The 18th summer course of Behavior Modeling
Final presentation

マルコフ決定過程に基づく経路選択行動のパラメータ推定
―自動車・自転車交通施策の検討―

Evaluation of car/bicycle traffic measures with a link choice model

University of Tokyo team A

Takuya Iizuka (M1)

Kenta Ishii (M1)

Shoma Dehara (M1)

Miho Yonezawa (M1)

1. Background

◆ Area: **松山市** Matsuyama city

Population: 512479 (2018.1.1.)

Area: 429.06 m^2

- Many people use private car.
- City projects are underway to increase activity in the central city.

http://udcm.jp/project/

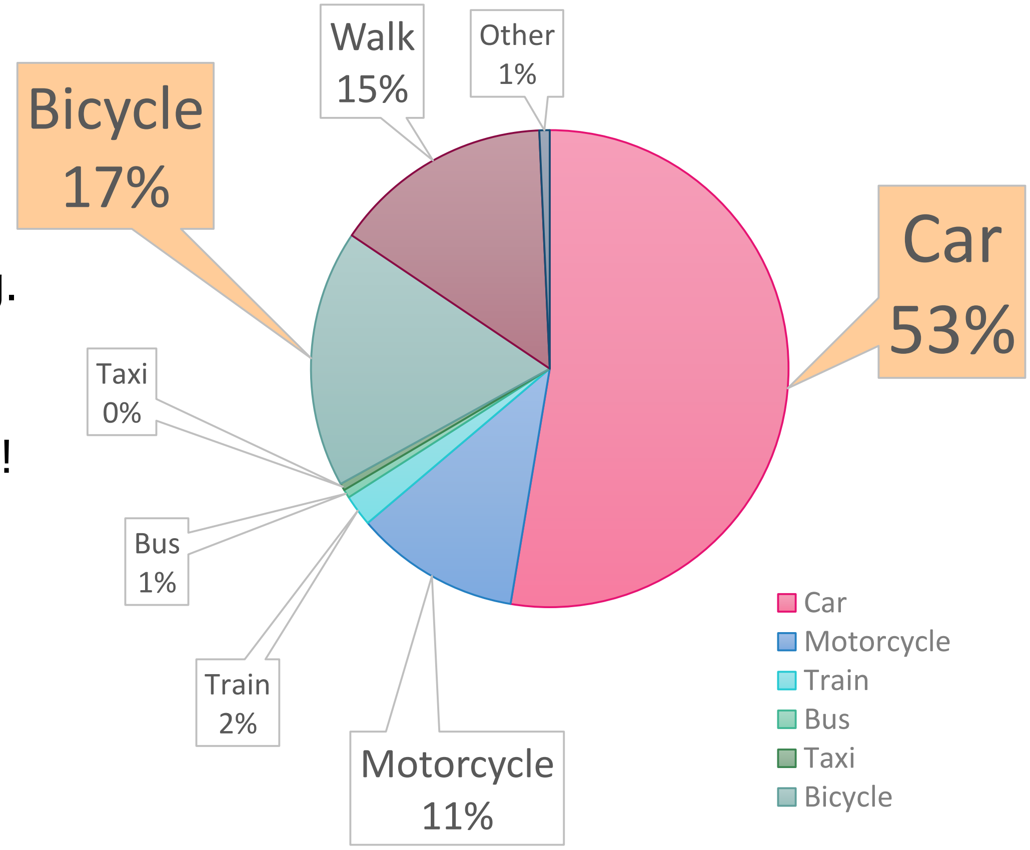
2. Basic Analysis

◆ Mode Choice

- Data: Matsuyama PP (2007 Feb.19 – Mar.23)
- High rate of Car & Bicycle use
- Car & Bicycle paths are overlapping.

→By providing bicycle lanes, traffic accidents can be suppressed !!

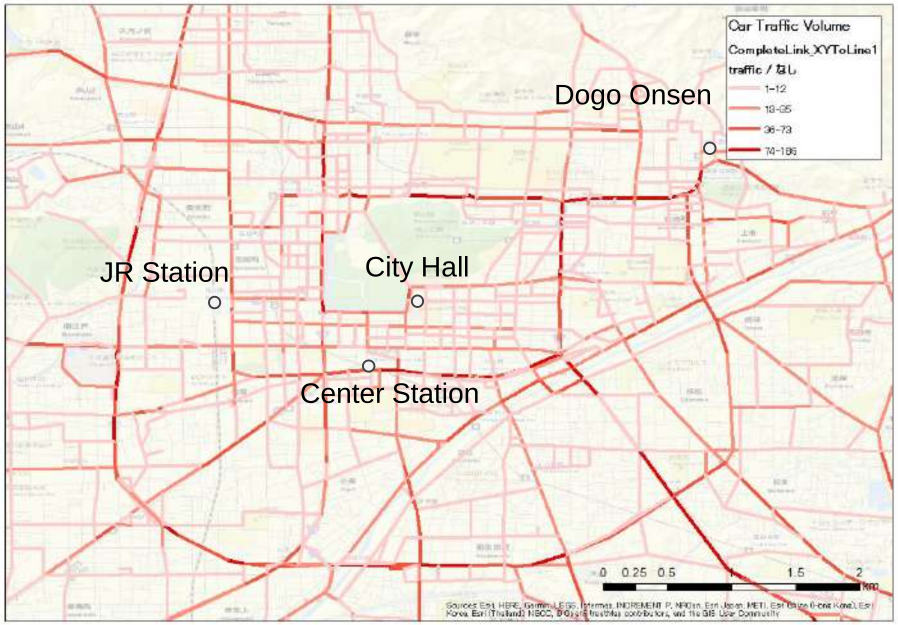
Representative Mode Choice in Matsuyama (n=7107)



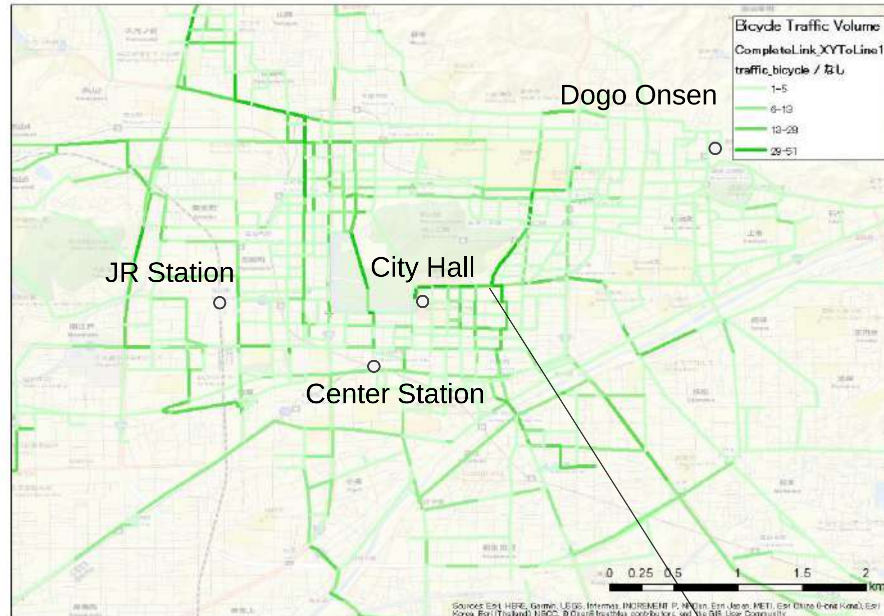
※経路情報が得られたトリップを抽出

2. Basic Analysis

◆ Traffic Volume in the center of Matsuyama

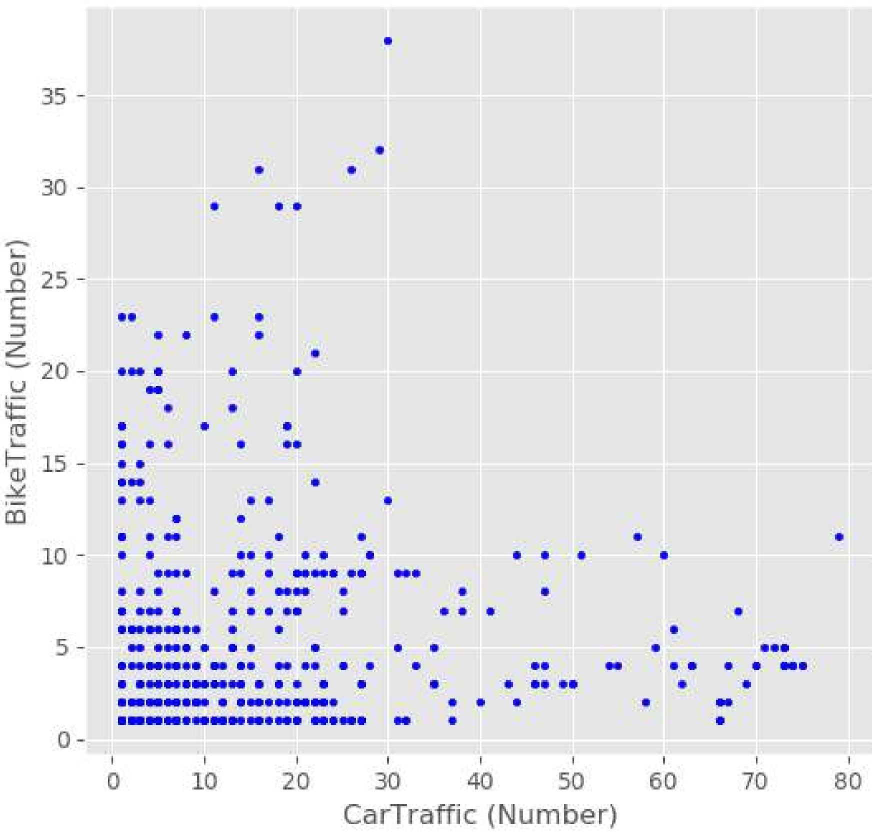


Car Trip



Bicycle Trip

- Most part of the center of Matsuyama, the car & bicycle trips are separated.
- At some roads, car & bicycle trips are overlapping!!



Car & Bicycle traffic of each link

The smaller the traffic of the car,
the more traffic of the bicycle.

On links with heavy car traffic,
sidewalks are maintained,
increasing bicycle traffic.

3. Target

◆ For Simulation

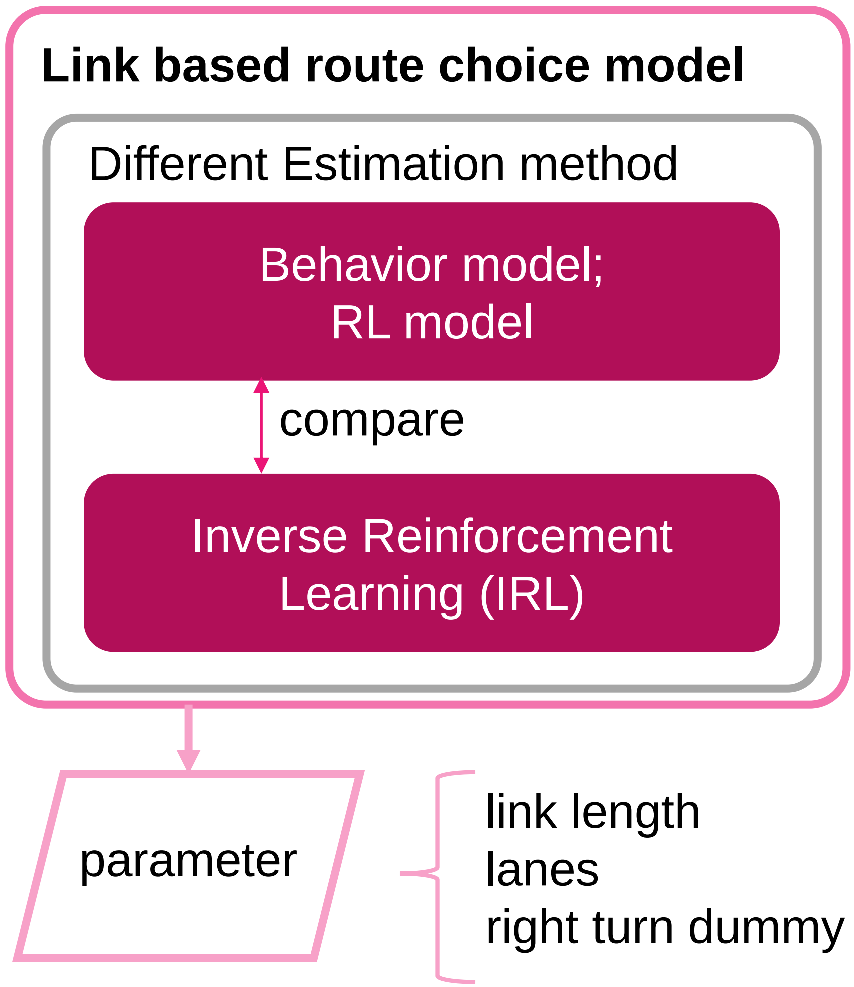
- Characteristics of each link (length, width, etc.) affect travelers' behavior.

→ We adopt Link Base Route Choice Model for analysis.

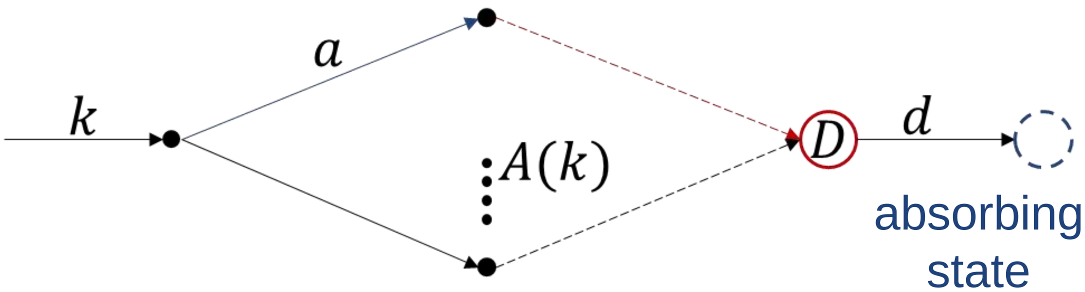
◆ Our Goal

- To clarify what is important element in the route choice behavior of car & bicycle
- To simulate transport policy and to verify the sensitivity of each parameter

◆ Estimation


Link based route choice model
Different Estimation method
Behavior model;
RL model
compare
Inverse Reinforcement
Learning (IRL)
parameter
link length
lanes
right turn dummy

◆ Sequential Route Choice Model: **Recursive Logit model (RL)** (Fosgerau et al., 2013)



Graph: $G = (A, \nu)$

A: set of links
ν: set of nodes

- Utility Maximization problem

$$v_n(a|k) + \mu\varepsilon_n(a) + \beta V_n^d(a)$$

An instantaneous utility

At each current state k, a traveler chooses an action a (next link).

$\varepsilon_n(a)$: error term (i.i.d. Gumbel distribution)
μ: scale parameter
β: discount rate

An expected downstream utility :value function

from the selected state a to the destination link d

The value function is defined by the **Bellman equation** (Bellman, 1957);

$$V_n^d(k) = E\left[\max_{a \in A(k)}\left(v_n(a|k) + \mu\varepsilon_n(a) + \beta V_n^d(a)\right)\right]$$
$\forall k \in A$

Link choice probability

$$P_n^d(a|k) = \frac{e^{\frac{1}{\mu}(v_n(a|k)+\beta V_n^d(a))}}{\sum_{a' \in A(k)} e^{\frac{1}{\mu}(v_n(a'|k)+\beta V_n^d(a'))}}$$

4. Compared IRL with RL

◆ Bellman equation

$$V^{\pi}(s) = E_{\pi}\left\{\sum_{k=0}^{\infty} \gamma^{k} r_{t+k+1} | s_t = s\right\}$$

$$= E_{\pi}\left\{r_{t+1} + \gamma \sum_{k=0}^{\infty} \gamma^{k} r_{t+k+2} | s_t = s\right\}$$

$$= \sum_{a} \pi(s,a) \sum_{s'} \mathcal{P}_{ss'}^{a} \left[\mathcal{R}_{ss'}^{a} + \gamma E_{\pi}\left\{\sum_{k=0}^{\infty} \gamma^{k} r_{t+k+2} | s_{t+1} = s'\right\}\right]$$

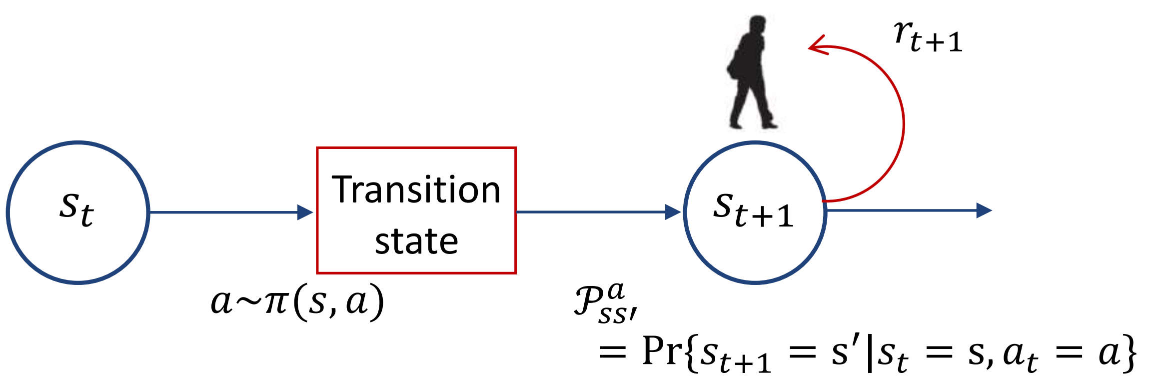
$$= Q(s,a)$$

$$= \sum_{a} \pi(s,a) \sum_{s'} \mathcal{P}_{ss'}^{a} [\mathcal{R}_{ss'}^{a} + \gamma V^{\pi}(s')]$$

γ: discount rate ($0 < \gamma \leq 1$)

$\mathcal{R}_{ss'}^{a}$: expected reward

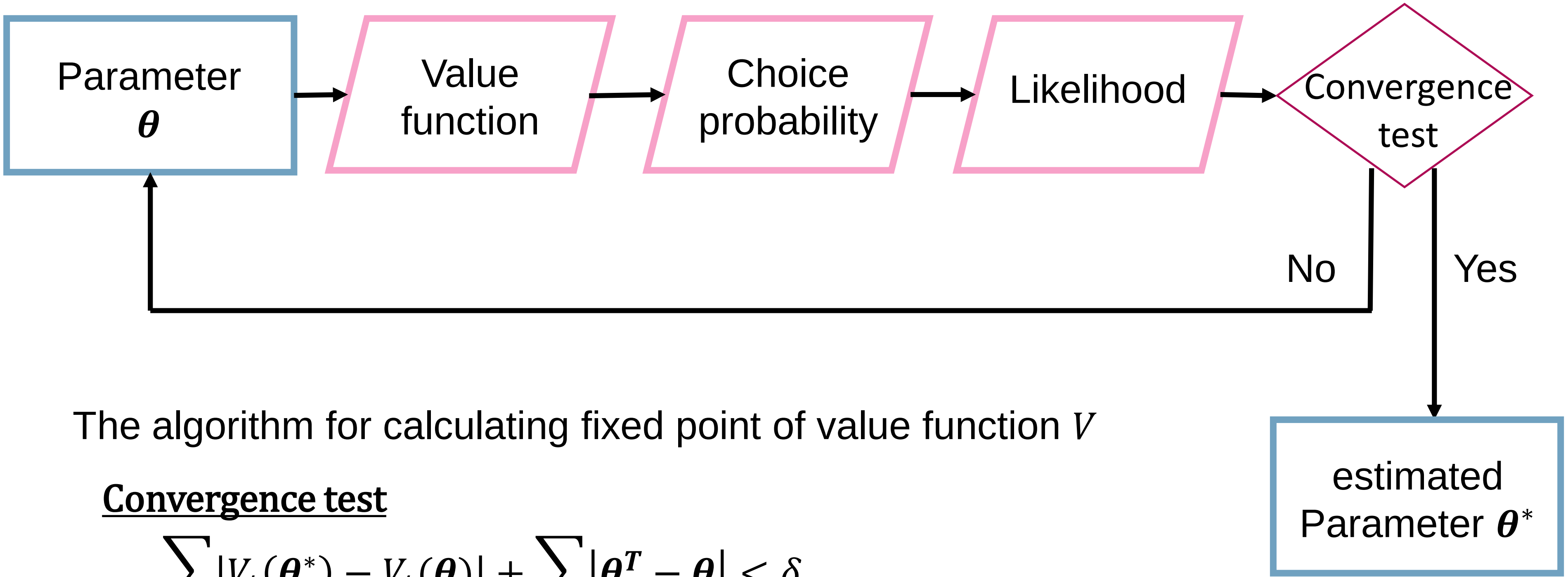
($= E\{r_{t+1} | s_t = s, a_t = a, s_{t+1} = s'\}$)



4. Compared IRL with RL

◆ The estimation method : **Recursive Logit model (RL) -NPL**

Reward (Instantaneous utility): $r_t = \boldsymbol{\theta}^T \boldsymbol{X}$



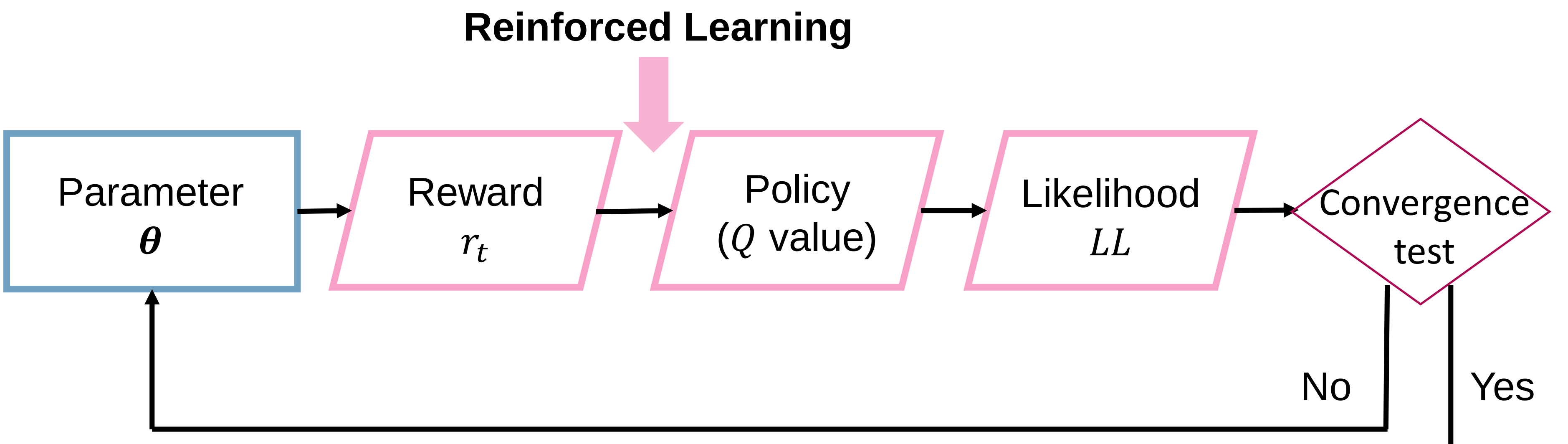
The algorithm for calculating fixed point of value function V

Convergence test

$$\sum_t |V_t(\boldsymbol{\theta}^*) - V_t(\boldsymbol{\theta})| + \sum_t |\boldsymbol{\theta}^T - \boldsymbol{\theta}| < \delta$$

◆ The estimation method : **Max entropy - Inversed Reinforced Learning (IRL)**

Reward: $r_t = \boldsymbol{\theta}^T \boldsymbol{X}$



Problem

$$\max_{\boldsymbol{\theta}} \sum_i \log P(\boldsymbol{\zeta}_i | \boldsymbol{\theta}) \quad st.\ Q_t = r_t + \gamma Q_{t+1}$$

where $\boldsymbol{\zeta}_i$ is the path of expert

$\boldsymbol{X}$ is the feature relating to link

5. Estimation Result

◆ RL estimation (car)

$\beta = 0.47$ (given)

Variables	Parameters	t-Value
Link Length	-0.03	-1.33
Right-Turn	-0.80	-6.49**
Lanes	0.37	2.76**

L(0)	-1179.29
LL	-1147.00
Rho-Square	0.03
Adjusted Rho-Square	0.02

◆ IRL estimation (car)

$\beta = 0.47$ (given)

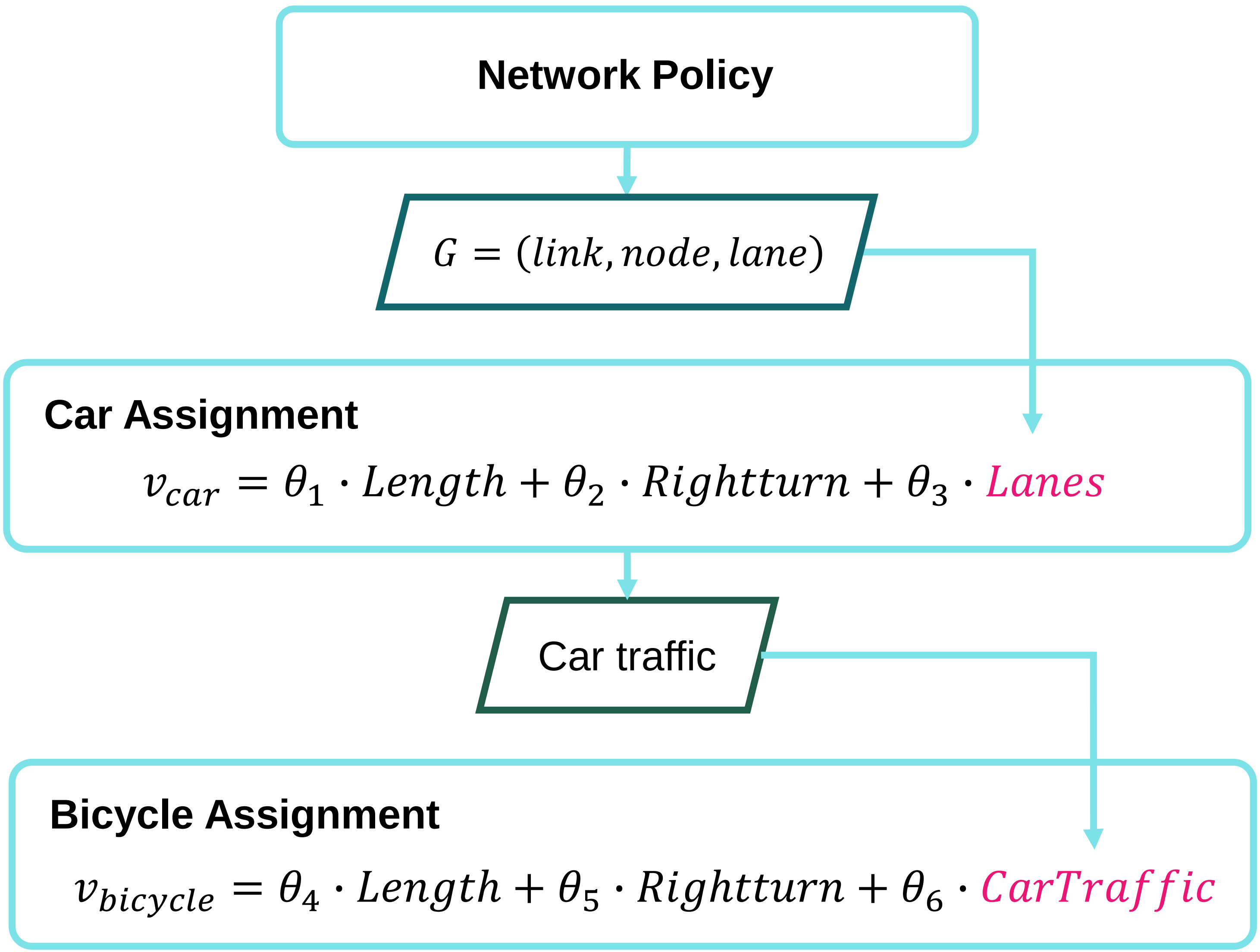
Variables	Parameters	t-Value
Link Length	-0.07	-9.72**
Right-Turn	-1.02	-8.53**
Lanes	-0.37	-5.64**

L(0)	-2080.67
LL	-1117.10
Rho-Square	0.46
Adjusted Rho-Square	0.46

◆ Recursive Logit estimation (bicycle)

Variables	Parameters	t-Value
Link Length	-0.00	-6.21**
Right-Turn	-0.19	-3.67**
Car Traffic	-14.37	-0.14
β	0.00	15.15**

L(0)	-4093.90
LL	-3861.56
Rho-Square	0.06
Adjusted Rho-Square	0.06


Network Policy
$G = (link, node, lane)$
Car Assignment
$v_{car} = \theta_1 \cdot Length + \theta_2 \cdot Rightturn + \theta_3 \cdot Lanes$
Car traffic
Bicycle Assignment
$v_{bicycle} = \theta_4 \cdot Length + \theta_5 \cdot Rightturn + \theta_6 \cdot CarTraffic$

←Bicycle traffic

Policy

Reduce the lanes of large bicycle traffic links

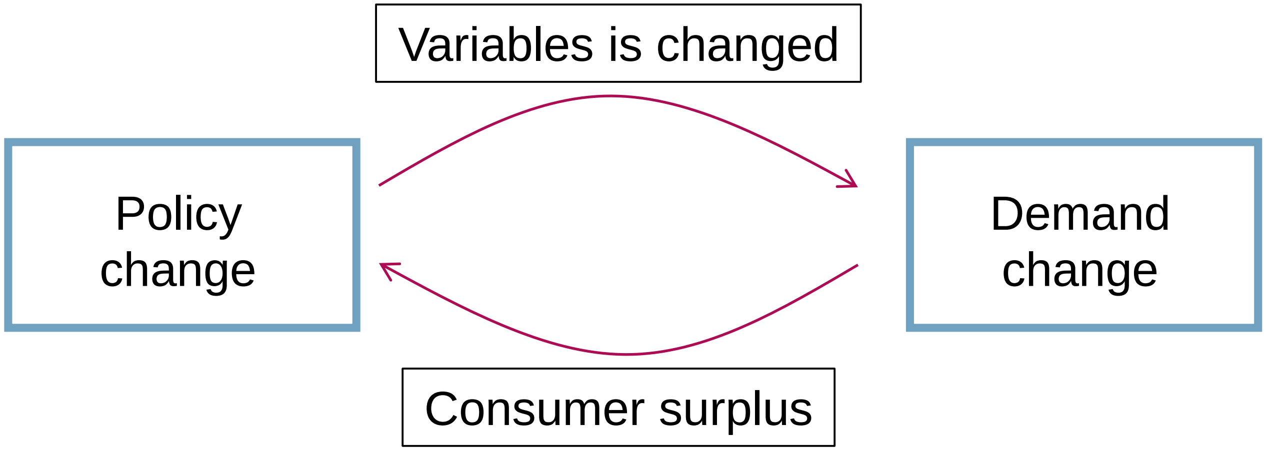
Private car/bicycle user's **logsum** value with/without policy

	Without policy	With policy (rode lanes are reduced)
Private car user	-2639	-2638
Bicycle user	-9297	**-1147** Increased!!

6. Future works

◆ Policies decided by Two-stage optimization

To decide the policy
by calculating the fixed point of demand of cars and bicycles



4. Frame & Model

◆ Estimation

Link based route choice model

Different Estimation method

Behavior model;
RL model

compare

Inverse Reinforcement
Learning (IRL)

parameter

- link length
- lanes
- right turn dummy

◆ Policy Simulation

Upper Problem: traffic network

- reduction of vehicle lanes (pedestrian/bicycle only)

traffic volume of each link

network

Lower Problem: route choice behavior

Car

Bicycle

Assign each OD volume