Using SEM in Analysing Travel Behaviour

Kaveh Jahanshahi[1]

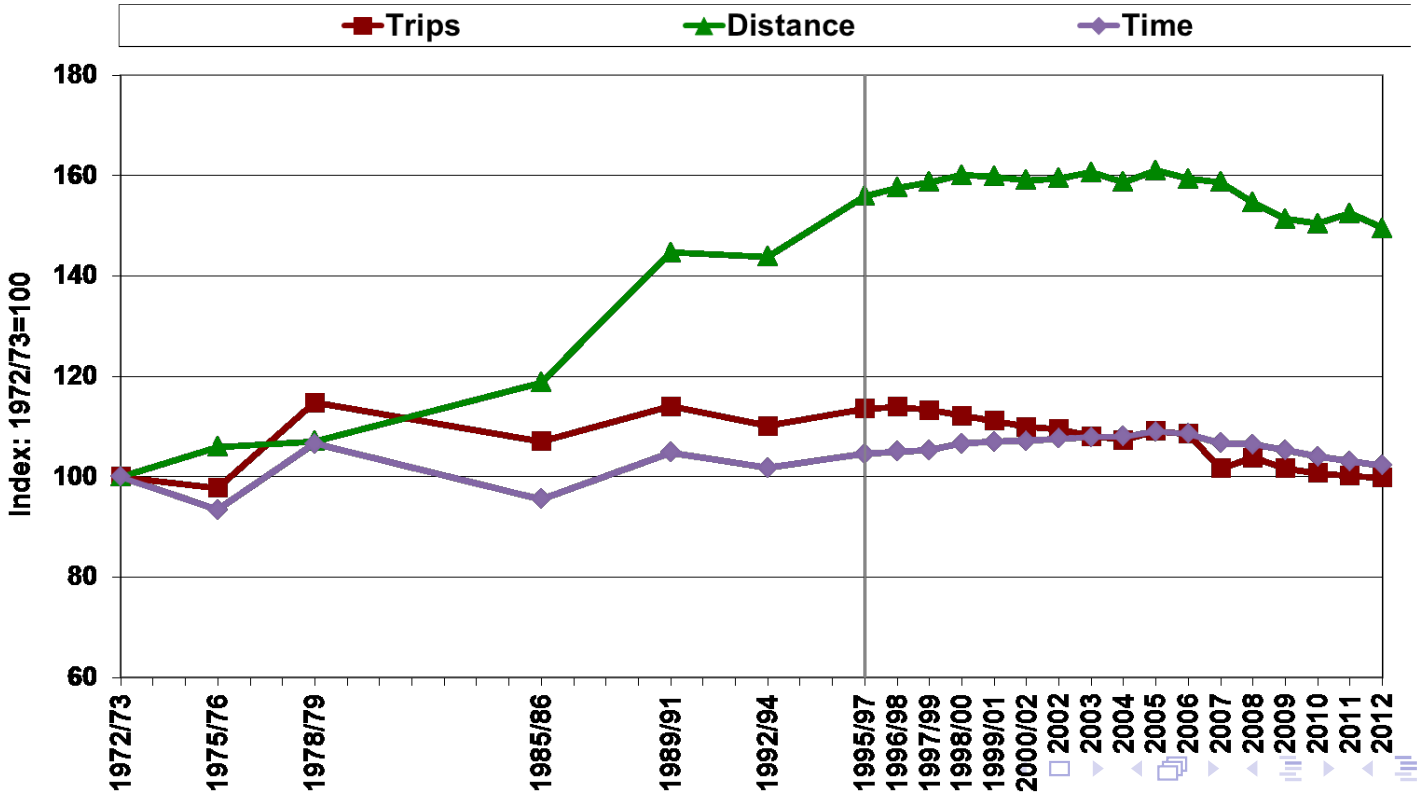
[1]PhD Student
University of Cambridge

MPLUS User Group Meeting
July 2014

Outline

Personal travel in the UK: a headline summary

- On average, each person makes 1000 trips a year, the daily average travel time is just over an hour. This hasn't changed much over the past 30 years
- A rapid increase in average travel distance till the late 1990s – now stabilizing around 7 miles/trip and 6700 miles/person.year



Where stands in Literatures

A large number of studies have examined the influence of land use patterns on travel behaviour

- Specifically car use and travel distance
- Only few on travel time

More recent studies aim to control the endogeneities

- examining the interdependency between travel patterns, travel attitudes, built environment characteristics, and car ownership
- Very few number of studies outside US

Main Research Questions

The observed trend raises few research questions:

- Are we at last giving up the aspiration to travel further, particularly by car ('Peak Car')?
- How do people spend their travel time?
  - Past research tend to focus on how the miles travelled; few looked at the time spent in travel and trip frequency
  - Travel time is directly related to traffic congestion, the ease to travel by public transport, transport energy efficiency, and how people live and work
  - Improve understanding of travel patterns by analysing all of its indicators

Any Trend-breakers?

- How much difference do urban planning and design make?
- Do people now behave differently since the Financial Crisis?

Outline

Research Gaps

- Studying travel time and trip frequency
- The heterogeneity between car ownership, socioeconomic factors and travel pattern indicators
  - e.g. Weis and Axhausen (2009) surprise finding of the absence of a significant income and car ownership effect
- Interaction among different travel purposes
- Explanation of change in travel pattern indicators over time making a use of improved model.
- Lack of comprehensive consistent dataset

Aim of study

In order to answer to the raised questions, we need to develop a new improved model which is capable for:

- Evaluating both direct and indirect influences
- Controlling for interactions among:
  - Travel purposes
  - Travel time, distance and frequency
  - Socio-demographic factors, built form and car ownership

Technically robust model

- Accounting for heterogeneities among household members
- Accounting for correlation among land use indicators
- Using weights at NTS data

Outline

Data

Great Britain National Travel Survey (NTS) dataset for the years 2002 to 2008

- household surveys designed to provide regular, up-to-date data on personal travel and to monitor changes in travel behaviour over time.

A weighting strategy for the NTS

- Household weights for adjustment for non-response household
- Trip weights
  - Drop off trips in course of the survey week
  - Recording of short walks
  - Underreporting in long distance trips

Variables to be used for the study

Land Use indicators

- Area Type
- Population Density
- Accessibility (Bus frequency, nearest shopping center etc)

Socioeconomic Factors

- Work Status
- Gender
- HH Income
- HH Size, etc

Travel data

- Travel time
- Travel distance
- Trip frequency

Outline

Choice of Technique

One combined framework to analyse the interactions among

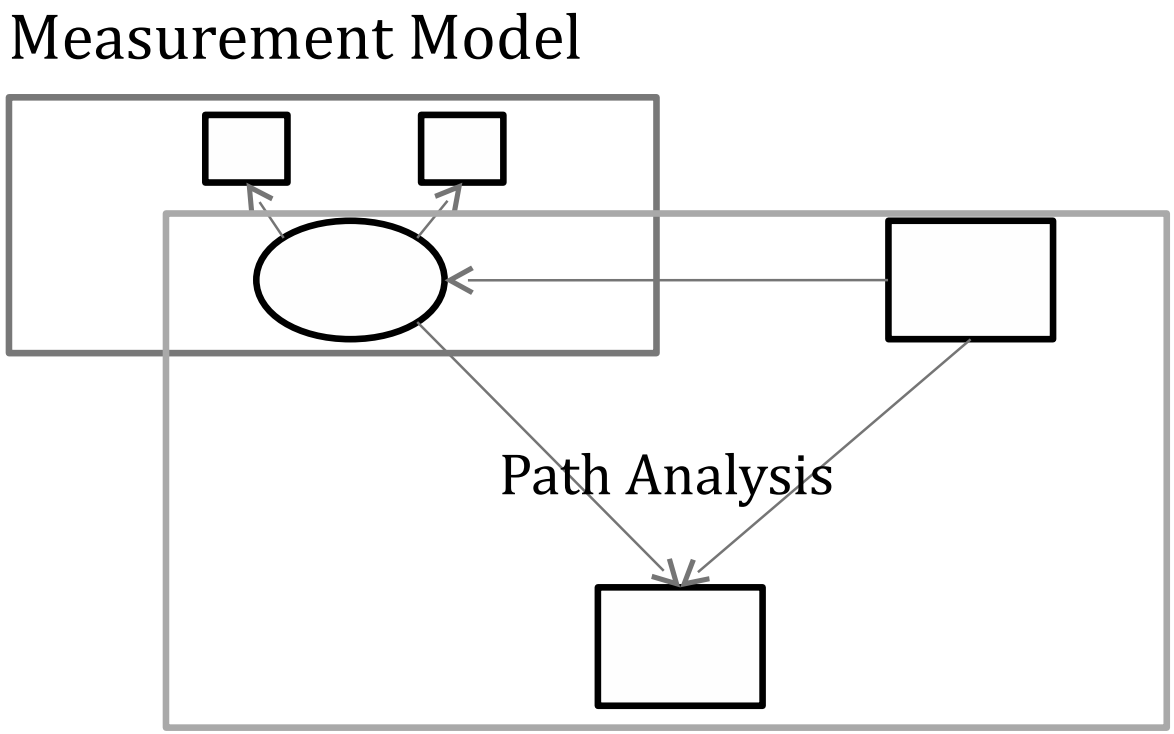
- travel purposes
- socio-economic and demographic factors
- accessibility and land use characteristics
- car ownership

Structural equation modelling (SEM) is a statistical technique for testing and estimating **causal relations** using a **combination of statistical data and qualitative causal assumptions**

SEM

SEM is the union of

- Path analysis (series of regression equations) and analysis of latent variables or measurement model; a model to define unobserved factor measured by some observed variables

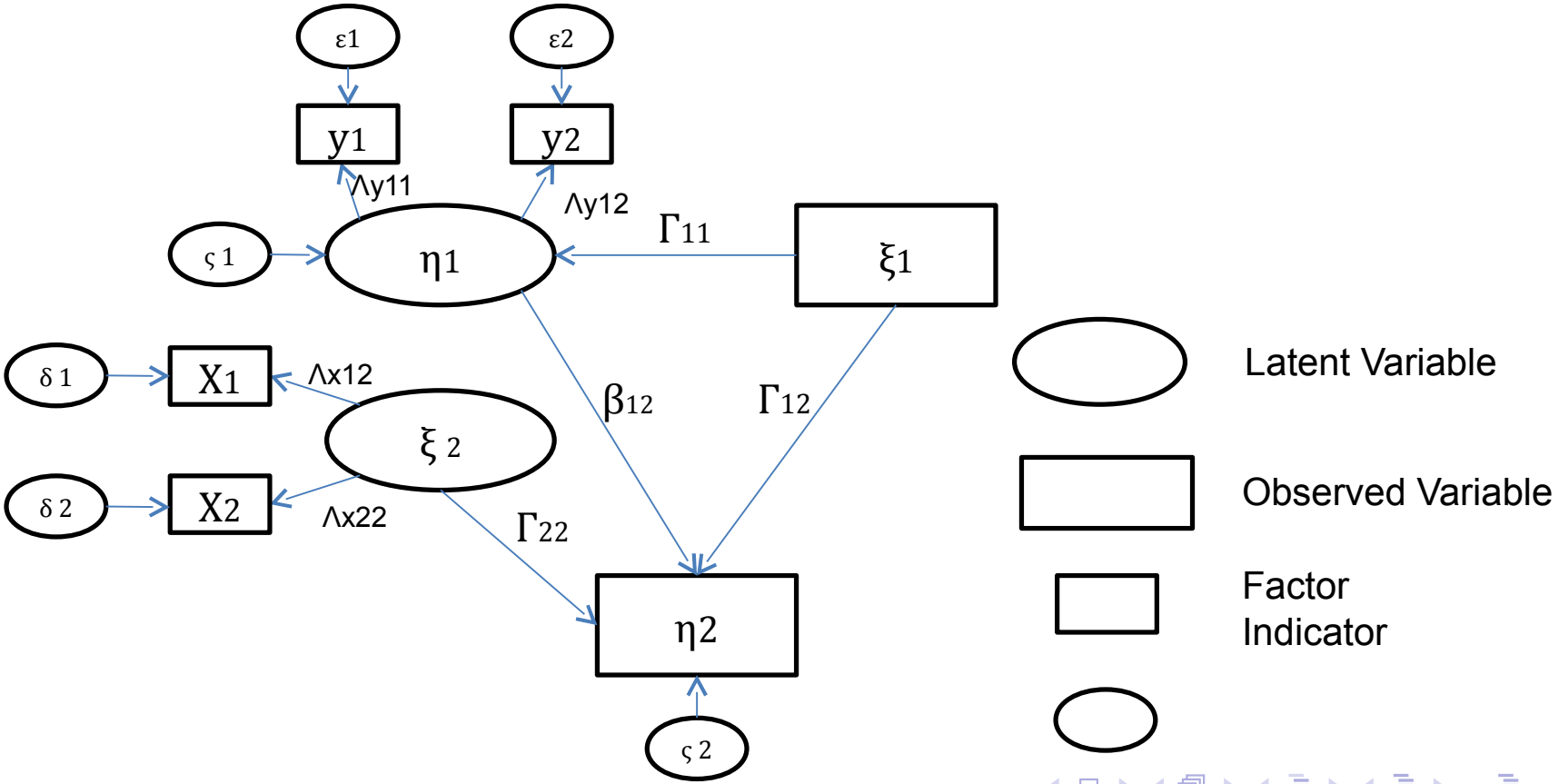


Notations

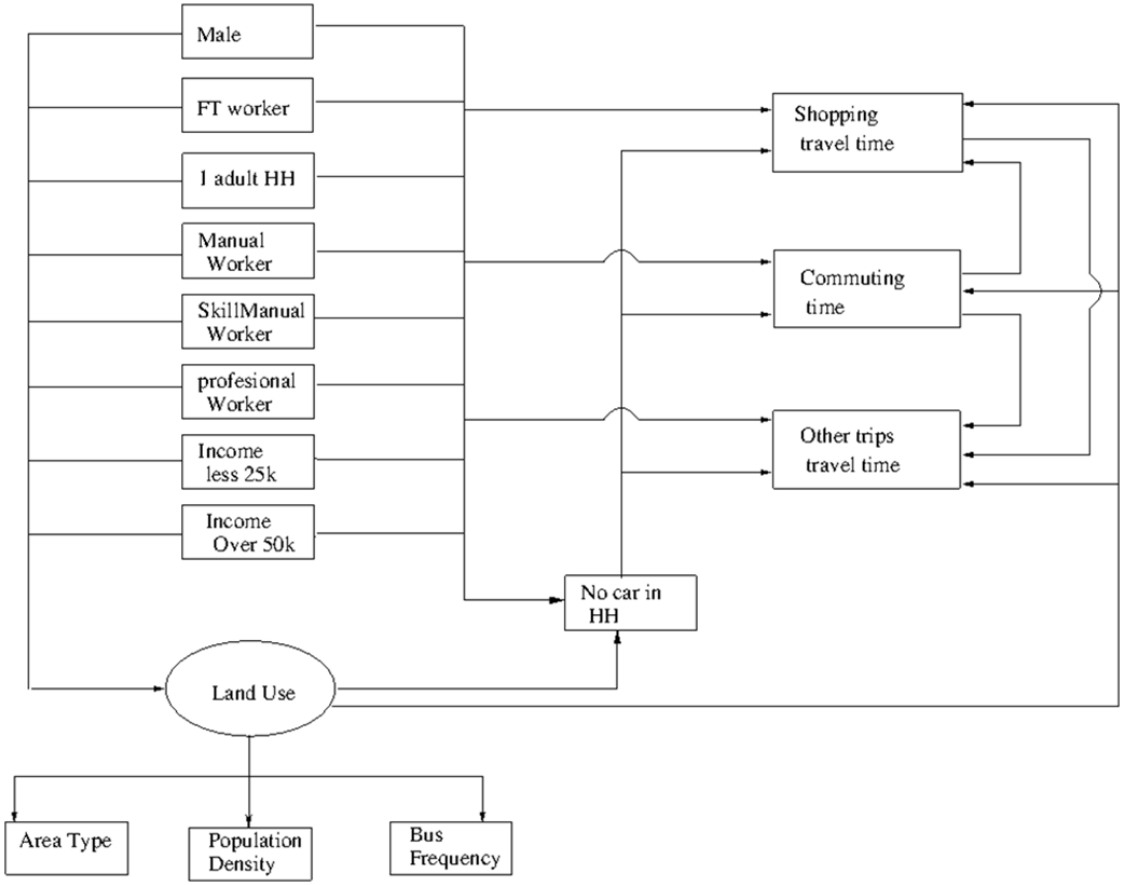
$$\eta = \alpha + \beta\eta + \Gamma\xi + \zeta \quad (1)$$

$$y = \tau y + \Lambda y\eta + \varepsilon \quad (2)$$

$$x = \tau x + \Lambda x\xi + \delta \quad (3)$$



Choice of Technique



- **Gender:** Male or Female
- **Work Status:** full or part time
- **Household size:** 1 or 2 or more adults
- **Skills:** Manual, Skilled Manual, Clerical, Professional
- **Household income:** Less than 25K, 25k-50K, over 50k
- **AreaType:** London, Metropolitan, Big Urban, Medium Urban, Small Urban, Rural
- **Car Ownership:** No car, 1 or more cars

MPLUS Input file

```
VARIABLE:
 NAMES =   HHCode Year YearSplit ...;
 Usevariable = HHCode YearSplit ...;
 Categorical = AreaType Bus_Freq PopDen CONo_Car;
 WEIGHT=W2;
 CLUSTER is HHCode;
 class is Year(3);
 knownclass is Year (YearSplit = 1 YearSplit = 2  YearSplit = 3);

Define:
TTT_HBW_Sc = TTT_HBW/100;
TTT_Oth_Sc = TTT_Oth/100;
TTT_Sh_Sc = TTT_Sh/100;

ANALYSIS:
  TYPE IS COMPLEX MIXTURE;
  ALGORITHM = INTEGRATION;
  processors = 4;
LINK=PROBIT;
MODEL:
%overall%
LU BY AreaType* Bus_Freq* PopDen*;
         LU@1; ...
%Year#3% !MODEL Before2007: ...
%Year#2% !MODEL After2007: ...

      MODEL CONSTRAINT:
      NEW (ind_Sh_HBW_CO);
      ind_Sh_HBW_CO = HBW_CO*Sh_HBW; ...


MODEL TEST:

HBW_CO_A07 = HBW_CO_B07; ! repeated for various others
```

- NTS HH Weight for weighted regressions
- Clustering by Households
- Use WLS for standardized coefficients
- Use probit regression to be comparable with WLS
- Multi group model by Year
- Dividing Travel Time by 100 for convergence

Outline

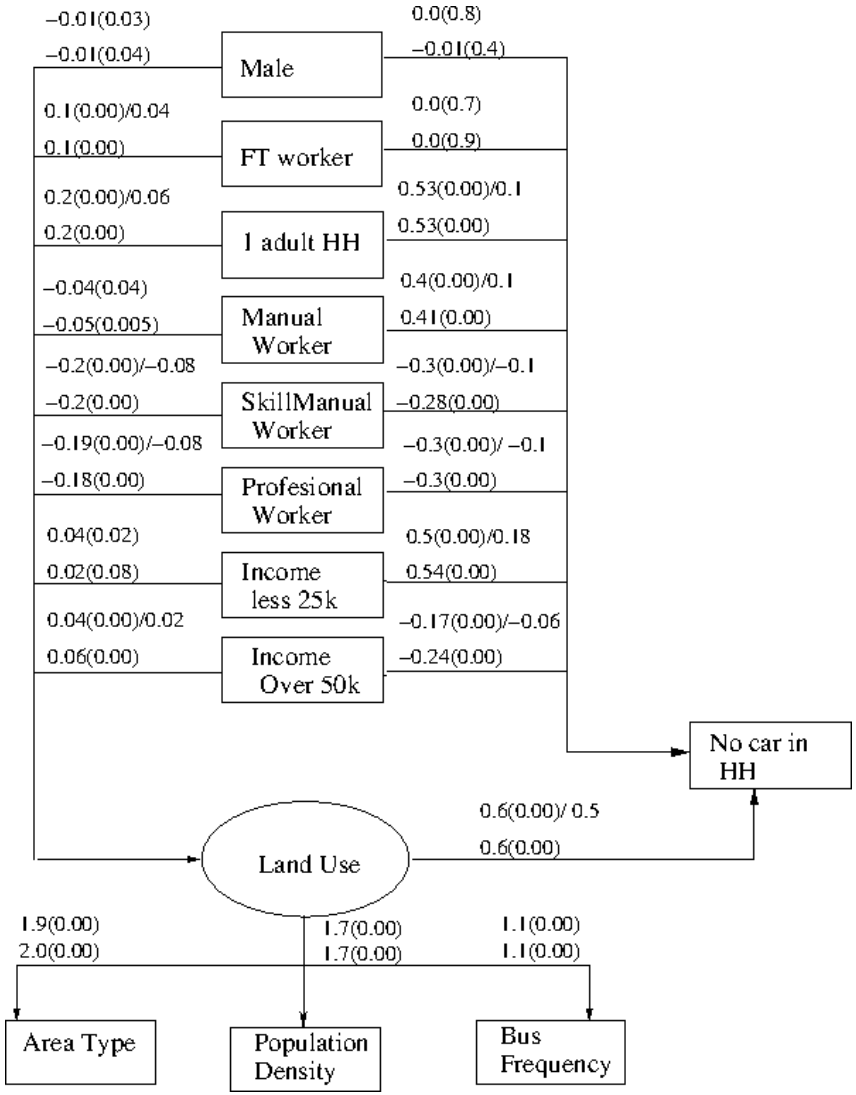
EFA to select land-use indicators

VARIMAX ROTATED LOADINGS

	1	2
WT_BS	0.263	0.063
FREQ_B	-0.720	-0.143
BT_RAIL	0.232	-0.155
RS_TYPE	0.190	0.766
AREATYPE	-0.879	-0.141
POPDEN	-0.825	-0.081

- PopDen and AreaType are highly correlated
- Minimum three indicators for just-identification
- The aim is to control for high correlations not best fitness for EFA

LU effects (WLS/MLR) on Car Ownership



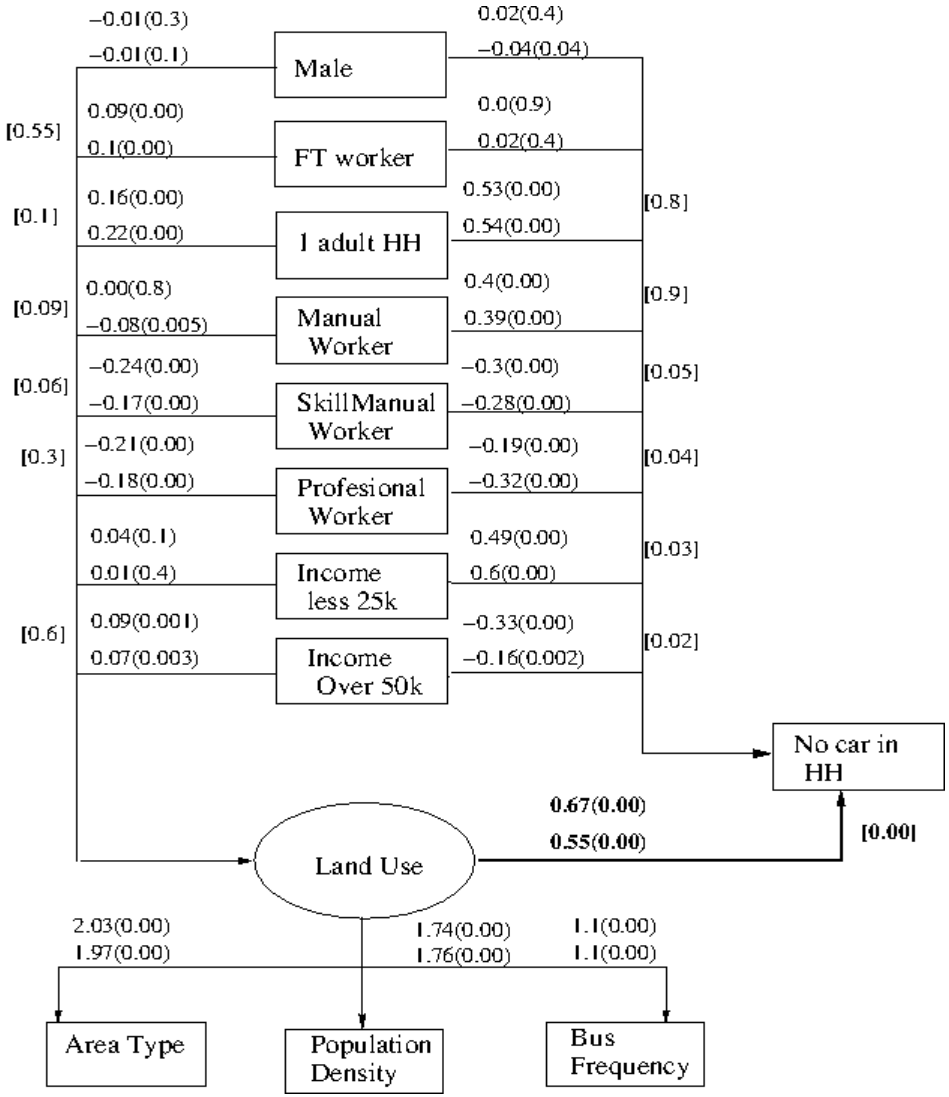
Comparing Constrained and Grouped model

Table : goodness of fit statistics

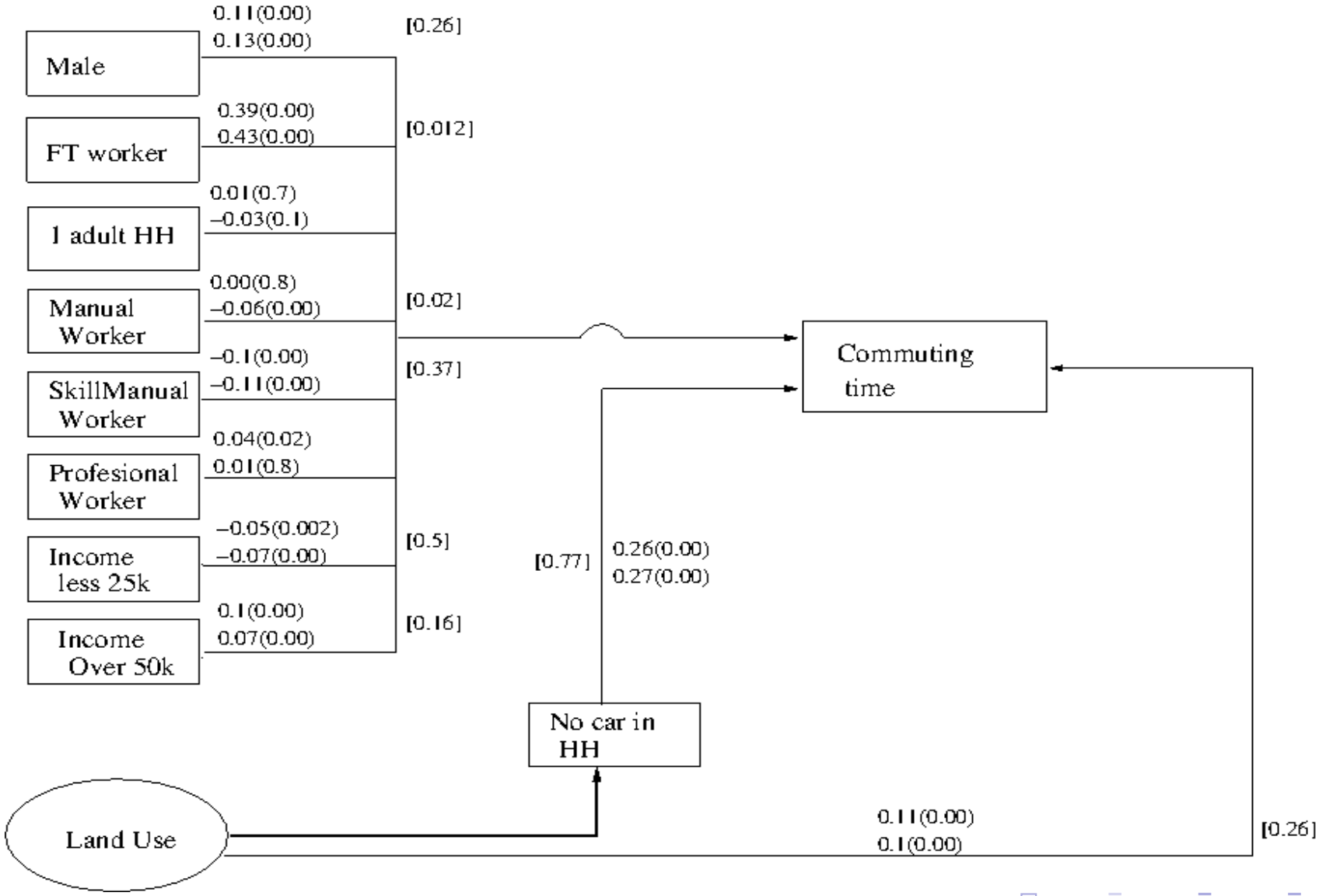
	AIC	BIC	ABIC
Constrained model	1,445,032	1,445,863	1,445,577
Grouped model	1,444,897	1,446,705	1,446,082

Both AIC and BIC are maximum likelihood estimate driven and penalize free parameters in an effort to combat overfitting. However, the penalty term is larger in BIC than in AIC

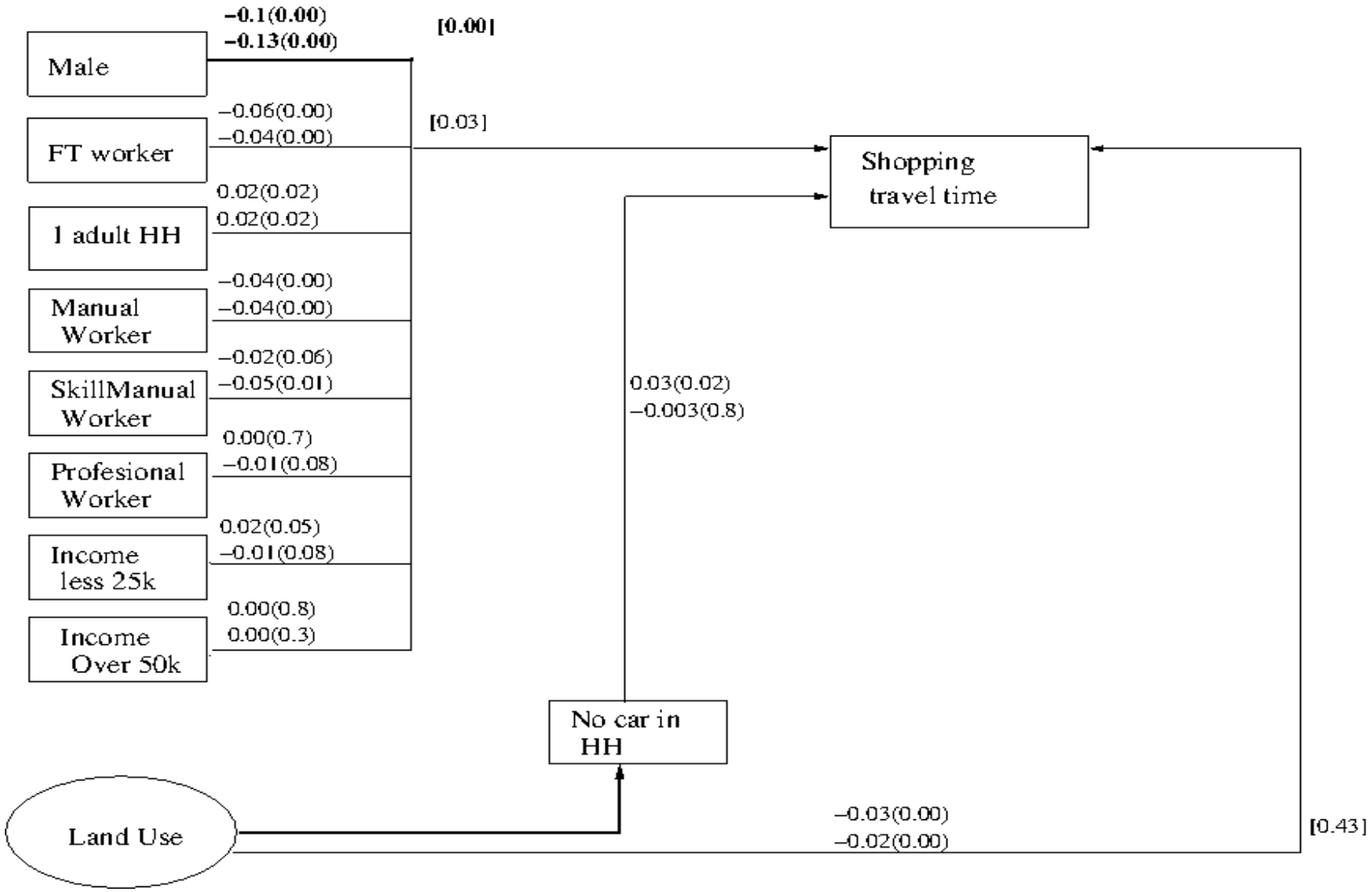
LU effects (B/A 07) on Car Ownership



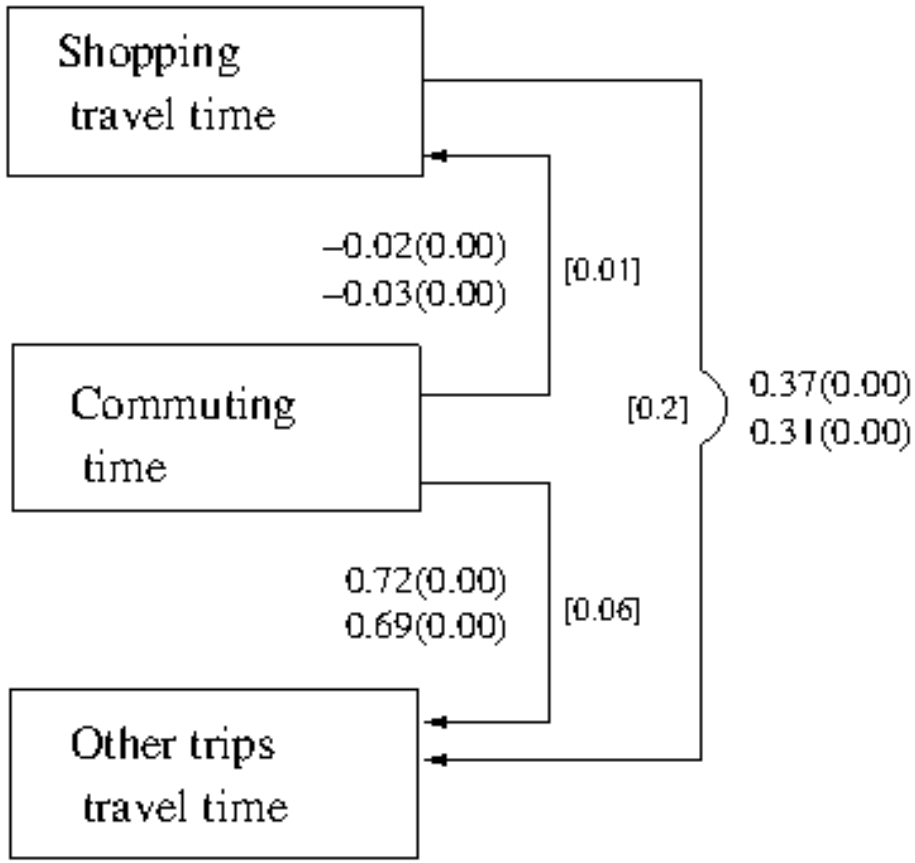
LU effects (B/A 07) on commuting time



LU effects (B/A 07) on shopping time



Interaction between travel purposes (B/A 07)



The importance of indirect effects

Table : Indirect influences

Direct Effect	Indirect effect	ML Estimator	
		Estimate	S.E.
FT->NoCar		-0.002 *	0.021
	FT->LU->NoCar	0.058	0.007
1adult->HBW		-0.013 *	0.012
	1adult->LU->HBW	0.021	0.002
	1adult->LU->NoCar->HBW	0.031	0.003
	1adult->NoCar->HBW	0.135	0.011
NoCar->Sh		0.008 *	0.008
	NoCar->HBW->Sh	-0.006	0.001

Comparing alternative handling of car ownership model

Table : Car ownership model

	Land use effect on commuting time		
	Direct effect	Indirect effect	Total Effect
Full model	0.105 (0.005)	0.154 (0.011)	0.259
Small Model	0.104 (0.005)	N/A	0.104
	The effect of having no car on commuting time		
	Direct effect	Indirect effect	Total Effect
Full model	0.253 (0.017)	N/A	0.253
Small Model	0.271 (0.017)	N/A	0.271
	Total effect for an average person		
Full model	0.512		
Small Model	0.375		

Outline

Key direct and indirect effects

The SEM approach taking account of all the influences does make a significant difference to our understanding of travel behaviour

- New insights into the different factors affect each other, and exert indirect influences on travel
- Particularly the influence of built form on car ownership – the drop of car ownership in dense areas with convenient public transport is a main driver to reducing car travel

The effect of recession

Shows that people changed little, but the changes in their circumstances have influenced their behaviour. For example:

- Living in a dense urban area has made it 23% more likely to forgo car ownership, and therefore making fewer trips by car
- But adult males still don't like going shopping – the data shows that they go shopping less because they are males, not because they spend more time commuting
- However, after 2007, adult males now spend on average 3 minutes more for going shopping

Outline

unobserved groups?