

# Data Transferability: Idealism or Realism

Kouros Mohammadian  
University of Illinois at Chicago

Peer Exchange on Data Transferability  
FHWA-TMIP  
December 16, 2003

1

## Data Transferability

- The idea is to use data collected in one context in a new context.
- This can reduce or eliminate the need for a large data collection for model development effort in the application context.
- While data transferability is an emerging research topic, travel demand model transferability has been subject of numerous studies.

2

## Travel Demand Model Transferability in the Literature

- Several empirical studies conducted to assess the effectiveness of model transferability.
- Some of these studies have examined model transfer from one spatial context to another
  - *Atherton and Ben-Akiva 1976, Galbraith and Hensher 1982, Koppelman et al 1985, McCoomb 1986, Gunn and Pol 1986, Tretvik 1989, Abdelwahab 1991, Algiers et.al. 1994, Karasmaa 1995, Sermons 2000,*
- Others have examined the temporal transfer of these models
  - *Talvitie and Kirshner 1978, Train 1979, McCarthy 1982, Hague Consulting Group 1990, Karasmaa and Pursula 1997, Karasmaa 1998, Badoe and Miller 1995a,b, Badoe and Miller 1998, Walker et al. 1998, Elmi et al. 1999.*

3

## Applications of Transferred Travel Data

- Examples of the use of transferred transportation data:
  - NCHRP 365, ORNL's NPTS transferability study, etc.
  - ITE trip generation rates for different land uses
  - ITE's *urban travel characteristics database*
  - FTA's *Characteristics of urban transportation systems*
  - Household travel survey data simulation studies (e.g., Stopher/Greaves works).

4

## Why Simulate Travel Data?

- Household travel data is critical to transportation planning and modeling
- Survey is expensive (\$150-200 per completed household survey)
- Need for data for small regions that local household travel data is not available
- Sampling rate is too low for the NHTS/NPTS to be directly used at a local level. (NHTS/NPTS data is not designed for trans. planning and modeling)
- Emerging modeling techniques (e.g., microsimulation models) are becoming available to planners that require much richer datasets that do not exist in most metropolitan areas.

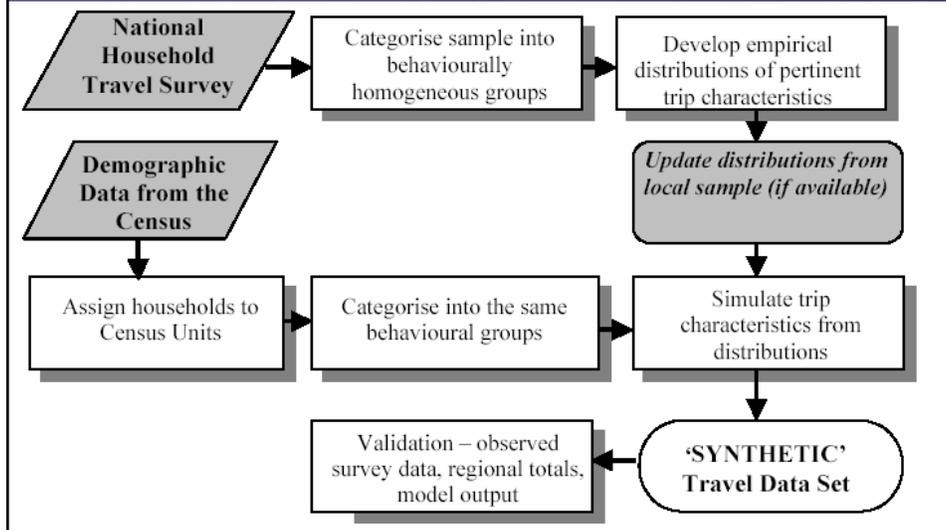
5

### *Basic Idea behind data simulation*

- Combine local socio-demographic data for individuals/households (from sources such as a census) with probability distributions of activity/travel patterns (from other travel survey, such as NHTS) to simulate local travel survey data.

6

## Overview of Methodology (Stopher and Greaves, 2004)



## Travel Attributes Selected for Simulation

- Several attributes can be selected to be simulated:
  - trip rates by purpose:
    - home-work, home-school, home-college, home-shop, home-other; non-home-work, non-home-other
  - mode shares
    - auto driver, auto passenger, bus, bike/walk, rail and other
  - departure times
    - e.g.: 6am-9am, 9am-4pm, 4pm-7pm, 7pm-6am
  - trip lengths
  - etc.

## *Sequence of Simulation*

- First the trip rates by purpose produced by the household
- Then the main mode of travel for each trip, given the purpose
- Then the time of departure, given the mode and purpose
- Finally, the trip lengths, given purpose, mode, and time of departure

9

## *Notes*

- Classification schema is a critical issue in simulating Household Travel Survey data.
- The further the population and the dependent variables are we subdivide into homogeneous groups, the better chance the simulation will ultimately have of reflecting these differences.
- By its nature, simulation is simply expanding/transferring existing data with all the inherent biases to a wider population.

10

# Transferability Project at UIC

- To develop a framework to facilitate transferability of HH survey data in calibrating and validating travel forecasting models. The project involves:
  - More detailed classification of NHTS and Census data using advanced clustering schemas.
  - Travel data simulation and further improving current literature by including tours, joint trips, etc.
  - Use synthesized and transferred data for model calibration and validation.