principle

Claim: When travel-time function is constant, the individual arrives at his / her t^*.

$$C(t_a) = \alpha \cdot tt(t_a) + \beta \cdot [t^* - t_a]^+ + \gamma \cdot [t_a - t^*]^+$$

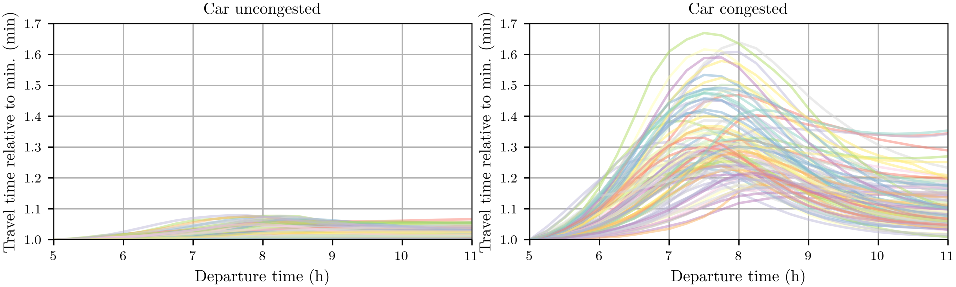
$$\text{if } tt(t_a) = \bar{tt} \quad \Rightarrow \quad \hat{t}_a = \arg\min_{t_a} C(t_a) = t^*$$

Consequence: The arrival-time distribution is equal to the t^* distribution *for individuals facing no congestion.*

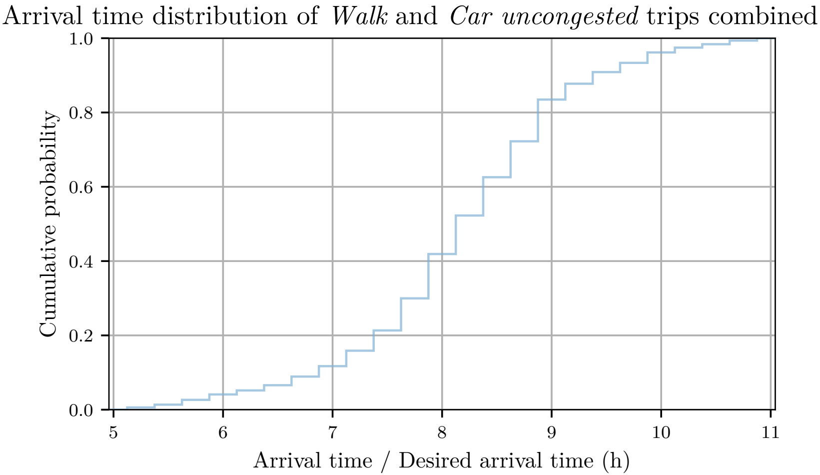
Trip categories

Three trip categories are analyzed:

- Walk: 672 trips
- Car uncongested (congestion index $\leq 2.46\,\%$): 1180 trips
- Car congested (congestion index $> 5.87\,\%$): 1180 trips



Estimated t^* distribution



Representativeness

The t^* distribution of *Walk* trips might be different from the t^* distribution of *Car uncongested* trips:

- Different mode chosen $\Rightarrow$ different demographic characteristics
- Different origin / destination $\Rightarrow$ different workplace

If the arrival-time distribution for *Walk* and *Car uncongested* trips are similar, then we can assume that the t^* distribution does not depend on the mode chosen and thus that **the $\mathbf{t}^*$ distribution of uncongested trips is representative of the $\mathbf{t}^*$ distribution for the whole population**.

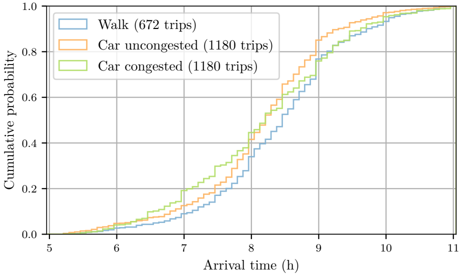
Comparing arrival-time distribution

- The **two-sample Kolmogorov-Smirnov test** can be used to compare two samples and assert if they come from the same probability distribution.
- Null hypothesis: “The values in the two samples are drawn from the same probability distribution”.

All trips, by mode

The null hypothesis is always rejected at the 1 % level (the distributions are different).

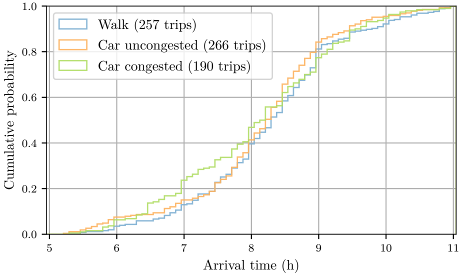
	KS statistic	p-value
Walk / Car uncong.	**0.1389**	**0.0000**
Walk / Car cong.	0.1302	0.0000
Car cong. / Car uncong.	0.0932	0.0001



Employees

The null hypothesis that *Walk* and *Car uncongested* have the same distribution **cannot** be rejected.

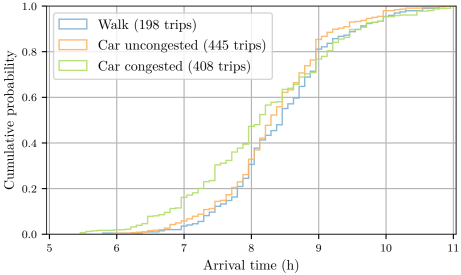
	KS statistic	p-value
Walk / Car uncong.	**0.0850**	**0.2791**
Walk / Car cong.	0.1242	0.0614
Car cong. / Car uncong.	0.1188	0.0806



Intermediate category

The null hypothesis that *Walk* and *Car uncongested* have the same distribution **cannot** be rejected.

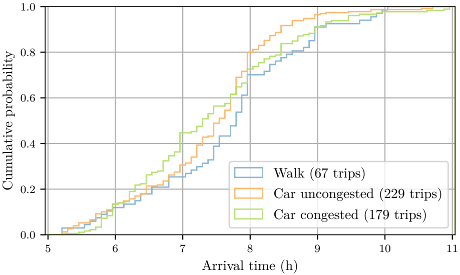
	KS statistic	p-value
Walk / Car uncong.	**0.0849**	**0.2587**
Walk / Car cong.	0.1987	0.0000
Car cong. / Car uncong.	0.1630	0.0000



Blue-Collar Workers

The null hypothesis that *Walk* and *Car uncongested* have the same distribution **cannot** be rejected.

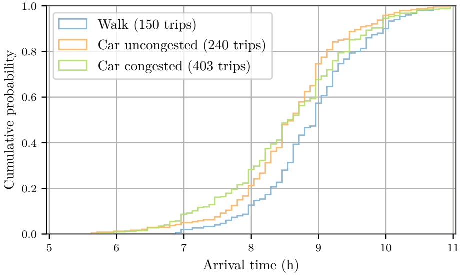
	KS statistic	p-value
Walk / Car uncong.	**0.1526**	**0.1585**
Walk / Car cong.	0.2024	0.0309
Car cong. / Car uncong.	0.1413	0.0320



Upper category

The null hypothesis that *Walk* and *Car uncongested* have the same distribution **can** be rejected.

	KS statistic	p-value
Walk / Car uncong.	**0.2058**	**0.0007**
Walk / Car cong.	0.2130	0.0001
Car cong. / Car uncong.	0.0887	0.1737



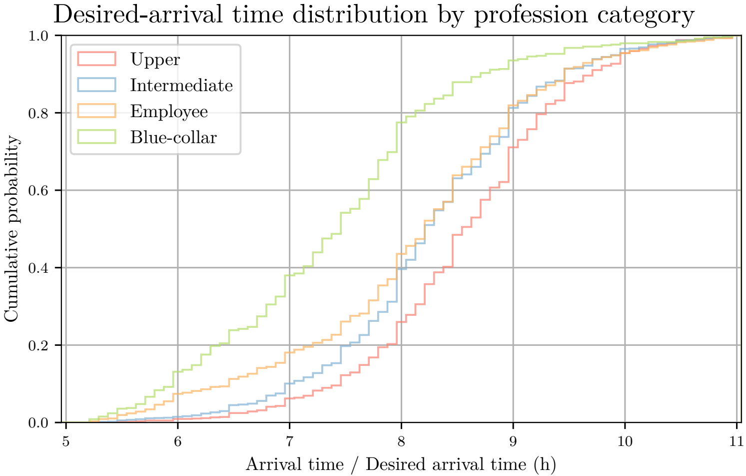
Upper category: By destination area

Average arrival time (trip count)

	Paris	Inner suburbs	Outer suburbs
Walk	9:10 (70)	8:50 (51)	8:38 (29)
Car uncongested	9:26 (3)	8:46 (51)	8:34 (185)
Car congested	8:45 (51)	8:32 (222)	8:38 (127)
Average	8:56	8:37	8:34

Ideally, t^* distributions should be split by profession category and destination area but sample size is too small.

Summary



Desired-arrival time distribution by profession category
Upper
Intermediate
Employee
Blue-collar
Cumulative probability
0.0
0.2
0.4
0.6
0.8
1.0
5
6
7
8
9
10
11
Arrival time / Desired arrival time (h)

Continuous Logit

- Utility of individual n for departure time t is

$$U_n(t) = V_n(t) + \varepsilon_n(t)$$

where

$$V_n(t) = \alpha \cdot tt_n(t) + \beta \cdot [t_n^* - t - tt_n(t)]^+ + \gamma \cdot [t + tt_n(t) - t_n^*]^+$$

and $\varepsilon_n(t)$ are i.i.d. extreme-value distributed.

- The probability to choose a departure time in interval $[t_j, t_{j+1})$ is

$$P(y_n \in [t_j, t_{j+1})) = \frac{\int_{t_j}^{t_{j+1}} e^{g(x_n(t), \cdot)} \, \mathrm{d}\, t}{\int_{t_0}^{t_1} e^{g(x_n(t), \cdot)} \, \mathrm{d}\, t}$$

Assumptions

- One estimation for each **profession category**: employee, intermediate category, blue-collar worker and upper category.
- **Only car trips.**
- **Mixture model:** $\{\alpha, \beta, \gamma\}$ are fixed coefficients, t_n^* is individual-specific.
- t_n^* has **categorical distribution** with probabilities given in the previous section.
- **Bayesian estimations** are used to estimate the models.

Results

Employee / Observations: 677

	Point estimate	90 % Confidence interval
α	11.45	[1.45, 22.25]
β	5.00	[3.88, 6.33]
γ	6.10	[4.64, 7.93]

Blue-collar worker / Observations: 587

	Point estimate	90 % Confidence interval
α	14.45	[2.57, 24.57]
β	5.24	[3.72, 7.03]
γ	5.54	[4.24, 7.18]

Intermediate category / Observations: 1294

	Point estimate	90 % Confidence interval
α	22.32	[13.45, 30.72]
β	6.01	[4.98, 7.24]
γ	7.09	[5.81, 8.59]

Upper category / Observations: 982

	Point estimate	90 % Confidence interval
α	16.38	[4.32, 28.03]
β	5.77	[4.65, 7.09]
γ	8.90	[6.90, 11.19]

Note: α: value of time; β: early penalty; γ: late penalty.

Takeaways

We estimate the $\mathbf{t}^*$ **distribution** using *arrival-time* distribution of individuals facing no congestion.

- The t^* distribution depends on the *profession category* and *destination area*.
- On average, t^* is earlier for *blue-collar worker* than for other profession categories.

We estimate α, β and γ using **Bayesian estimations** and a **Continuous Logit model**.

- β is similar for all profession categories (between 5 and 6).
- γ is larger for high-qualification jobs (9 for upper category, 7 for intermediate category) than for low-qualification jobs (5.5 for blue-collar workers, 5 for employees).
- α cannot be estimated accurately.

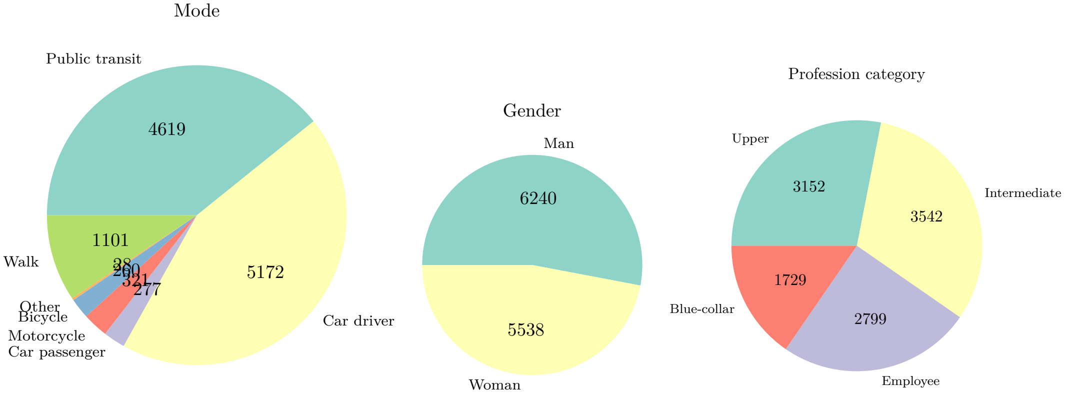
Future works

- **Evening commute** (desired departure time from origin)
- **Trip chaining** (with t^* at intermediate stop and at destination)
- Day-to-day **travel-time variability**
- Integrated **mode** and **departure-time choice**

Thank you

Slides available at `lucasjavaudin.com`

Characteristics of home-work trips



Travel-Time Data

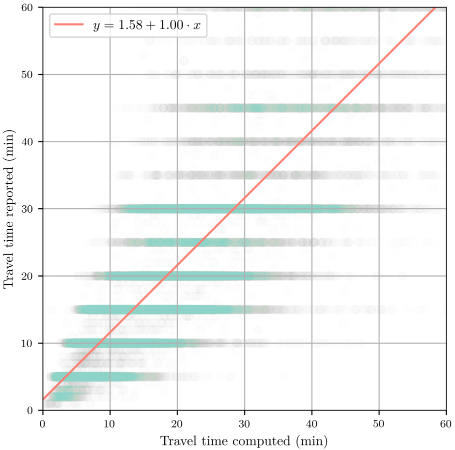
Reported travel time in the travel survey can be well predicted by the computed travel time with HERE data ($R^2 = 66\,\%$).

Dep. Variable:	EGT tt	**R-squared:**	0.664
Model:	OLS	**F-statistic:**	8.274e+04
No. Observations:	41934	**Prob (F-statistic):**	0.00

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
cst	1.5758	0.087	18.061	0.000	1.405	1.747
HERE tt	1.0003	0.003	287.650	0.000	0.993	1.007

Note: Travel-time penalties at intersections are calibrated to reach a slope close to 1.

Travel-Time Data



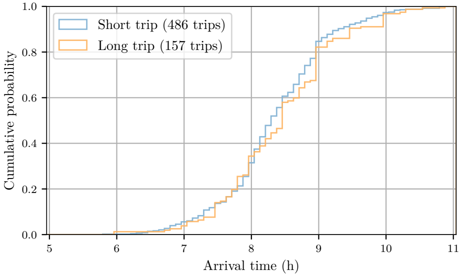
Robustness check: travel time

Comparing long / short trips (intermediate category; walk and car uncongested only).

Long trip: Travel time is longer than 30 minutes.

The null hypothesis that *Short trip* and *Long trip* have the same distribution **cannot** be rejected.

	KS statistic	p-value
Short / Long	0.0923	0.2453



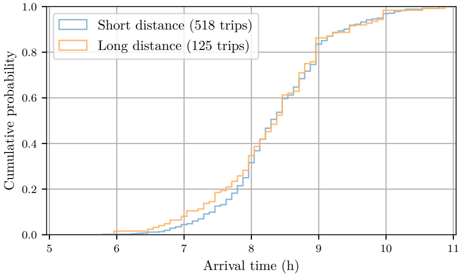
Robustness check: distance

Comparing long / short distance trips (intermediate category; walk and car uncongested only).

Long distance: Euclidian distance between origin and destination is greater than 10 kilometers.

The null hypothesis that *Short distance* and *Long distance* have the same distribution **cannot** be rejected.

	KS statistic	p-value
Short / Long	0.0588	0.8531

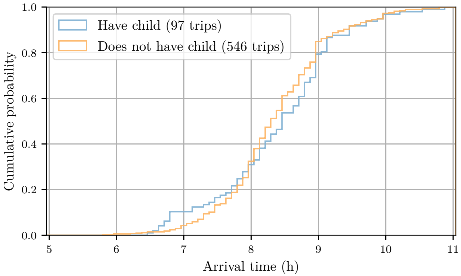


Robustness check: children

Comparing trips of people with / without child (intermediate category; walk and car uncongested only).

The null hypothesis that *Male* and *Female* have the same distribution **cannot** be rejected.

	KS statistic	p-value
Child / No child	0.0914	0.4662

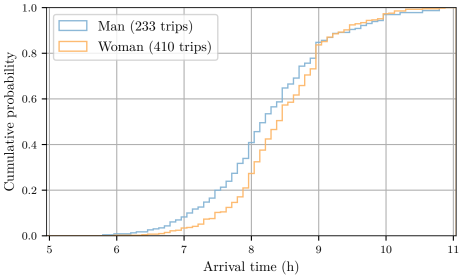


Robustness check: gender

Comparing trips of men / women (intermediate category; walk and car uncongested only).

The null hypothesis that *Man* and *Woman* have the same distribution **can** be rejected.

	KS statistic	p-value
Man / Woman	0.1469	0.0029



Bayesian Estimations

- Values are drawn from the posterior distribution using **Gibbs sampling**:
  1. Draw $(t_n^*)^{\tau+1}$, $\forall n$ given $\{\alpha^\tau, \beta^\tau, \gamma^\tau\} \rightarrow$ Metropolis-Hastings algorithm

     $$K(t_n^* | \{\alpha, \beta, \gamma\}; y_n) \propto L(y_n | \{\alpha; \beta; \gamma\}; t_n^*) f(t_n^* | \theta), \quad \forall n$$

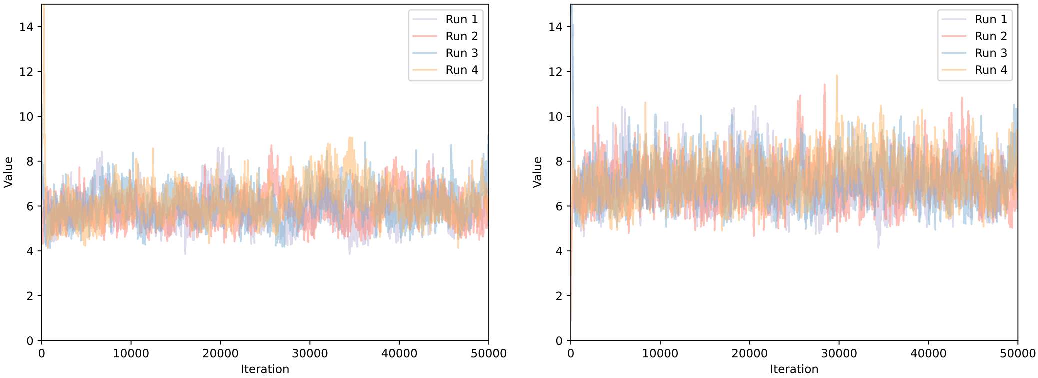
  2. Draw $\{\alpha^{\tau+1}, \beta^{\tau+1}, \gamma^{\tau+1}\}$ given $(t_n^*)^{\tau+1} \rightarrow$ Metropolis-Hastings algorithm

     $$K(\{\alpha, \beta, \gamma\} | t_n^*, \forall n; Y) \propto \prod_n L(y_n | \{\alpha, \beta, \gamma\}; t_n^*)$$

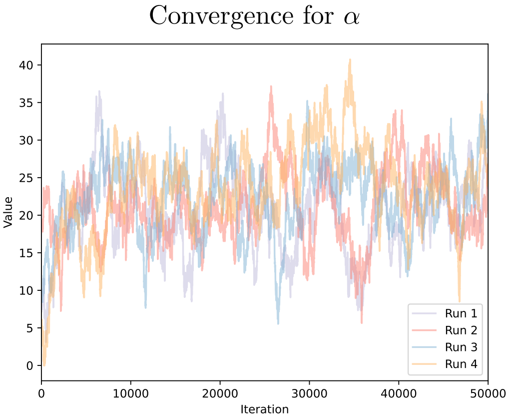
- We run 4 simulations with different initial conditions. Each simulation consists in 50 000 iterations of Gibbs sampling.

Results: Intermediate category

Convergence for β and γ



Results: Intermediate category



Results: $\beta < \alpha < \gamma$ inequality

Contrarily to most other studies, we find $\gamma < \alpha$.

- The desired arrival time might not be equal to the starting time of work
- We do not consider (day-to-day) travel-time variability

Departure-Time Probability Comparison

Example individuals with $t^* = 8$AM and constant travel time of 30 minutes, for all profession categories.