

Taste Heterogeneity as a Source of Residential Self-Selection Bias in Travel Behavior Models: A Comparison of Approaches

Patricia L. Mokhtarian

plmokhtarian@ucdavis.edu

www.its.ucdavis.edu/telecom/

The Joy of the Journey:

Celebrating the Life and Work of Ryuichi Kitamura

University of California, Davis, June 28-30, 2009

Ancestry of this Project

Importance of
residential location
& lifestyle
(Kitamura, 1988)

Residential location is extremely relevant to the discussion... But what is the relation between housing preferences and household travel behavior? Unfortunately, little is known that might answer this question...

household car ownership... is desired that the car ownership variable be replaced by another variable more intrinsically related to travel behavior.

Other factors that are less frequently used in travel behavior analysis but are nonetheless relevant here include education, ethnicity, and residence location. Education is used by Salomon as one of the dimensions along which the life-style orientation is defined (9). Past studies have found certain associations between education and travel behavior (60). Allaman et al. (4) found education to be associated with the time spent working, eating away from

A very fundamental question is whether travel behavior is conditionally independent of life-style orientation, given residence location, car ownership, and other measurable factors.

caused by shifts in life cycle, or by economic... little is known about the association between ethnicity and travel, possibly due to the multitude of ethnic and cultural backgrounds. In addition, the frequently

observed correlation between ethnic background and socioeconomic status makes the isolation of an ethnic effect a difficult task.

Residential location is extremely relevant to the discussion if suburban life-style is the cause of urban congestion problems. Relatively little attention has been paid in travel demand analysis to the factors that influence choice of residential location. The key issue that needs to be addressed is how households trade off among commuting distance (and time), housing price, and various amenities that vary greatly by location. Among the preferred housing attributes are "better" neighborhood quality, better schools, a new house in a well-established neighborhood, a housing unit all on one floor, and a large lot (77). Obviously these preferences are correlated with the life cycle and life-style of the household. But what is the relation between housing preferences and household travel behavior? Unfortunately, little is known that might answer this question.

The existing results indicate that trip generation is negatively correlated with population density, suggesting geographical variations in life-style within a metropolitan area (4, 78). The significant effect of home ownership found in several studies (51) may again be viewed as an indication of the association between population density and travel behavior. An analysis of daily travel patterns concludes that (36) "those residing in low-density areas are significantly more likely to undertake a multistop daily pattern."

Although some insights are already available, more detailed and extensive examination is desirable on the subject of life-style and residential location.

Analyses in this subject area have not advanced, possibly because of the conviction that there exist spatially invariant trip generation rates and that trip generation models are geographically transferable. Attempts to determine the effects of accessibility on trip generation have shown only minor results (79, 80). Thus, the conventional planning models implicitly assume that households of given characteristics will exhibit the same travel patterns no matter where they are located.

Still another problem is that many attributes of a metropolitan area, including transportation service levels and characteristics of household members and their preferences, are spatially correlated. Because of this, cases of ecological fallacy are the likely consequences of a marginal analysis of travel characteristics in which a limited number of contributing factors is controlled for. Because of these limitations, it is not possible to determine whether a particular travel pattern exhibited by a household is due to the household's life-style aspiration or to residence location and car ownership.

A very fundamental question is whether travel behavior is conditionally independent of life-style orientation, given residence location, car ownership, and other measurable factors. An answer to this question will determine the

Ancestry of this Project

RK dissertation:
taste variations
(Kitamura, 1981,
crediting Ben-Akiva,
Hensher, Stopher,
Golob, Tardiff, etc.)

Importance of
residential location
& lifestyle
(Kitamura, 1988)

PKL(M) dissertation:
importance of attitudes
to explaining travel behavior

CARB project:
impact of attitudes &
lifestyle on relation-
ship between built
environment and
travel behavior

Further work on
residential self-
selection (w/ Bagley,
Schwanen, Handy, Cao,
Circella)

Taste heterogeneity
in shopping channel
intention (w/ Tang)

Role of taste heterogeneity
in accounting for residential
self-selection

Overview

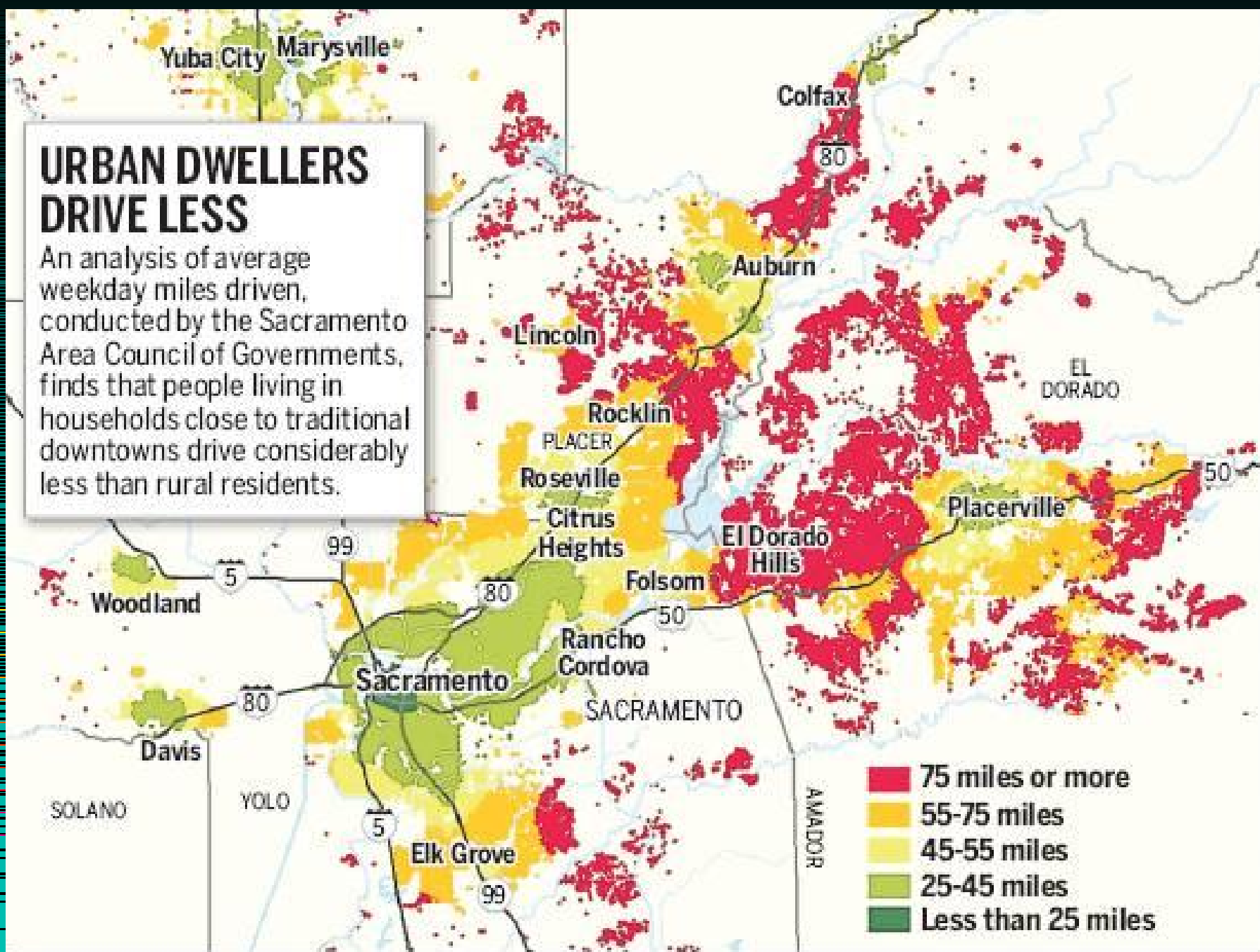
- n Motivation: The problem of residential self-selection
- n Past treatments of the problem
- n Why taste heterogeneity matters
- n Description of proposed research

Motivation

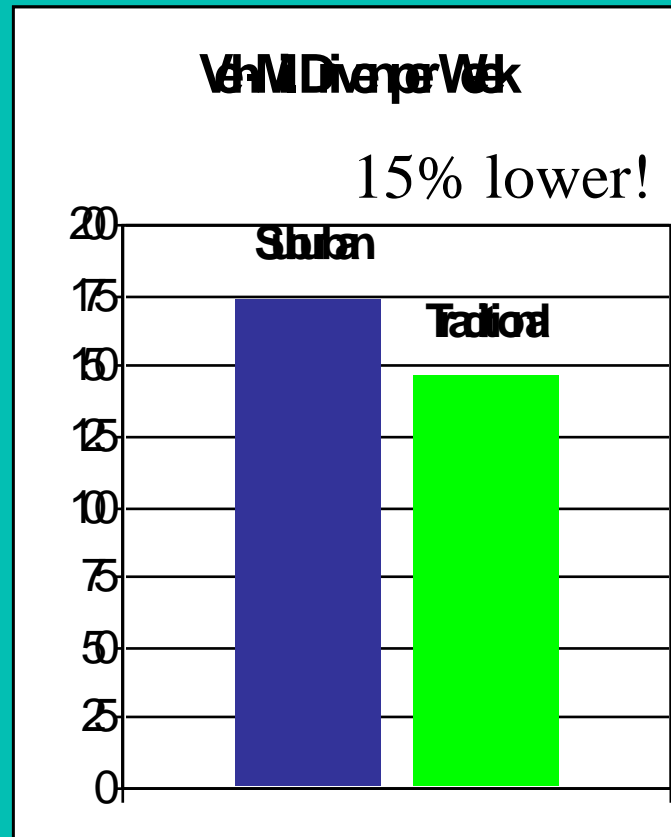
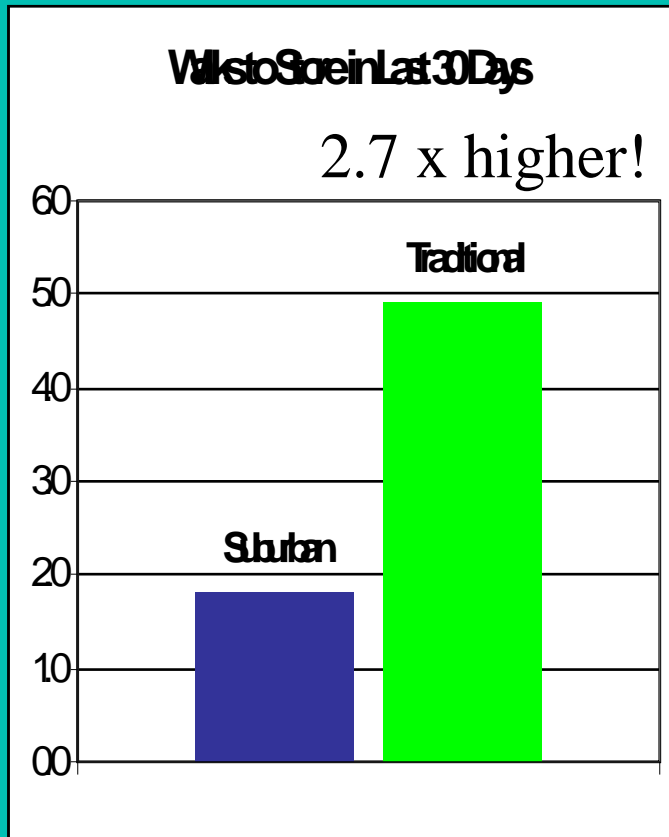
- n Many studies have compared travel behavior (TB) of residents of “suburban” versus “urban” neighborhoods
- n and found that **suburban dwellers walk less and drive more,**
- n supporting the rationale for more compact urban forms

URBAN DWELLERS DRIVE LESS

An analysis of average weekday miles driven, conducted by the Sacramento Area Council of Governments, finds that people living in households close to traditional downtowns drive considerably less than rural residents.

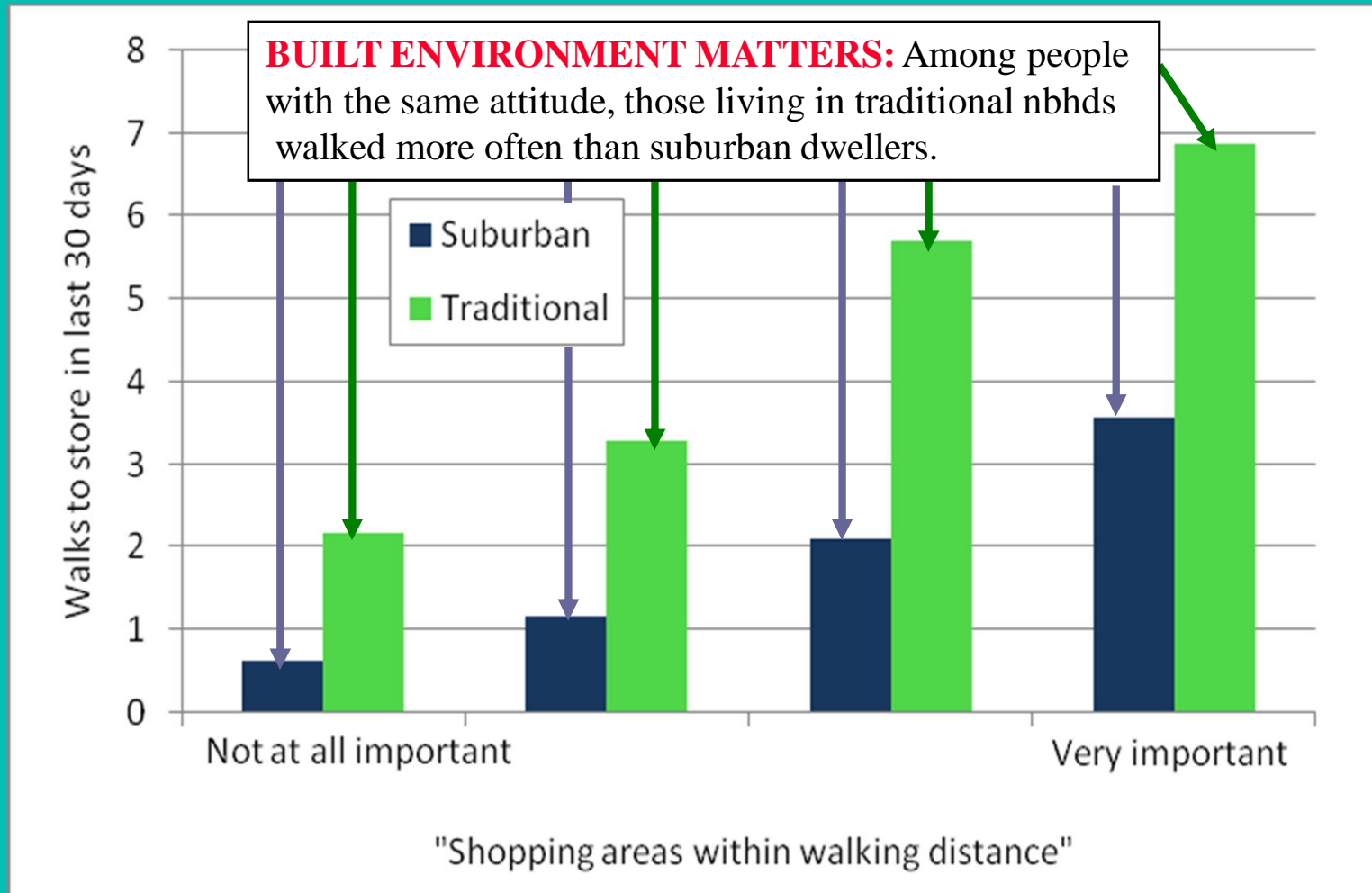


Another Northern California Example



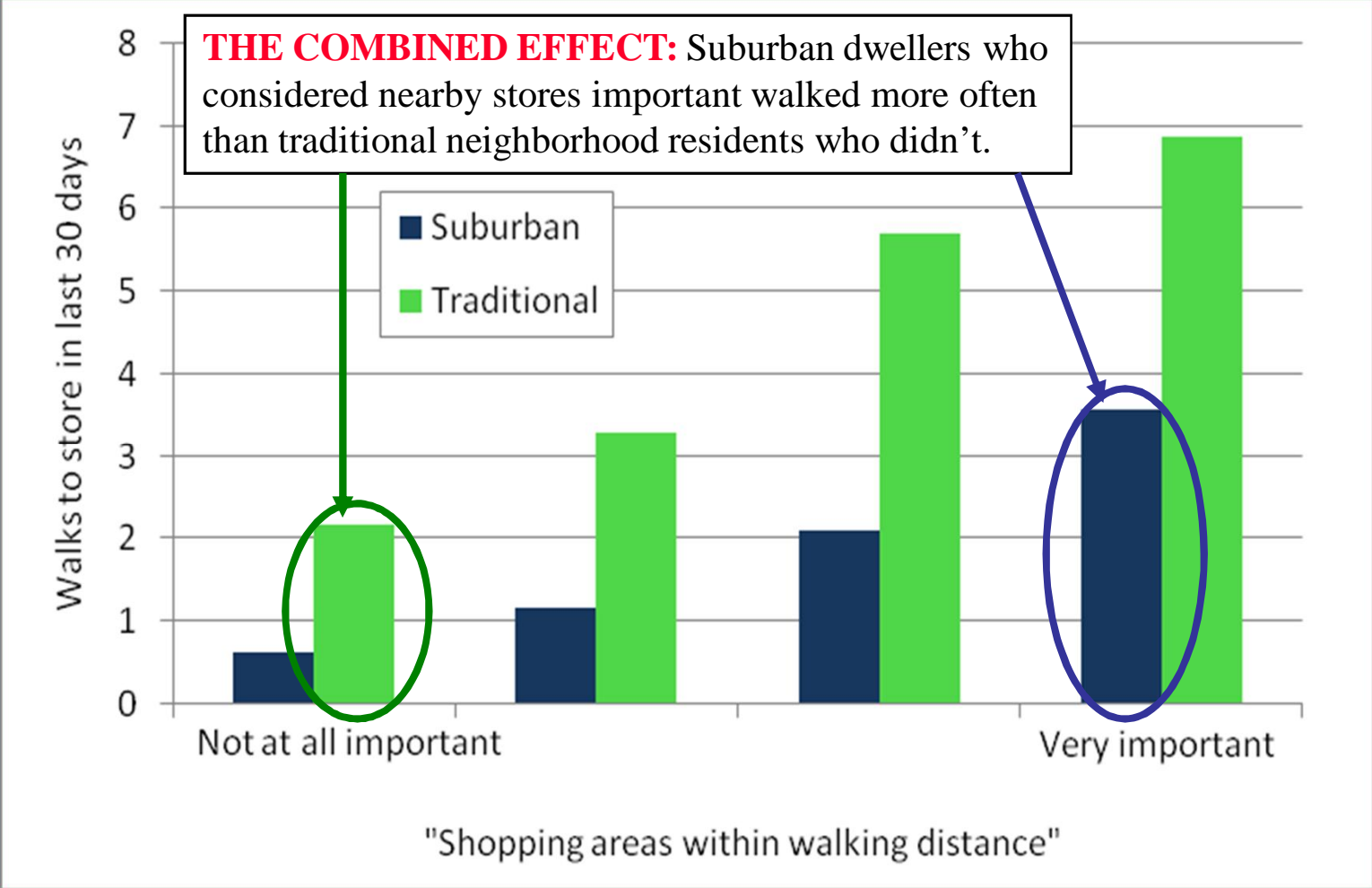
(Handy et al., 2005, 2006)

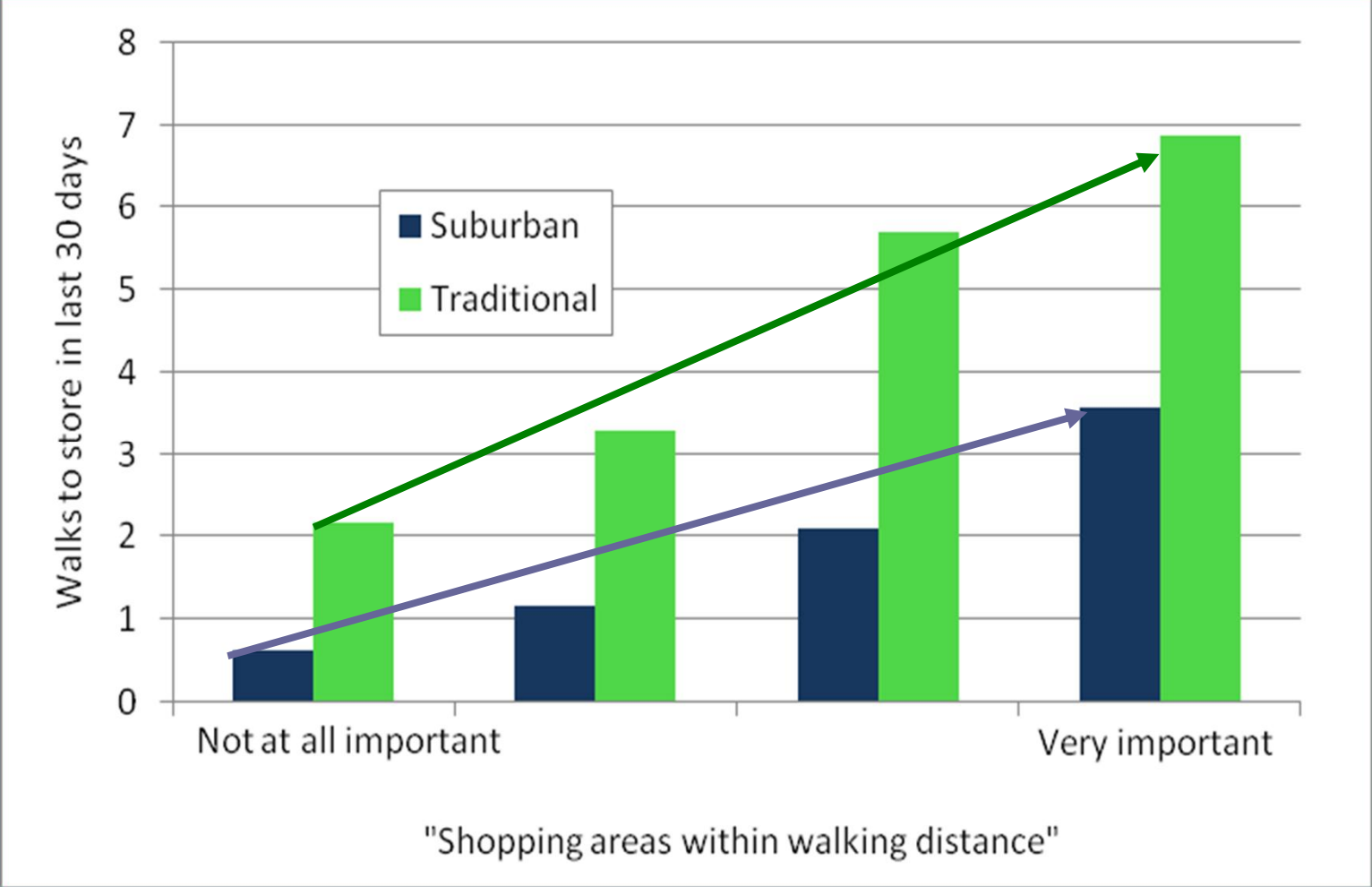
Great! What's the Problem?



ATTITUDE MATTERS: Among people living in the same type of nbhd, those who consider having nearby shops to be very important walked (~ 4x) more often than those who don't.







The Problem (cont'd)

- n So are the observed TB differences because of
 - a true independent influence of the built environment (BE)?
- or because
 - people who like walking (or, want to minimize driving) choose to live in neighborhoods supportive of that desire (AT)?
- or
 - some of both?

Hey, What's the Difference?!

- n *Objection 1:* Doesn't the built environment (BE) matter, either way?
 - "... if households self-select into areas that meet their travel preferences, it seems self-evident that urban structure matters" (Naess, 2009)
- n *Objection 2:* As long as there is unmet demand for urbanist neighborhoods, self-selection is irrelevant – we need to increase the supply

The Problem (cont'd)

- n If the effect of BE on TB is primarily due to attitudinal predispositions (AT), then policies promoting denser, more diverse land use patterns may not have the desired effect
- n For example, if a “car-lover” lands in an urban neighborhood because of policy incentives (e.g. financial), s/he may still drive like the typical suburban dweller

The Problem (cont'd)

- n Costs of being wrong:
 - potentially diminished quality of life
 - » inability to satisfy preferences
 - » disadvantages of crowding (e.g. lack of privacy, lack of children's play space/green space, congestion, tensions, contagion)
 - opportunity costs – time, money, & political capital could have been spent on more useful policies
- n Thus, to evaluate the effectiveness of (proposed) LU policies, it's important to know the relative roles of BE and AT in influencing TB

Nine Approaches for Addressing Self-selection

1. direct questioning
2. statistical controls (Kitamura et al., 1997)
3. instrumental variables models
4. sample selection models
5. propensity score models
6. joint discrete choice models
7. structural equations models
8. mutually-dependent discrete choice models
9. longitudinal designs

A Tenth Approach: Accounting for Taste Heterogeneity

- n Little-used to date
 - Exceptions: Bhat and co-authors (including Hensher, Pendyala), in the context of random-coefficient discrete choice models
- n Even less-discussed, at a conceptual level
 - Mentioned without elaboration by van Wee (2009)

What's the Problem (Mathematically)?

n Stereotypical relationship:

$$TB = \beta_0 + \beta_1 SED + \beta_2 BE + \varepsilon,$$

where

- **TB** is a measure of **travel behavior** (continuous in our application),
- **SED** represents (multiple) **socioeconomic/ demographic** variables, and
- **BE** represents (multiple) measures of the **built environment**

Stereotypical Form of RSS

- n Standard techniques (regression, discrete choice) require that observed variables (SED, BE) be uncorrelated with unobserved ones (ε)
- n Otherwise, the resulting *endogeneity bias* means that coefficient estimators of SED & BE will be *biased* and *inconsistent*

- n But if

$$TB = \beta_0 + \beta_1 SED + \beta_2 BE(AT) + \varepsilon(AT),$$

where **AT = attitudes**,

then this requirement is violated

Straightforward Remedy

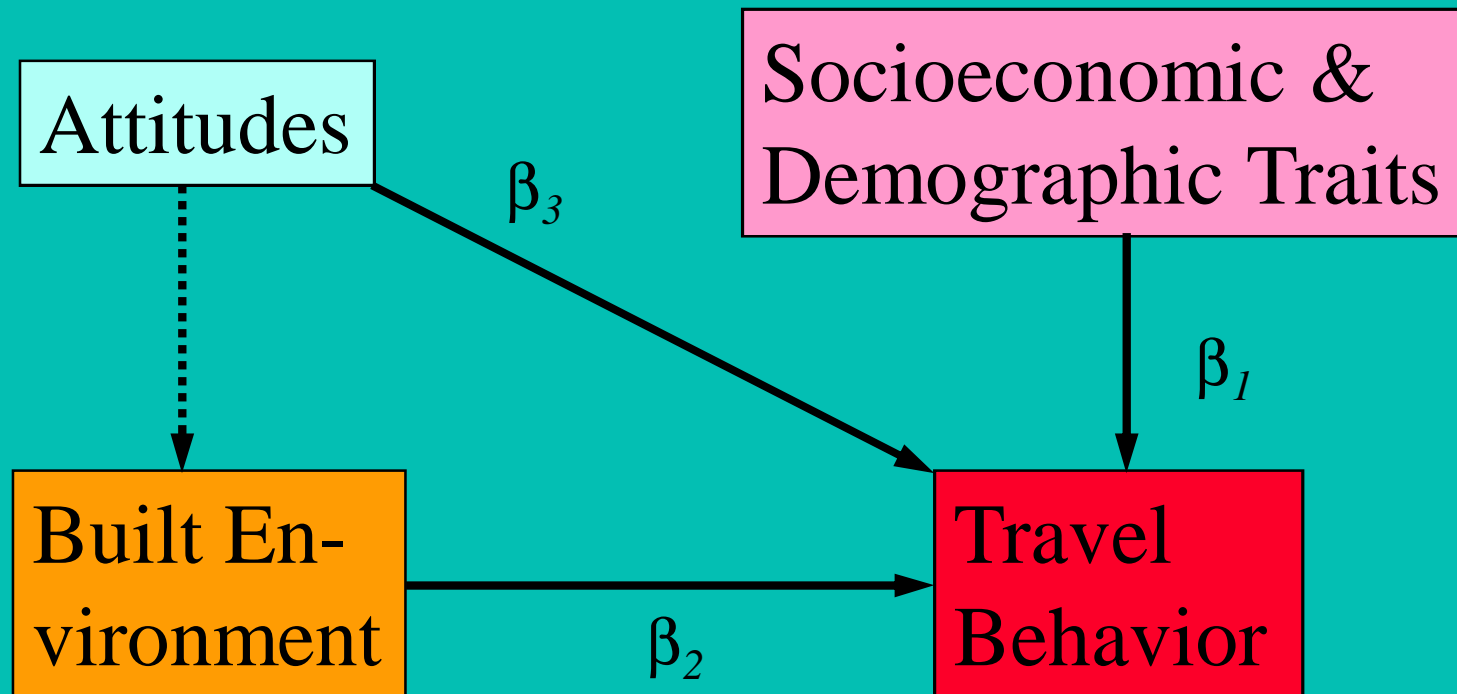
$$TB = \beta_0 + \beta_1 SED + \beta_2 BE + \beta_3 AT + \varepsilon$$

- n Observing & including AT removes them from ε and therefore removes correlation between BE and ε (“statistical controls”)
- n BE is called a *mediator* between AT & TB
- n Comparing β_2 with AT out of and in the model indicates separate importance of BE

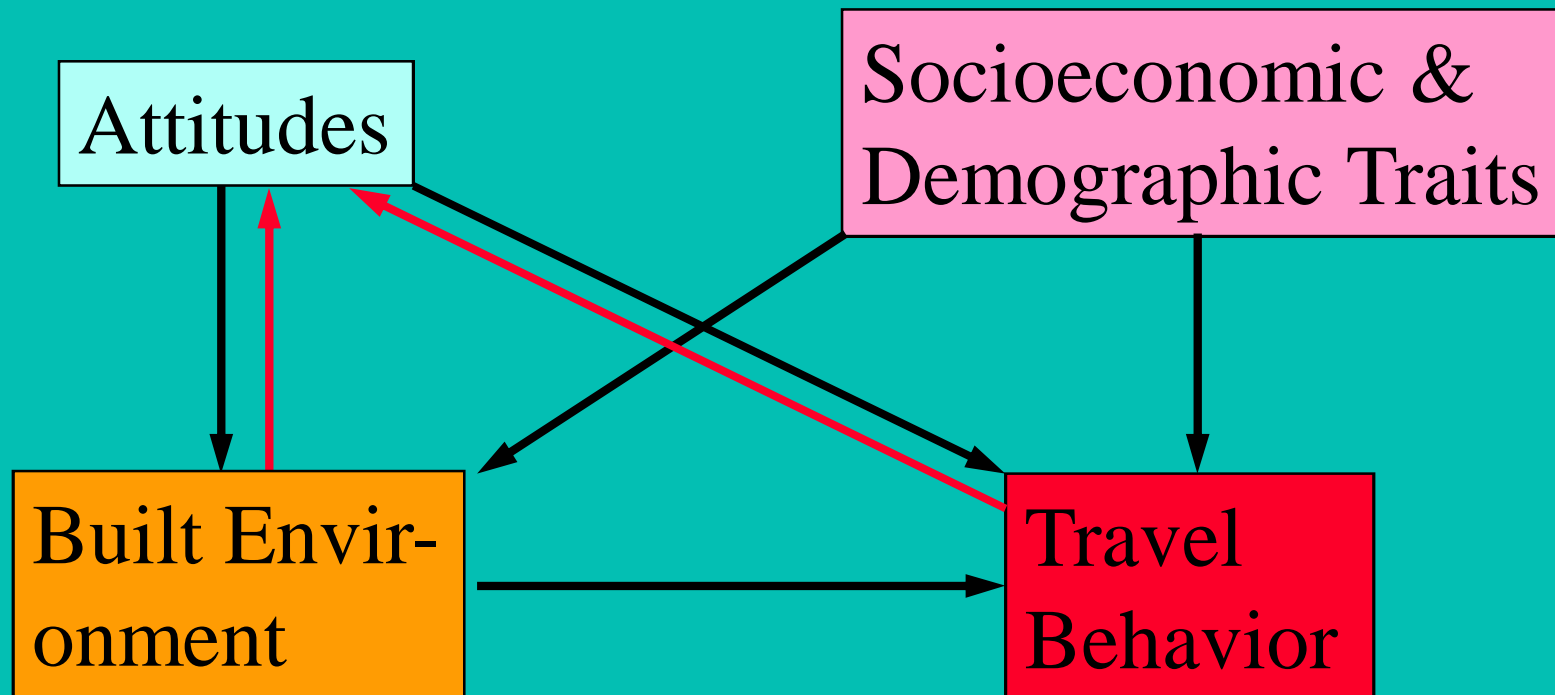
(Baron & Kenny, 1986)

(Kitamura et al., 1997)

Built Environment as a Mediator of Attitudes



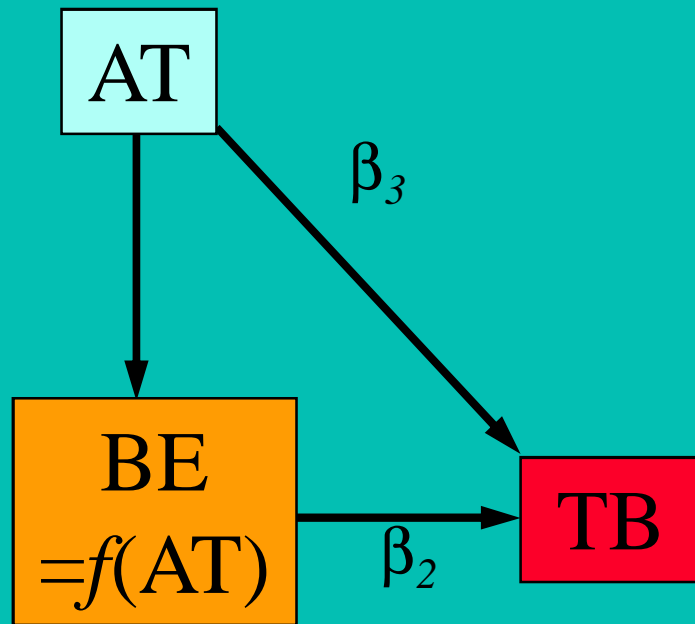
A More Complex Representation



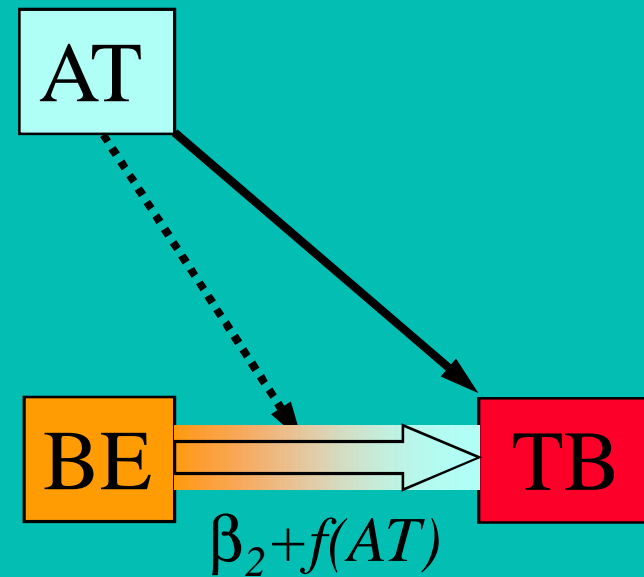
The Taste Heterogeneity Issue

- n But what if β is a function of AT, rather than (or as well as) BE being a function of AT?
 - The impact of the identical BE on TB (i.e. the value of β_2) could differ considerably between
 - » the person living there because she actively wants the transportation environment it provides, and
 - » the one living there because of other factors (job location, family members, location amenities)
- n AT called a *moderator* of the impact of BE (a “dimmer switch” on β_2 ; Wu & Zumbo, 2008)

BE as Mediator of AT



AT as Moderator of BE



In that Case, We Have...

$$TB_n = \beta_0 + \beta_1 SED_n + (\beta_2 + \eta_2^n) BE_n + \varepsilon_n$$

(where n sub/superscripts the individual),

n and if estimating the standard model, we have:

$$TB_n = \beta_0 + \beta_1 SED_n + \beta_2 BE_n + (\varepsilon_n + \eta_2^n BE_n)$$

n But this clearly results in correlation of BE and the error term,

n and if BE is also a function of AT, the correlation is even greater (since ζ_2^n , by assumption, is also a function of AT).

And It's Worth Noting...

- n ... that if taste heterogeneity is a factor, simple statistical controls won't resolve it:

$$TB_n = \beta_0 + \beta_1 SED_n + \beta_2 BE_n + \beta_3 AT_n + (\varepsilon_n + \eta_2^n BE_n)$$

- n still leaves us with

- an error term that includes BE,
- and the unobserved ζ_2^n correlated with the observed AT

- n So it's important to test for, and resolve, taste heterogeneity effects

Possible Approaches

- n Kitamura (1981, p. 473): “There are ... at least three approaches to the problem” of taste variations:
 1. random-coefficient models
 2. explicitly modeling the coefficients (creating interaction terms)
 3. market segmentation (“stratification”)
 - He considered deterministic segmentation
 - We will also explore latent class modeling

Our Proposed Study

- n Will use data previously collected from 8 neighborhoods in Northern California
- n Dependent (*TB*) variable: $\ln(\text{wklyVMD}+1)$
- n Explanatory variable categories:
 - *BE*: perceived (factors) & objective neighborhood traits
 - *AT*: travel-related factors, e.g. pro-bike/walk, pro-transit, car-dependent; and residential preferences for (e.g.) outdoor spaciousness & accessibility
 - *SED*: age, gender, employment status, HH size, income, etc.

Comparison of Three Specifications of Taste Heterogeneity

1. Modeling the coefficient(s) of BE, through creating interaction terms:
 - By adding the interaction term $AT_n \times BE_n$ to the model, we are effectively
 - replacing $\beta_2 BE_n$ with $(\beta_2^* BE_n + \beta_3 AT_n BE_n)$,
 - and thereby modeling β_2 as $(\beta_2^* + \beta_3 AT_n)$ (making it individual-specific, β_{2n})

Three Specifications of Taste Heterogeneity (cont'd)

2. Deterministically segmenting the sample on AT (and comparing BE coefficients across segments):
 - Will experiment with simple partitions of the sample with respect to one or two AT at a time (e.g. high v. low pro-bike/walk x high v. low outdoor spaciousness),
 - as well as cluster analysis to identify segments with similar bundles of ATs

Three Specifications of Taste Heterogeneity (cont'd)

3. Creating latent class regression models of VMD
 - Assumes the existence of G segments (classes) in the population, with segment-specific coefficients in the model
 - Class membership of a given person is unknown; a separate equation models membership (here, as a function of AT)
 - Segmented model coefficients are estimated so as to explain as much as possible about VMD

Basic Work Plan

- n Develop best baseline (unsegmented) models of VMD, with and without AT
- n Develop best segmentation models using each of the 3 methods just described
 - Compare goodness of fit, predictive ability
 - Analyze how the estimated impact of BE differs
- n For each model, quantify relative contributions of SED, BE, and AT to the explanation of VMD
- n Assess policy implications

Conclusions

- n Ryuichi's work has helped set the stage for this proposed project, as he stressed the importance of
 - Understanding the influence of residential location on TB
 - Accounting for taste heterogeneity in modeling TB
 - Understanding how attitudinal predispositions (residential self-selection) affect the impact of the BE on TB
- n I will feel him looking over our shoulders (in the friendliest possible way) as we conduct it!

References

- n **Bagley, M. N. and P. L. Mokhtarian** (2002) The impact of residential neighborhood type on travel behavior: A structural equations modeling approach. *Annals of Regional Science* **36(2)**, 279-297.
- n **Baron, R. M. and D. A. Kenney** (1986) The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* **51(6)**, 1173-1182.
- n **Bhat, C. R. and J. Y. Guo** (2007) A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. *Transportation Research B* **41**, 506-526.
- n **Cao, X., P. L. Mokhtarian, & S. L. Handy** (2009) Examining the impacts of residential self-selection on travel behavior: A focus on empirical findings. *Transport Reviews* **29(3)**, 359-395.
- n **Cao, X., P. L. Mokhtarian, & S. L. Handy** (2008) *Examining the Impacts of Residential Self-selection on Travel Behavior: Methodologies and Empirical Findings*. UC Davis Institute of Transportation Studies, Research Report UCD-ITS-RR-08-25.
http://pubs.its.ucdavis.edu/publication_detail.php?id=1194.
- n **Handy, S., X. Cao, and P. L. Mokhtarian** (2005) Correlation or causality between the built environment and travel behavior? Evidence from Northern California. *Transportation Research Part D* **10**, 427-444.
- n **Handy, S., X. Cao, & P. L. Mokhtarian** (2005) Self-selection in the relationship between the built environment and walking: Empirical evidence from Northern California. *Journal of the American Planning Association* **72(1)**, 55-74.
- n **Kitamura, R.** (1981) A stratification analysis of taste variations in work trip mode choice. *Transportation Research A* **15**, 473-485.

References (cont'd)

- n **Kitamura, R.** (1988) Life style and travel demand. In *A Look Ahead, Year 2020, Special Report 220*, Transportation Research Board, Washington, D.C., pp. 149-189.
- n **Kitamura, R., P. L. Mokhtarian, and L. Laidet** (1997) A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area. *Transportation* **24(2)**, 125-158.
- n **Mokhtarian, P. L. and X. Cao** (2008) Examining the impacts of residential self-selection on travel behavior: A focus on methodologies. *Transportation Research Part B* **43(3)**, 204-228.
- n **Pinjari, A. R., C. R. Bhat, & D.A. Hensher** (forthcoming) Residential self-selection effects in an activity time-use behavior model. *Transportation Research Part B*.
- n **Pinjari, A. R., N. Eluru, C. R. Bhat, R. M. Pendyala, and E. Spissu** (2008) Joint model of choice of residential neighborhood and bicycle ownership: Accounting for self-selection and unobserved heterogeneity. *Transportation Research Record* **2082**, 17-26.
- n **Pinjari, A. R., R. M. Pendyala, C. R. Bhat, & P. A. Waddell** (2007) Modeling residential sorting effects to understand the impact of the built environment on commute mode choice. *Transportation* **34(5)**, 557-573.
- n **Tang, W. & P. L. Mokhtarian** (2009) Accounting for taste heterogeneity in purchase channel intention modeling: An example from Northern California for book purchases. Paper presented at the International Choice Modeling Conference, Leeds, UK, March 30-April 1, 2009.
- n **van Wee, B.** (2009) Self-selection: A key to a better understanding of location choices, travel behaviour, and transport externalities? *Transport Reviews* **29(3)**, 279-292.
- n **Walker, J. and J. Li** (2007) Latent lifestyle preferences and household location decisions. *Journal of Geographical Systems* **9**, 77-101.
- n **Wu, A. D. and B. D. Zumbo** (2008) Understanding and using mediators and moderators. *Social Indicators Research* **87**, 367-392.

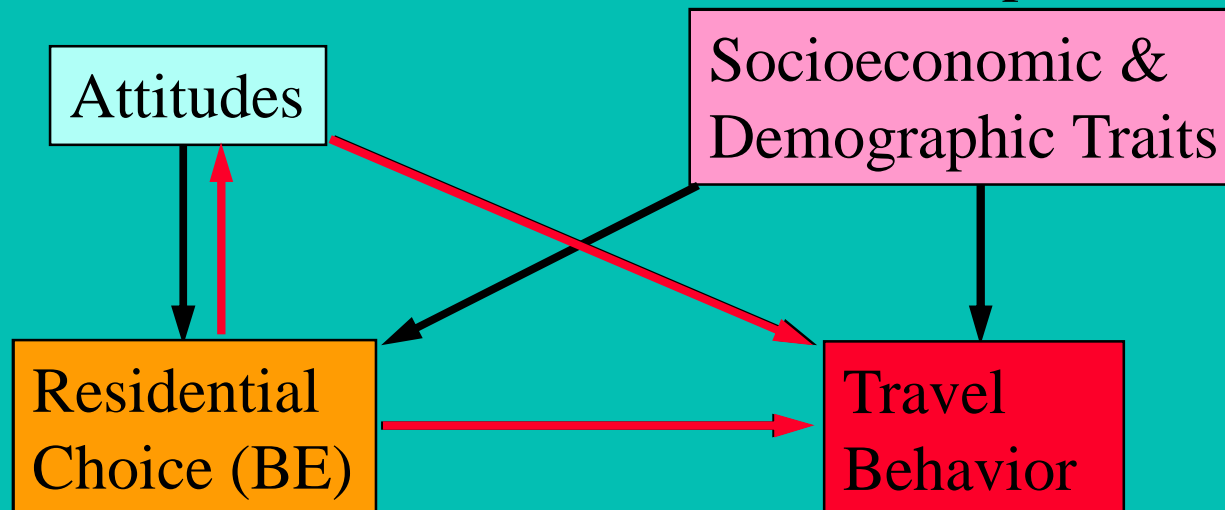
Questions?



plmokhtarian@ucdavis.edu
www.its.ucdavis.edu/telecom/

Some Policy Implications

- n There is some benefit to motivating people to live in more urbanist areas, even if they are attitudinally mismatched
 - The BE does have an independent effect on TB
 - And it may also, over time, affect attitudes to be more urbanist, which in turn have a direct impact on TB



Policy Implications (cont'd)

- n And changing attitudes to be more urbanist/less auto-oriented is almost surely a “two-fer”:
- ATs have a direct impact on TB, even if residentially mismatched
 - Urbanist ATs increase the likelihood of living in an urbanist area, which then exerts an additional influence on TB

