Advances in Choice Modeling and Asian Perspectives

Toshiyuki Yamamoto, Nagoya Univ.
Tetsuro Hyodo, Tokyo Univ. of Marine Sci. & Tech.
Yasunori Muromachi, Tokyo Inst. of Tech.

Outline

- Recent developments in econometric choice modeling
- Characteristics of transport modeling in Asian cities
- Inaccuracy of transport demand models

Outline

- Recent developments in econometric choice modeling
- Characteristics of transport modeling in Asian cities
- Inaccuracy of transport demand models

Recent developments in econometric choice modeling

- GEV (generalized extreme value) model
- MMNL (mixed multinomial logit) model
- VTTS (value of travel time saving)
- Discrete-continuous model

GEV model：Basic

- Has flexible error correlations by relaxing IIA property of MNL model
  - MMNL model also has the same flexible structure
- Maintains a closed form in representing choice probability, thus are free from numerical integrations
  - Numerical integrations, vulnerable to simulation error, are adopted by MMNL model
- Only a few members have been exploited
  - The appropriate types of GEV models should be selected or created

GEV model：Extension

- CNL model is reformulated as a generalization of the two-levels hierarchical logit model, and shown to reproduce any hypothetical homoscedastic covariance matrix（Papola, 2004）

- GNL model is extended to include covariance heterogeneity and heteroscedasticity of the observations（Koppelman & Sethi, 2005）

- An operationally easy way of generating new GEV models are proposed by using RNEV (recursive nested extreme value) model and the network structure of the correlation of the error terms（Daly & Bierlaire, 2006）

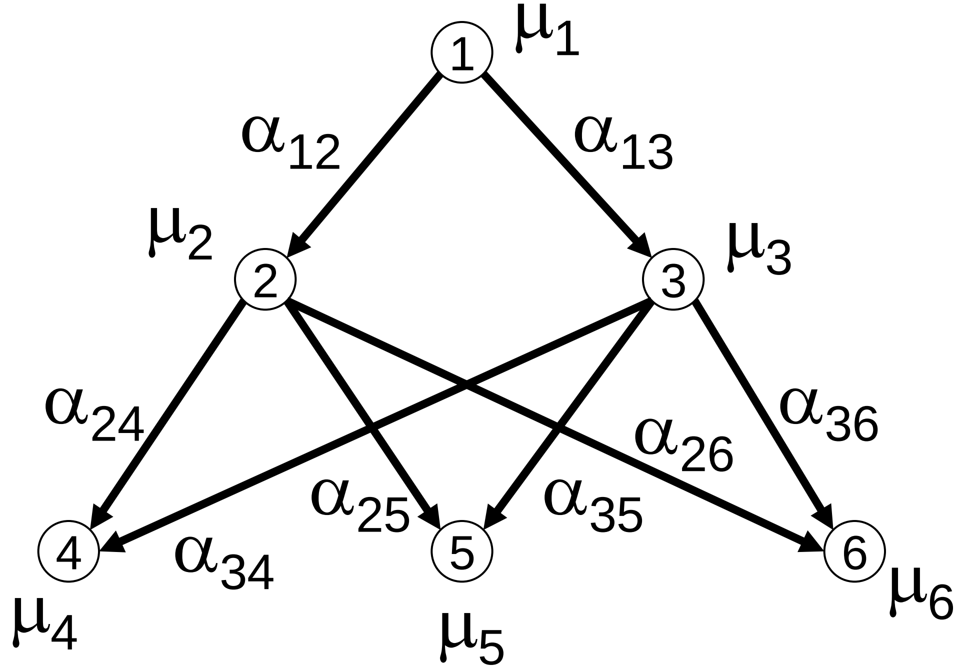
GEV model: Extension

RNEV + network GEV

$$P_j(i) = \sum_{k \in S_j} \Omega_{jk} P_k(i),$$

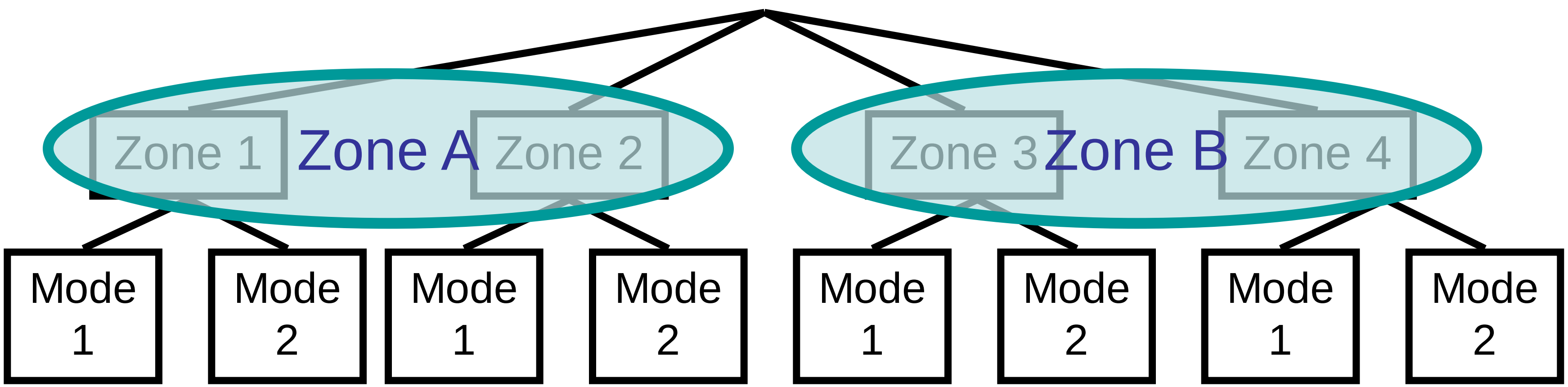
where

$$\Omega_{jk} = \frac{\alpha_{jk} \exp(\mu_j V_k)}{\sum_{l \in S_j} \alpha_{jl} \exp(\mu_j V_l)}$$



GEV model：New properties

- A set of rules allowing the consistent aggregation of alternatives is derived for NL model of joint choice of destination and travel mode（Ivanova, 2005）



GEV model：New properties

- With choice-based samples, ESML estimator is shown to give consistent estimates of parameters except alternative specific constants even in NL model（Garrow & Koppelman, 2005）
  – WESML estimator is consistent, but not asymptotically efficient

- Both studies extend the well-known properties of ML model to NL model

MMNL model: Basic

- Incorporates error components to ML model
  - Represents any types of correlations among alternatives
  - Represents taste heterogeneity

$$P_{ni} = \int \left(\frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}} \right) f(\beta)\, d\beta.$$

- Choice probability does not maintain closed form, so numerical integration is required. Simulation techniques are applied

$$\check{P}_{ni} = \frac{1}{R} \sum_{r=1}^{R} L_{ni}(\beta^r)$$

MMNL model：Basic

Simulation techniques:

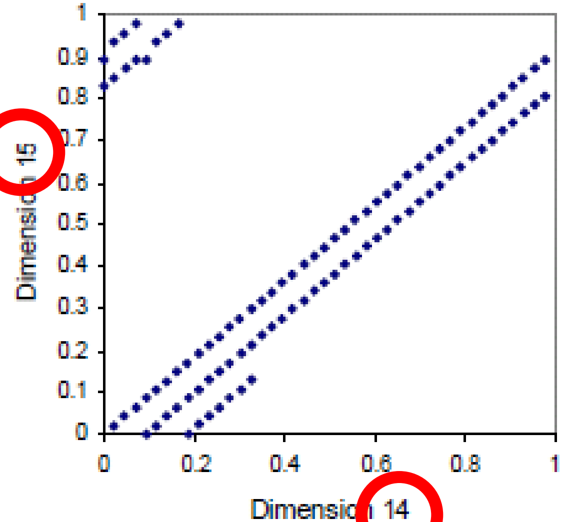
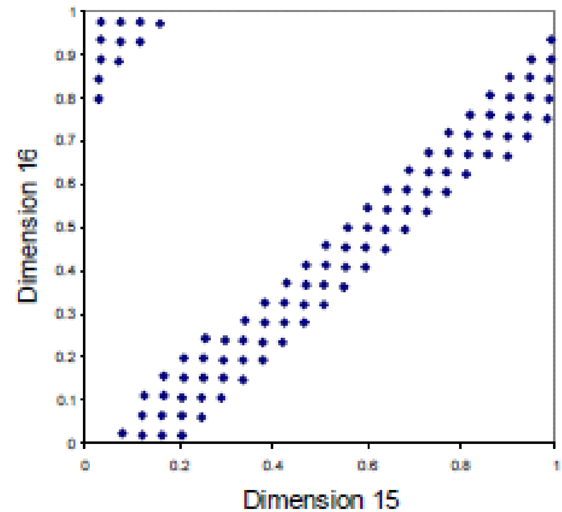
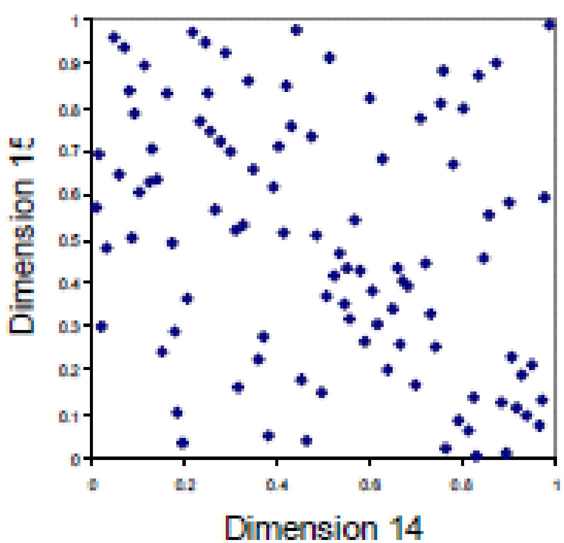
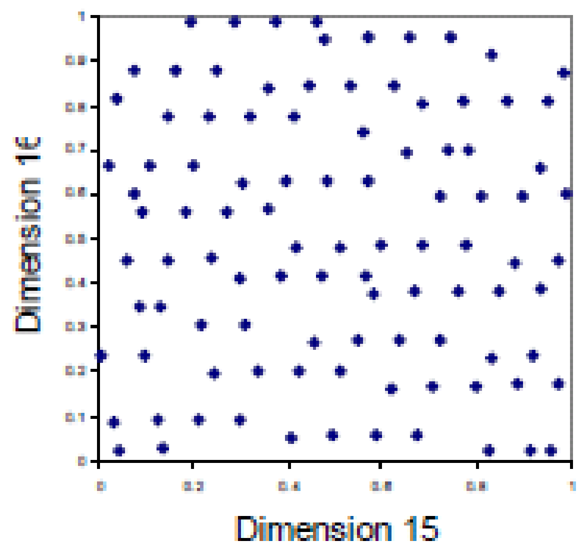
- Pseudo-random sequence
  - Independent random draws: deterministic pseudo-random sequence is used in computer
- Quasi-random sequence
  - Non random sequence to provide better coverage than independent draws
- Hybrid method
  - Quasi-random sequence with randomization (scramble, shuffle, etc.)

MMNL model：Efficient numerical integration

- (t, m, s)-nets is more efficient than Halton sequence（Sándor & Train, 2004）

- Based on the comparison of Halton sequence and Faure sequence (a special case of (t, m, s)-nets), their scrambled versions and LHS, scrambled Faure sequence is the most efficient（Sivakumar, et al., 2005）

- MLHS (modified Latin hypercube sampling) is more efficient than standard, scrambled and shuffled Halton sequence（Hess, et al., 2006）
  - MLHS is not yet compared with Faure sequence

	Halton sequence	Faure sequence
Standard		
Scrambled		

Sivakumar et al. (2005)

MMNL model：Efficient algorithm

BTRDA (basic trust-region with dynamic accuracy) algorithm

- Variable number of draws in each iteration in the estimation of the choice probabilities, which gives significant gains in the optimization time（Bastin, et al., 2006）
- BTRDA with MLHS performs better than BFGS algorithm with pure pseudo-Monte Calro sequence（Bastin, et al., 2005）

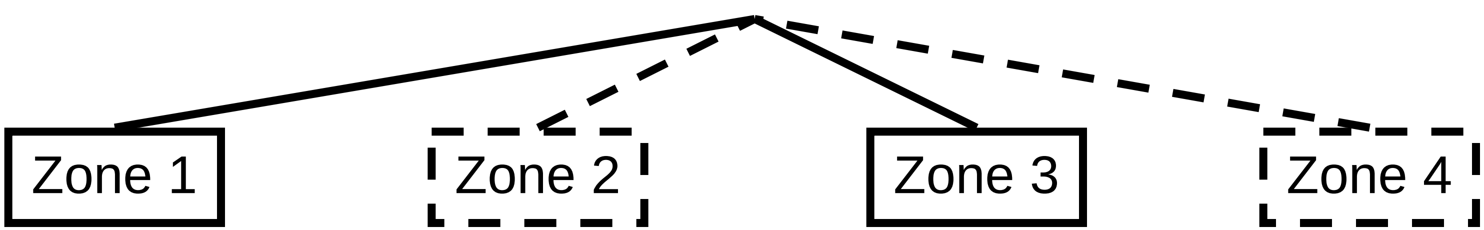
MMNL model：Comparison with MNP

- In the context of panel analysis with fewer than 25 alternatives, MNP model with GHK simulator is sperior to MMNL model with pseudo-random sequence（Srinivasan & Mahmassani, 2005）

- Based on simulation data, both MMNL model with pseudo-random sequence or Halton sequence and MNP model with GHK simulator require 8000 sample cases to recover correlations of error structure adequately（Minizaga & Alzarez-Dazian, 2005）

MMNL model: Sampling of alternatives

- Consistent for MNL model, but it does not hold for MMNL model
- For empirical accuracy, safe to use a fourth to half for MMNL and eighth to fourth for MNL (Nerella & Bhat, 2004)



VTTS：Basic

- Fundamental factor to evaluate the transportation policy measures
- Can be calculated from the estimated discrete choice models by taking the ratio of the time coefficient to the cost coefficient in linear-in-variables utility function

$$V_i = \alpha_i - \beta_c C_i - \beta_t T_i \qquad \frac{\partial V_i / \partial T_i}{\partial V_i / \partial C_i} = \frac{\beta_t}{\beta_c}$$

- Distribution of the time coefficient provides distribution of VTTS

VTTS：Distribution of VTTS

- Usually, MMNL models use normal distribution for random coefficient, but it causes a negative VTTS for a part of individuals
- Several distributions are examined: truncated normal, log-normal, bounded uniform, triangular, Johnson's S_B, etc.

$$c = a + (b - a) \cdot \frac{\exp(\xi)}{1 + \exp(\xi)}, \quad \xi \sim N(\mu, \sigma)$$

- Nonparametric and semiparametric methods are applied to investigate the distribution of VTTS（Fosgerau, 2006）
- Accounting for variance heterogeneity produces better model fits（Greene, et al., 2006）

VTTS：Reliability of SP data

- Based on the literature review，VTTS is underestimated by using SP data （Brownstone & Small, 2005）

- Dimensionality of the stated choice design affects the decision rules, resulting the underestimation of VTTS if the dimensionality is not accounted for （Hensher, 2006）

Discrete-continuous model：Basic

- Choice of continuous amount as well as discrete choice is represented by theoretical models consistent with random utility theory
- Standard discrete-continuous model treats one discrete choice and choice of continuous amount simultaneously
  - Automobile type and VMT, heating type and usage, telephone charge plan and usage, etc.

Discrete-continuous model: Extension

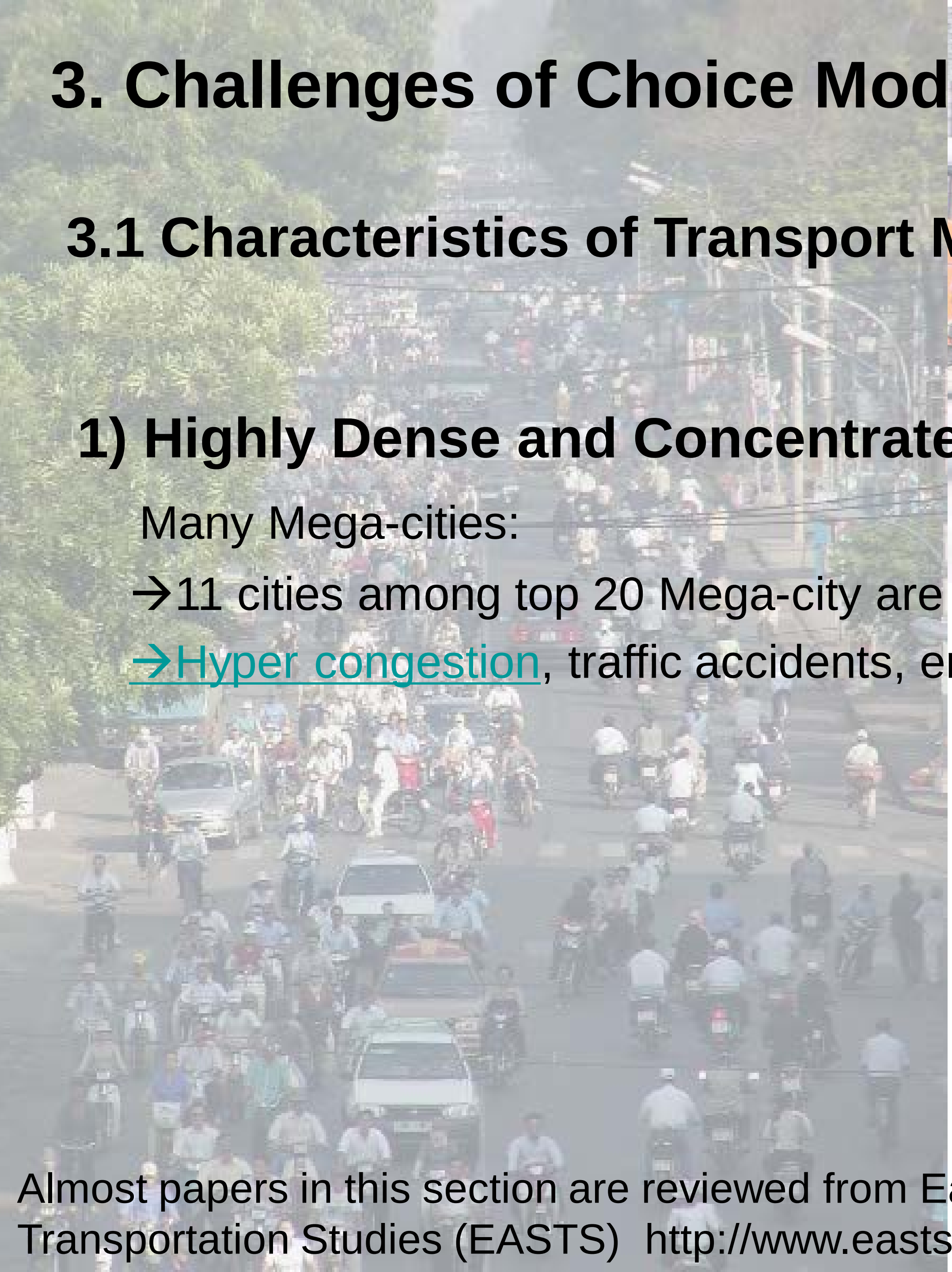
- Discrete-continuous model is extended to incorporate the chioce of multiple alternatives simultaneously
  - Activity types and durations, automobile types of multiple car household and VMTs, etc.

- Bayesian approach with Metropolis-Hasting method is used including unobserved heterogeneity among individuals by Kim, et al. (2002). GHK simulator is used for multivariate normal integral

- Gumbel distribution is applied, and scrambled Halton sequence is used for heteroscedasticity and error correlation across alternative utilities by Bhat (2005)

Outline

- Recent developments in econometric choice modeling
- Characteristics of transport modeling in Asian cities
- Inaccuracy of transport demand models



3. Challenges of Choice Modeling in Asia

3.1 Characteristics of Transport Modeling in Asian Cities

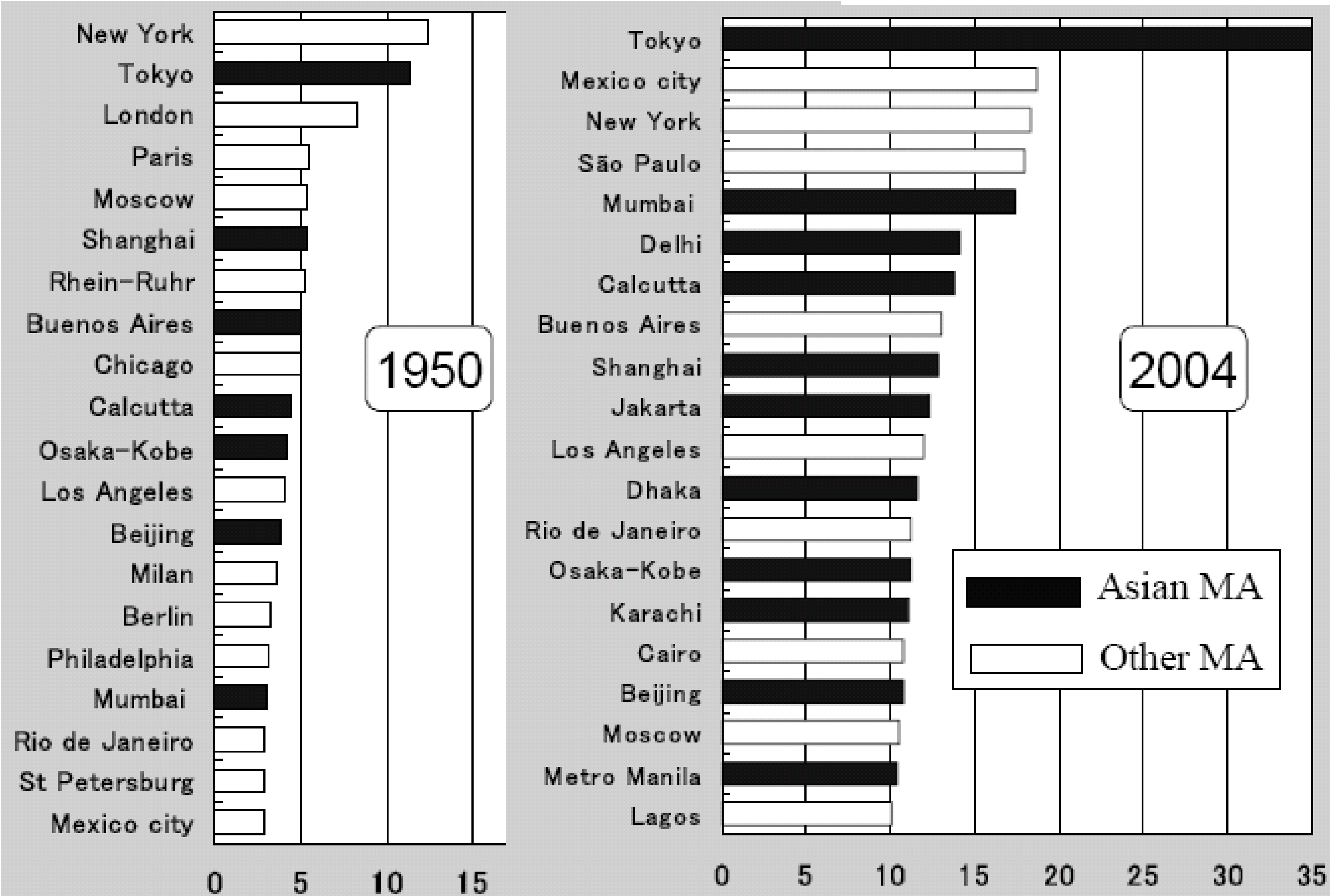
1) Highly Dense and Concentrated Population

Many Mega-cities:

→11 cities among top 20 Mega-city are in Asia in 2015

→Hyper congestion, traffic accidents, environmental issues…

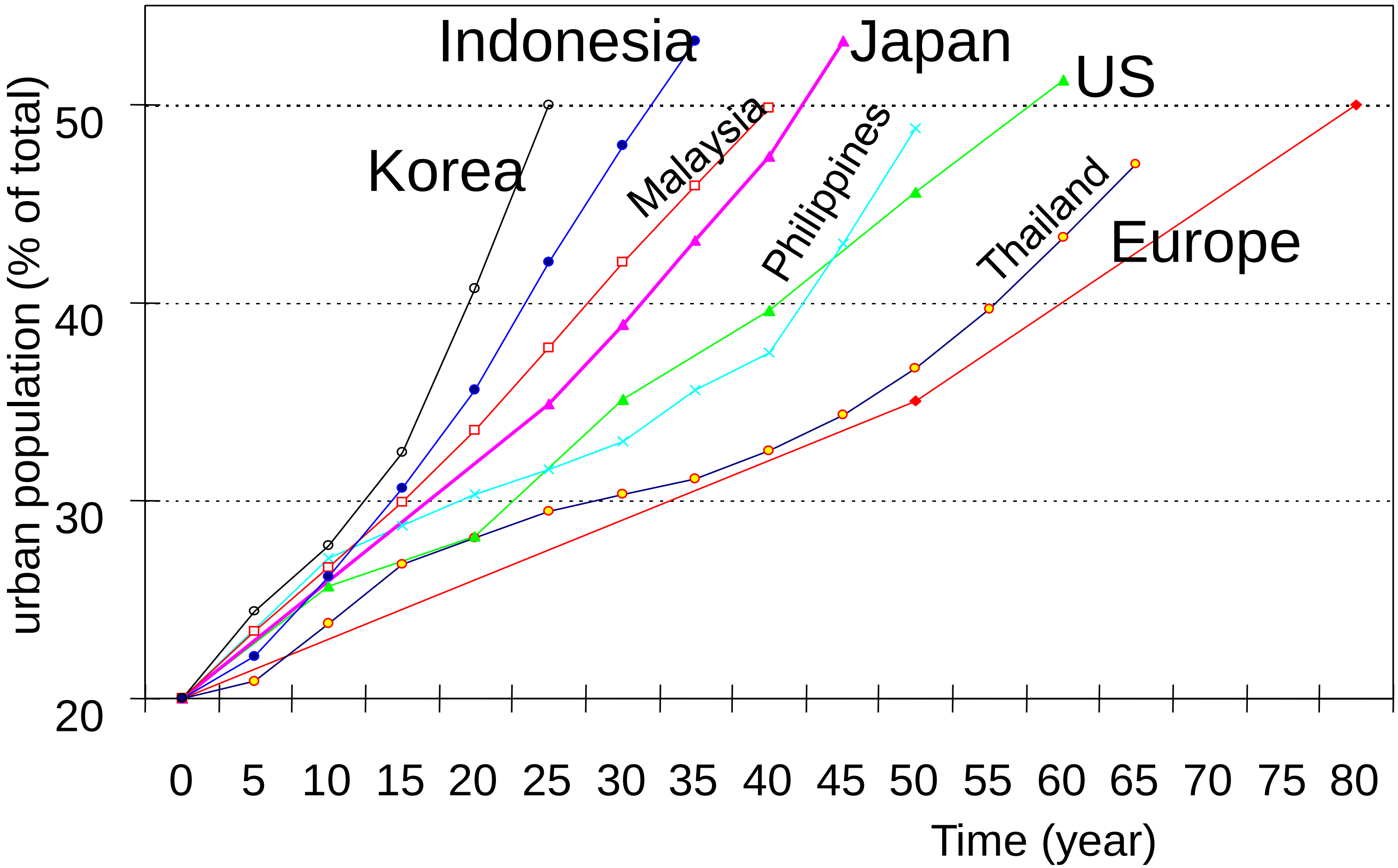
Almost papers in this section are reviewed from Eastern Asia Society for Transportation Studies (EASTS) http://www.easts.info/index.html



 Population in World’s 20 Largest Metropolitan Areas (Morichi, 2005)

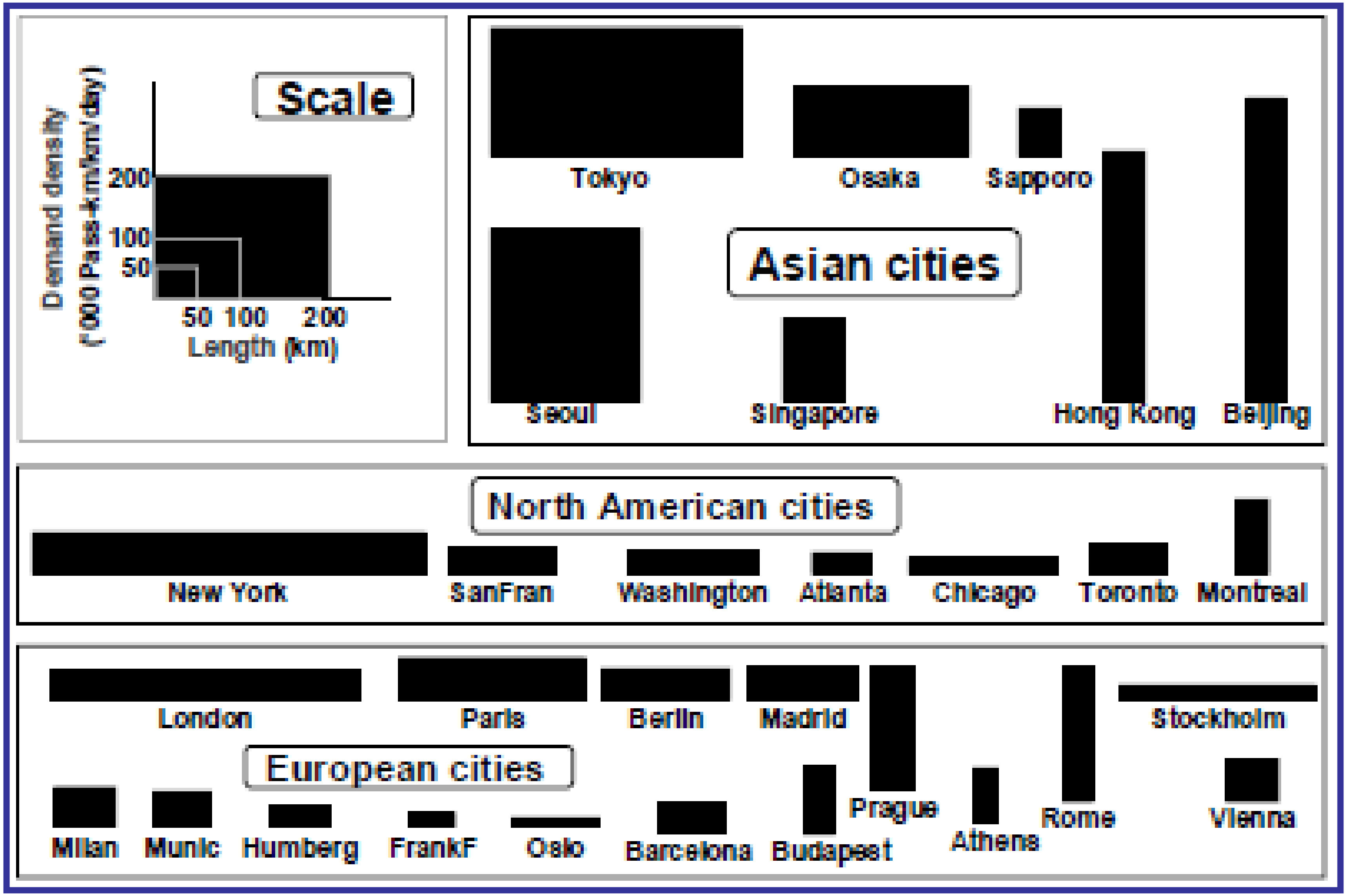
Rapid Urbanization in Asia

Speed of Urbanization: Years taken for 20 % to 50 %



Years from 20 to 50 % Urbanization : Europe (80 yr), US (60 yr), Korea (25 yr), Indonesia (32 yr), Japan (42 yr)

Morichi (2005)



Network Length and Demand Density of Subways (Morichi, 2005)

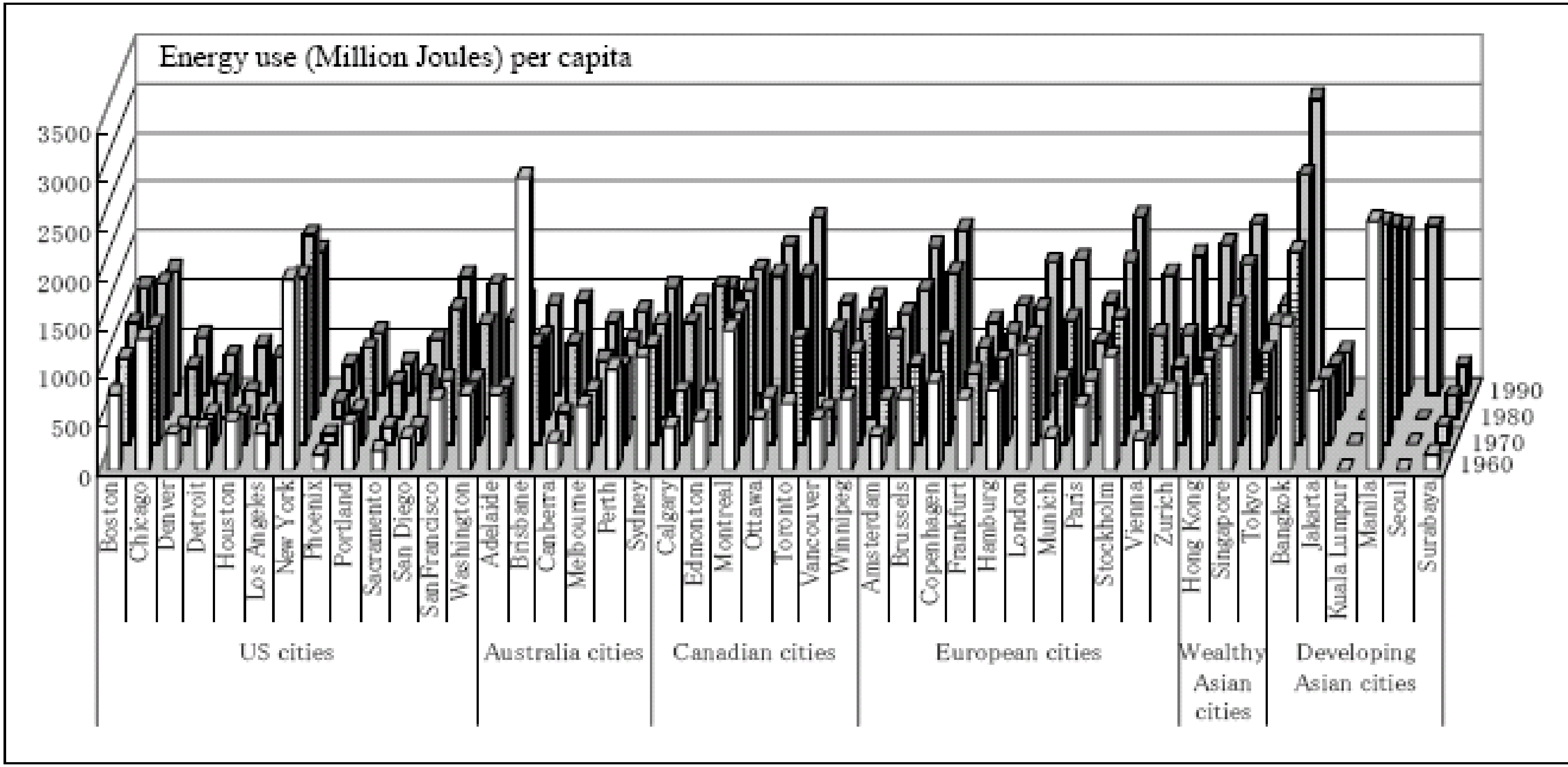


Figure 10. Energy use by public transport systems in developed and developing cities

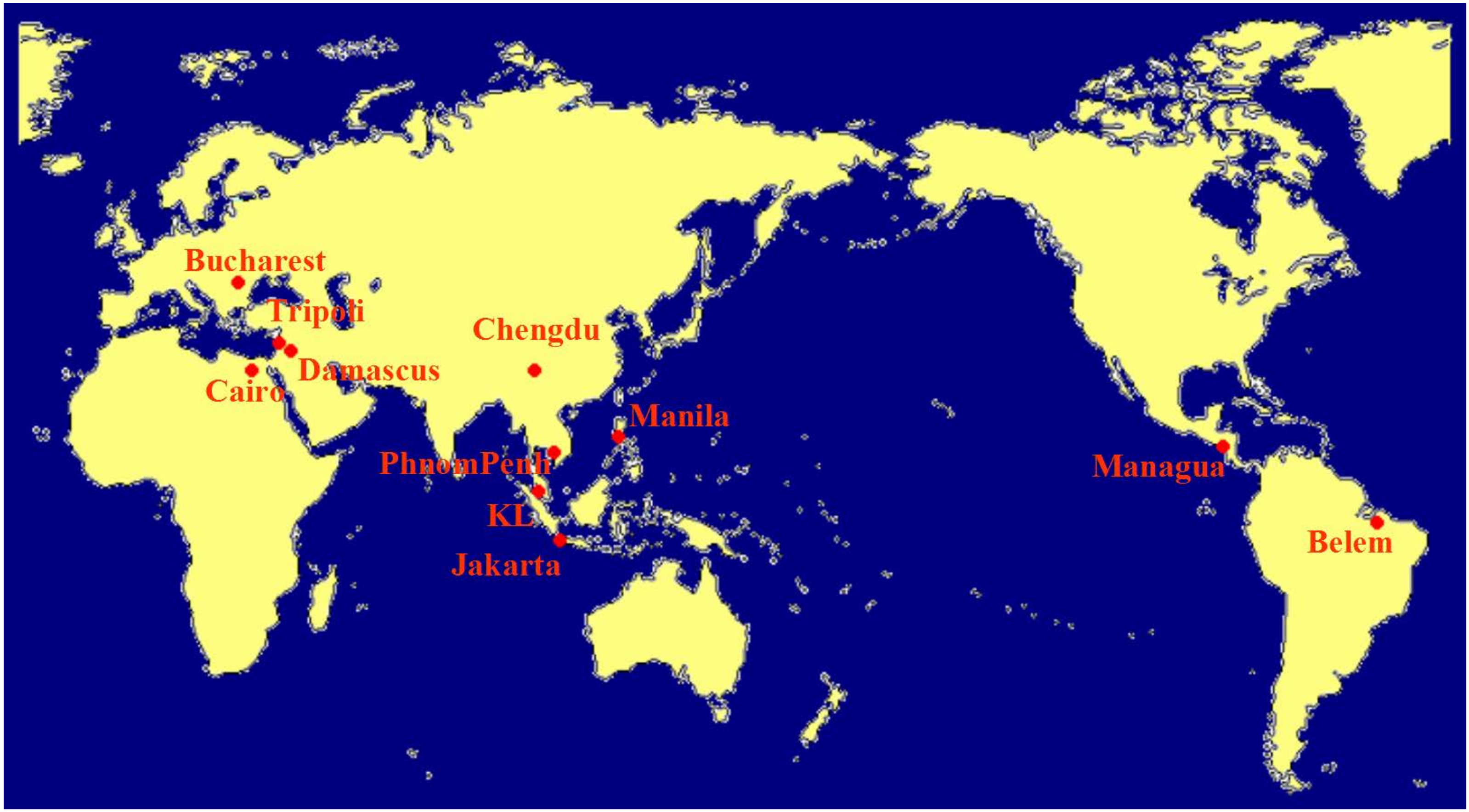
Fujiwara et al.(2005) provides interesting comparative results by "Kenworthy data"

2) Diversity of Transportation Modes

JICA (Japan International Cooperation Agency) summarized the past household interview surveys (HIS) in 11 developing countries

→ They are opened for academic researches

Hyodo et al. (2005) introduced the aggregation results



■Various Mode

01Tripoli
1Passenger Car
2Taxi / Service
3Light Bus / Pass. Van
4Pick-up / Cargo Van
5Truck 2-Axle
6Truck 3-Axle
7Truck 4-Axle or more
8Large Bus
9Bicycle / Motorcycle
0Walking
00Others

06Managua
1Walk
2Car
3Truck(small)
4Truck
5-
6Taxi
7-
8Micro bus
9Bus
10Motor cycle
11Bicycle
12Other

03Damascus
1Walking
2Bicycle and Motorcycle
3Passenger Car
4Taxi
5Microbus
6Bus
7Truck
8Others

07Belem
1Bus
2Micro Bus
3Alternative
4Car Driver
5Car Ride
6Taxi
7Rented Bus
8School Bus
9Motor Bike
10Cicro Motor
11Bike
12By Foot
13Boat
14Truck
15Other

04Manila
1Walking
2Pedicab
3Bicycle
4Motorcycle
5Tricycle
6Jeepney
7Mini-bus
8Standard Bus
9Taxi
10HOV Taxi
11Car/Jeep
12School/Co./Tourist Bus
13Utility Vehicle
14Truck
15Trailer
16LRT
17PNR
18Water Transport
19Others

09Cairo
1On-Foot
2Bicycle
3Motorcycle
4Private Car Driver
5Private Car Passengers
6Pickup for Passengers
7Taxi
8Shared Taxi
9Public Minibus
10Public Bus
11Public A/C Bus
12Cooperative Minibus
13Company (Work) Car
14Factory/Company Bus
15School Bus
16Truck for Passengers
17Nile Bus
18Tram
19Heliopolis Metro
20Underground Metro
21ENR Train
22Animal Drawn
23Other
99No Answer

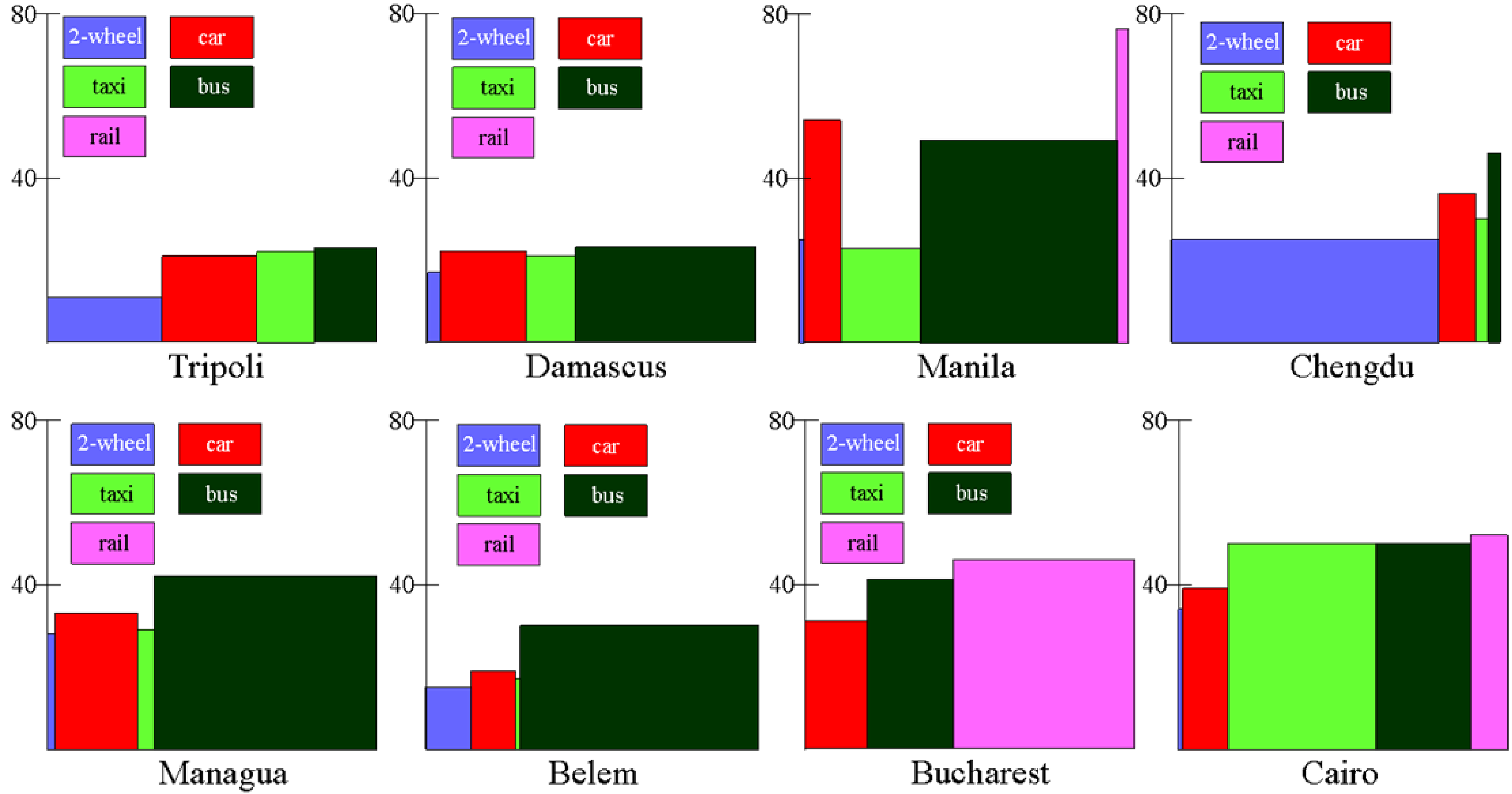
05Chengdu
1Walking
2Bicycle
3Tricycle by man
4Motorcycle
5Tri-motorcycle
6Taxi
7Passenger Car
8Middle Car
9Large Car
10Light Truck
11Large Truck
12Large Bus
13Middle Bus
14Rail

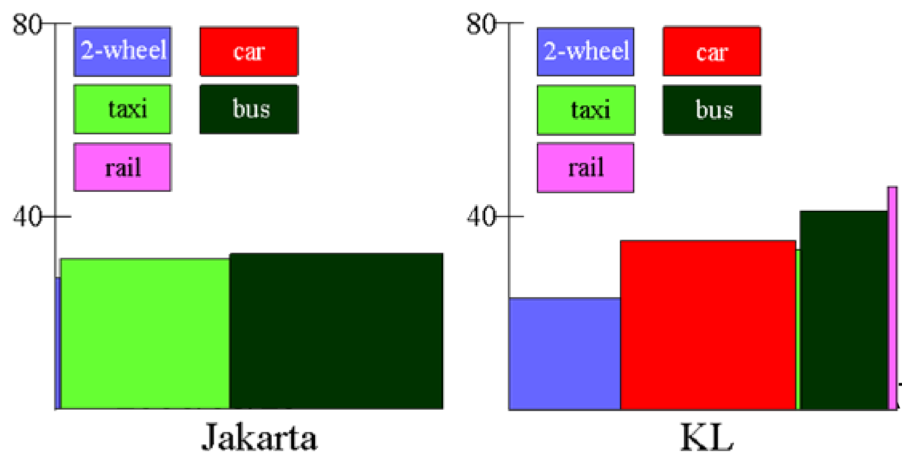
10Jakarta
1Walking to final destination
2Walking for transfer
3Bicycle
4motorcycle
5Sedan, jeep, kijang
6Colt, mini cab
7Pick up
8Truck
9Rail(express)
10Rail(economy)
11 Patas AC
12Large bus (patas, regular)
13Medium bus
14Mini bus(Angkot or mikrolet)
15Taxi
16Bajaj
17Ojek
18Becak
19Omprengan
20Company bus, school bus, tour bus
21Others

2Phnom Penh
1Passenger Car
2Taxi
3Light Bus/Pass.Van
4Pick-up/Cargo Van
5Truck/Trailer
6Large Bus
7Mortorcycle
8Mortodop
9Motorumo
10Cyclo
11Bicycle
0Walking
00Others

11KL
1Walking
2Bicycle
3Motorcycle
4Car
5Small Van(For Passenger)
6Taxi
7Mini Bus
8Feeder Bus to/from KTM or STAR station
9Intrakota
10Park Mmay/City Liner
11Other Stage Buses(with A.C.)
12Other Stage Buses(without A.C.)
13Factory Bus
14School Bus
15Other Buses
16Small Lorry(light 2-Axles)
17Other Lorries
18STAR(LRT)
19KTM Train
20Others
*A.C. : Air Condition

08Bucharest
1Walk
2Bicycle
3Motorcycle
4Automobile
5Pickup, Van, Freight Vehicle less than 1.5 Tons
6Medium truck (1.5 - 3.5 Tons Capacity)
7Heavy Truck (over 3.5 tons Capacity)
8Taxi
9Maxi Taxi
10RATB Bus
11Express Bus
12Private Minibus, Company Bus
13Trolley Bus
14Tram
15Metro (Subway)
16Train (Railway)
17Other





TBR

Average trip duration vs. modal share

Area means total trip time

→It relates environmental emissions.

3) Demand Models for Big Projects in Asia

Korea Train eXpress (KTX)

Taiwan High Speed Rail (THSR)

Wen (2003) applied GNL for Inter-regional modal choice in Taiwan

Yang (2005) also analyzed comparative analysis on:

MMNLogit model, heterogeneous logit model, latent class model…

Major Airport in Asia:

-New Hong Kong International Airport (1998)

-Kuala Lumpur International Airport (1998)

-Shanghai Pudong International Airport (1999)

-Incheon airport in Korea (2001)

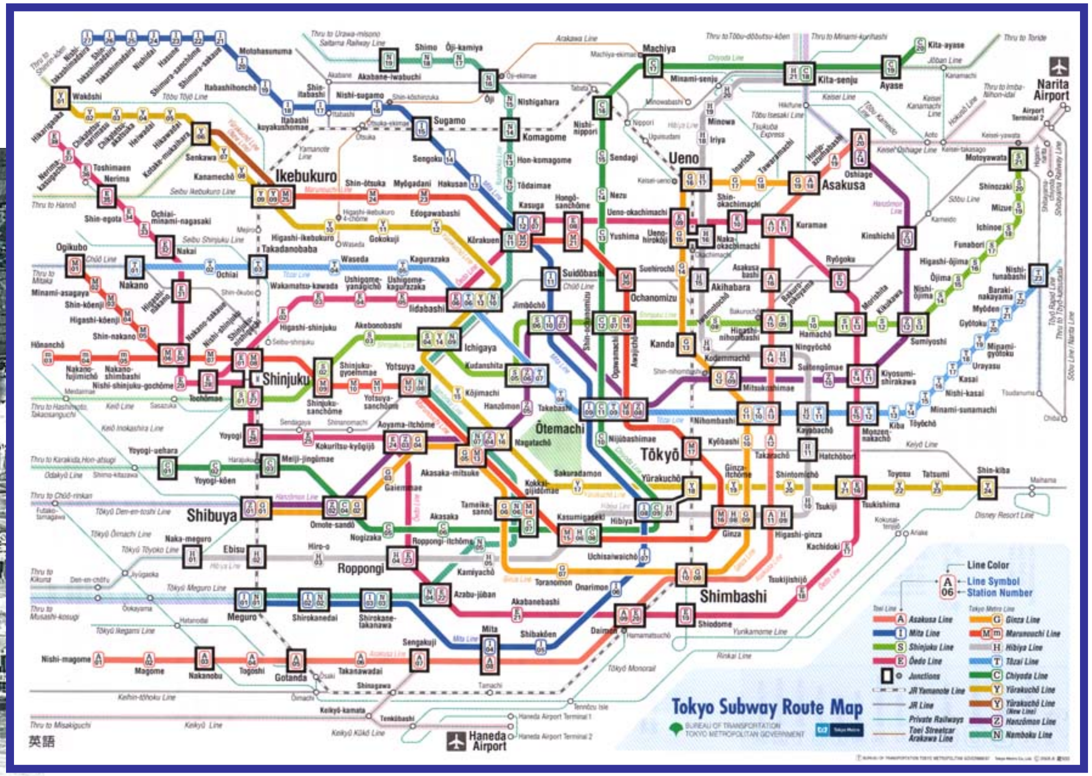
-Centrair airport in Nagoya (2005)

4) Advanced Modeling for Dense Transit Network in Asia

a) A number of stations and lines generate enormous alternatives

→ "Structured Probit Route Choice Model" (Yai et al., 1997) was applied for future master plan of railway in TMA

→Hibino et al. (2004) also examined comparative analysis with Probit model, MMNL model and C-logit model

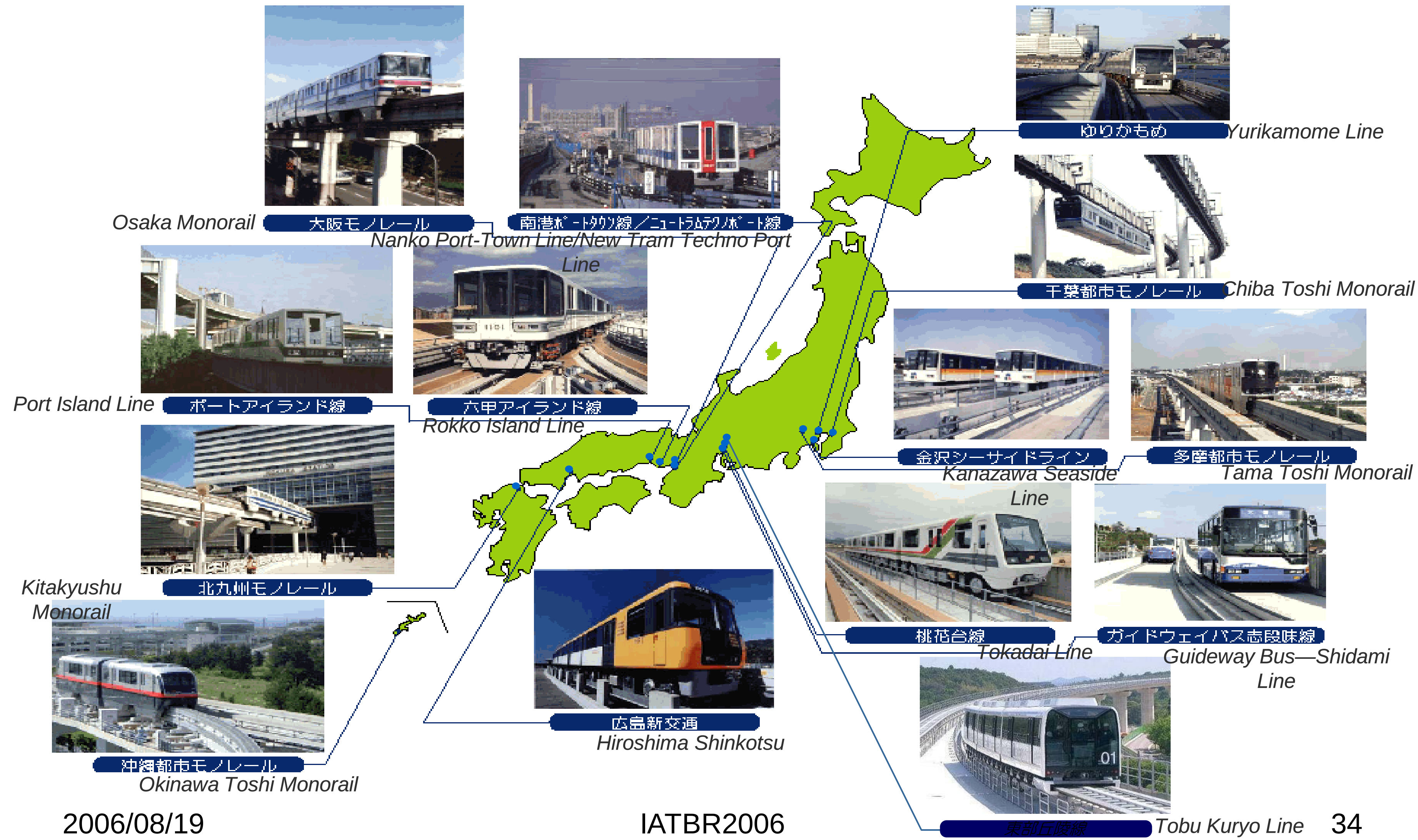


b) Railway/Subway stations have many access/egress modes

→ Hierarchal modeling techniques are required

- Muromachi (2003) introduced GNL model for route and parking location choice model
- Mizokami (2003) also estimated GNL or CNL model and C-logit for park and ride behavior

New Transportation, Urban Monorail and Guideway Buses in Japan



c) Analyses on New transportation policies

- Peak load pricing, variable (flexible) fare structure …
- Iwakura et al. (2003) developed a departure time choice model
  → The error covariance structure among departure time utility by a MMNL model

Hyper congestion at Tokyo station (1970')

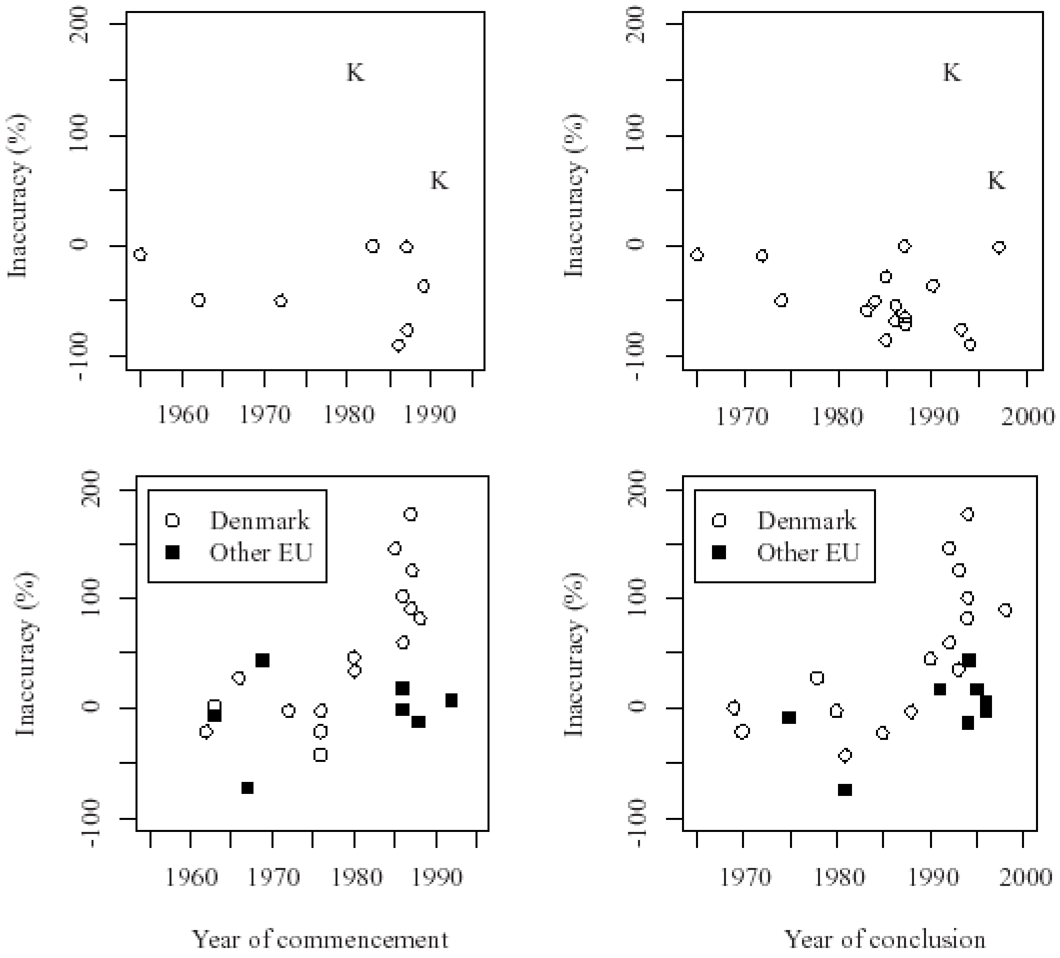
Outline

- Recent developments in econometric choice modeling
- Characteristics of transport modeling in Asian cities
- Inaccuracy of transport demand models

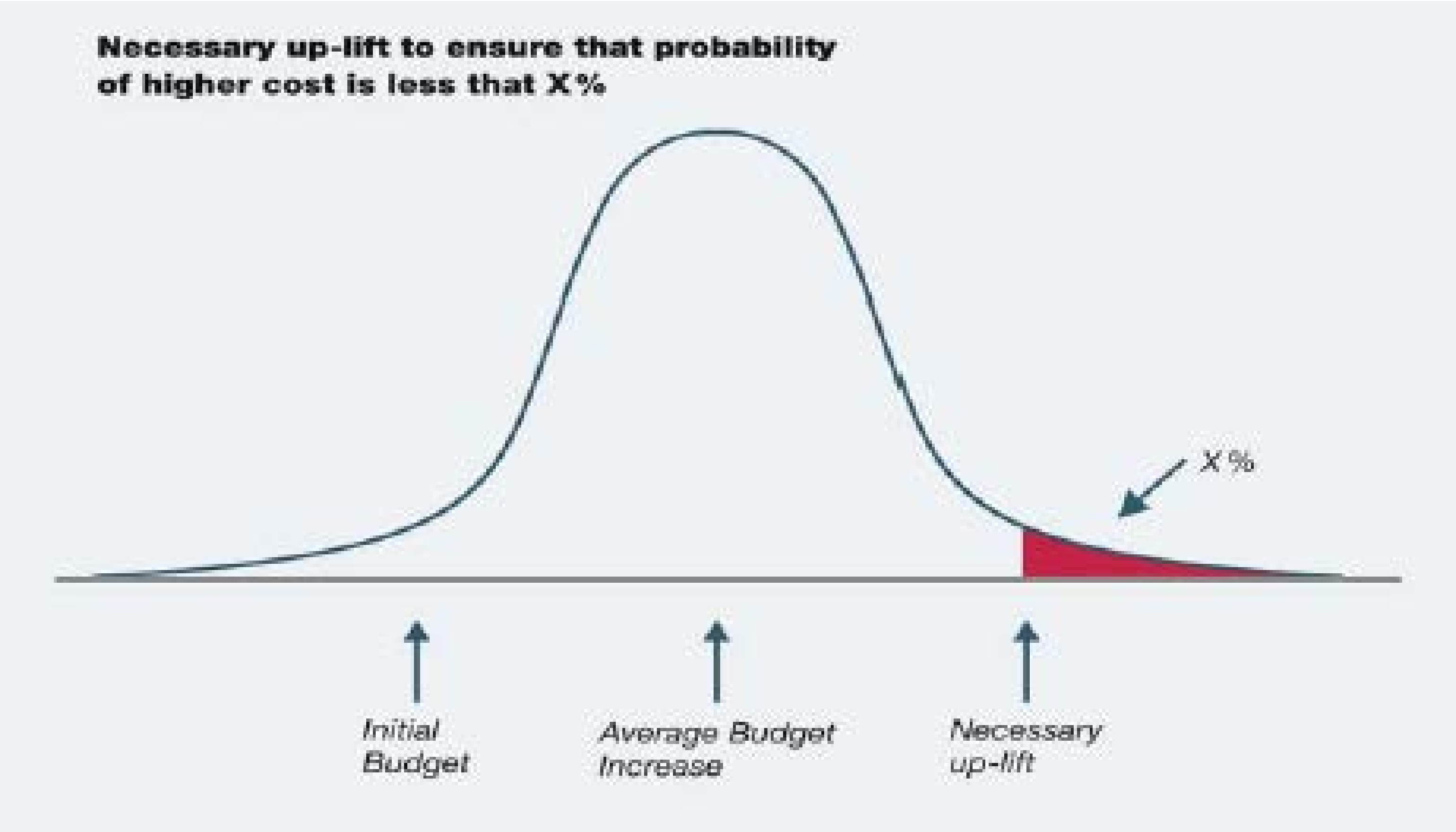
Inaccuracy of Transport Demand Models

- Flyvbjerg et al. (2005) investigated 210 road and rail projects worldwide and found that the number of cases for a large difference between predicted and observed demand is not small.
- Flyvbjerg et al. also concluded that accuracy in transport demand forecasting has not improved over time, which might undervalue continuous theoretical development of transport demand models.
- If planners are to get forecasts right, Flyvbjerg et al. recommended a new forecasting method called “reference class forecasting” to reduce inaccuracy and bias. Reference class forecasting uses “outside view” on the particular project being forecast that is established on the basis of information from a class of similar projects.

Inaccuracy Over Time in Forecasts for Rail and Road Projects(2005)



Procedures for Dealing with Optimism Bias in Transport Planning



Japanese Cases

- The outputs transport demand models produce are major inputs into cost-benefit analysis of transport infrastructure projects in Japan, as is in most other countries.
- For some projects, the discrepancy between predicted and observed demand has incurred severe criticism.
- Inaccuracy of transport demand forecasting even became one of the major agendas during the privatization process of Japan Highway Public Corporation.
- In coupled with some corruption cases by government officials and long economic slump during the 1990s, inaccuracy of transport demand forecast for some large transport infrastructure projects made the public trustless to transport demand models.

The Aqualine

- The new bridge and tunnel crossing the Tokyo Bay, the Aqualine, carried only about forty percent of the number of vehicles predicted when it opened in 1997.

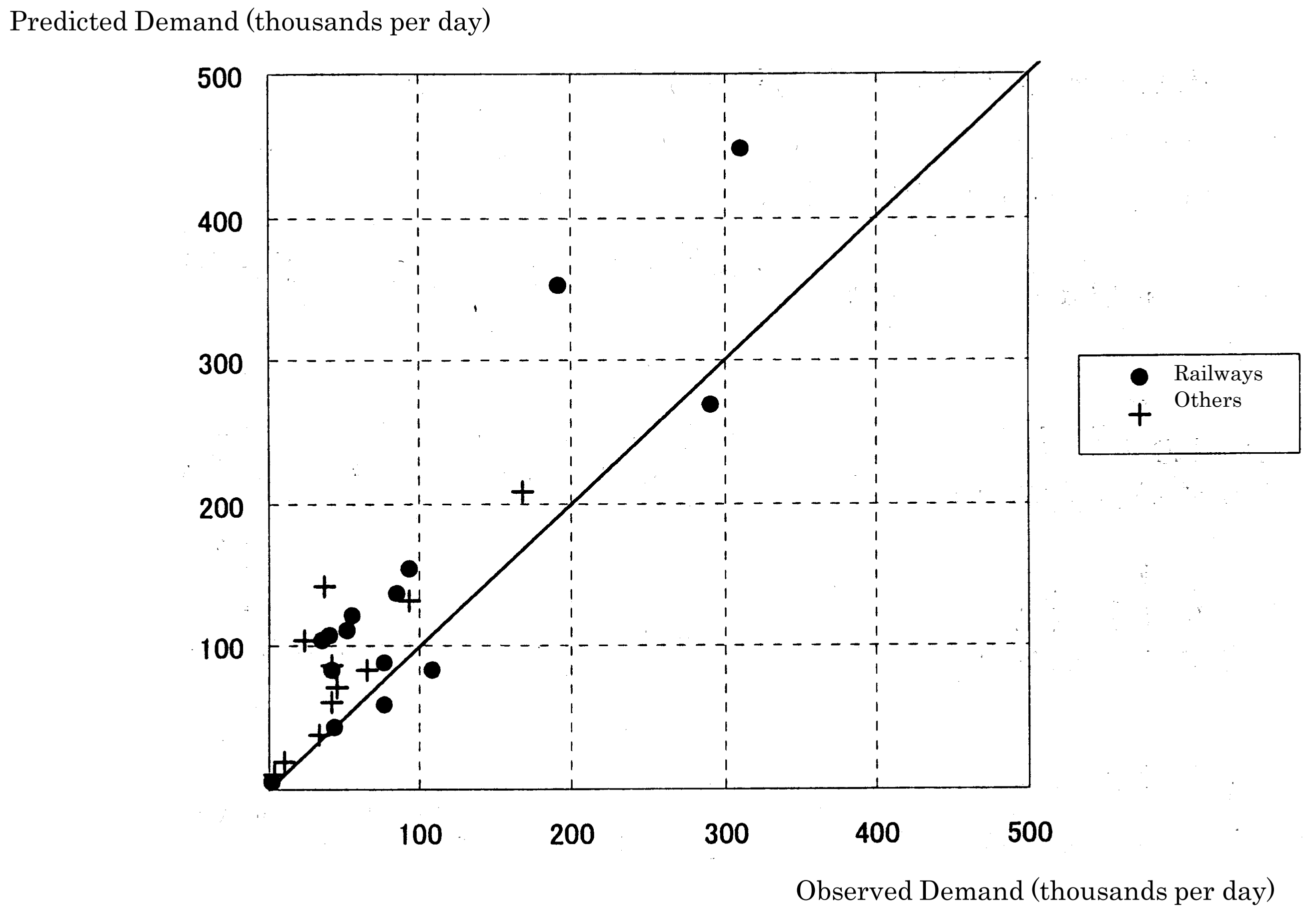
Ex-post Evaluation of Transportation Planning Group (1987)

- EETPG considered three types of uncertainty in relation to transport planning: UE (uncertainty about the related planning environment), UR (uncertainty about the related decisions) and UV (uncertainty about value judgments).
- Investigating the discrepancy between predicted and observed demand for the metropolitan transport study and the road project cases, EETPG concluded that one of the most important estimates was total transport demand, or control total.
- EETPG also found that root mean square error at the step of trip distribution was the largest of the four step transport demand models and needed further studies.

Institution for Transport Policy Studies (ITPS) (2001)

- ITPS investigated predicted and observed demand for 26 railway segments recently opened. ITPS found that prediction error was within 20 percent for 5 segments, more than 20 to 100 percent for 10 segments and more than 100 percent for 10 segments.
- ITPS found that while prediction error of some segments was mainly ascribed to population overestimate, prediction error of other segments might be generated by other factors such as demand forecasting method.
- ITPS concluded that prediction errors generated by modal split and route choice steps were larger than the errors by the other steps. The inappropriate premises of the level of service for railways and cars and of the restructuring of bus network also caused large prediction errors.

The Comparison between Predicted and Observed Demand



How Would We Do?

- Doi et al. (1997) studied past demand forecasting for Tokaido Shinkansen and concluded that premises of national income and Shinkansen fare, disregard of competition with air, and time required for switching to new mode just after the opening made the difference between predicted and observed demand.
- After investigating about 14.5 times higher predicted than observed demand for new public transport system, Morikawa et al. (2004) concluded that, of four step transport demand models, generation step, or population input, made the difference by about 1.7, modal split step about 6.6 to 7.3 and others about 1.2 to 1.3 times.
- The trust by the public in transport demand models and transport infrastructure planning must be recovered. Yai, et al. (2006) proposed giving predicted demand with distribution and studied its acceptability by the public.

- It is inappropriate to ascribe the discrepancy between predicted and observed demand for some large transport infrastructure projects only to the deficiency of transport demand models.

- However, it might also be inappropriate to free transport demand models from any charges against the discrepancy.

- Future studies still need to give more insight into human (travel) behavior on which any transport demand models should depend

Thank you