Center for Teaching Old Models New Tricks (TOMNET)

A USDOT Tier 1 University Transportation Center

PROJECT PROPOSAL

Title: Assembling Integrated Data Sets for Analyzing Connections between Travel Behavior, Attitudes, and the Built Environment

Principal Investigator: Deborah Salon, Assistant Professor, School of Geographical Sciences and Urban Planning, Arizona State University

---

1. Introduction/Problem Statement

Important travel behavior outcomes include distance, mode, route, and time of day for a particular trip, as well as number of trips taken overall by an individual or a household. Observed travel behavior depends on some combination of attitudes and preferences about these factors; attitudes and preferences about vehicle ownership, vehicle type, and residential location; and physical (both individual capability and the built environment) and budget constraints (Figures 1 & 2).
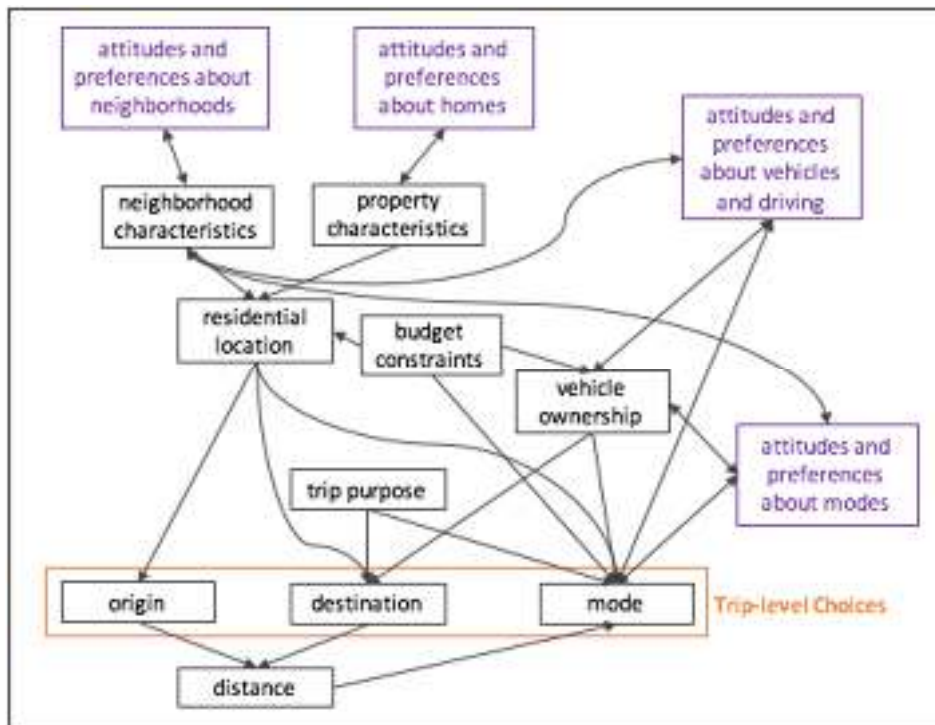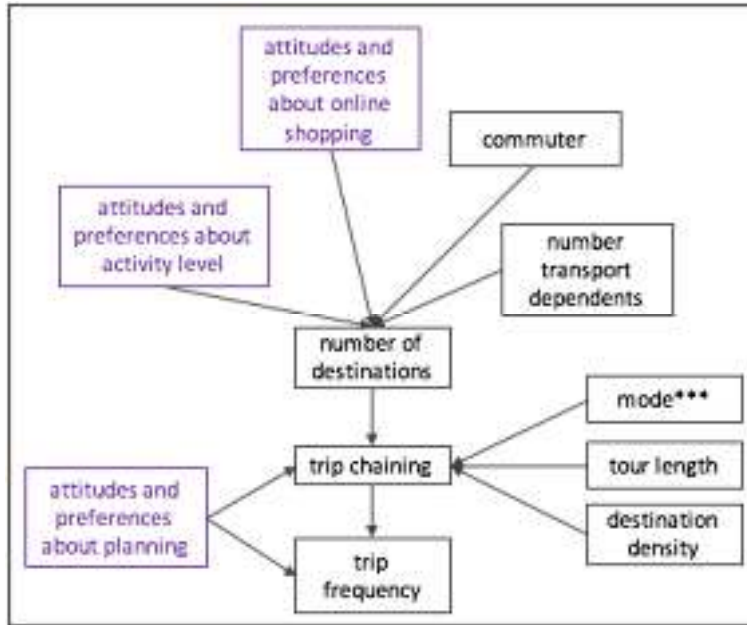

Figure 1: Simplified Conceptual Model of the Determinants of Trip Distance and Mode Choices

*** Mode is both affected by and affects multiple other factors - see Figure 1
Figure 2: Simplified Conceptual Model of the Determinants of Trip Frequency

Travel surveys -- that are used to quantify these relationships and form the basis for regional travel demand models -- contain only a portion of the relevant variables. Most travel diary-style datasets include details of trip origins, destinations, modes, time of day, purpose, frequency, and vehicle ownership. Many also include residential neighborhood information, and socioeconomic and demographic data about the households. Few include attitudes and preferences at all. If included, attitudes and preferences about modes are the most common. We are not aware of any travel diary-style datasets that include detailed home property characteristics and attitudes and preferences about neighborhoods and homes.

Separate surveys do exist that include questions regarding attitudes and preferences about modes, vehicles, and neighborhoods. However, these detailed attitudinal surveys do not collect detailed travel behavior information, chiefly based upon the (probably correct) concern that the respondent burden would be too great.

2. Project Objectives

The proposed project will begin to address this data gap by fusing existing datasets that are rich in attitudinal information with vehicle ownership and use, home neighborhood, and individual property characteristics datasets. These combined datasets will allow us to investigate questions about how attitudes influence residential location choice, and how the combination of attitudes and location choice influences travel behavior. The development of fused datasets will be the deliverable from Year 1 of this project. We will propose to analyze these data in Year 2. Below is a brief preview of the research questions these data will help us address.

We hypothesize that much of the influence of attitudes on travel behavior may come from attitudes affecting location choice -- whether or not those attitudes are related to travel (Zhang, 2014) -- rather than attitudes directly affecting travel choices such as vehicle ownership, trip frequency, and mode choices. Thus, establishing the link between attitudes and location choice is a critical step to untangling the role of attitudes in travel behavior.

Other researchers have found significant relationships between attitudes and residential location in data-limited contexts. For example, Bagley and Mokhtarian (1999) found that attitudes are associated with the choice of urban or suburban neighborhoods. Fu et al. (2014) found attitudes to be significantly related to a number of variables reflecting residential location. Neither of these models estimate

residential location directly, however. Instead, they estimate properties of the residential location (e.g. whether it is an urban neighborhood, or housing type). The dataset produced by this project would allow more detailed models of residential location choice.

Walker and Li (2007) found that estimating a joint latent class-residential location choice model substantially improved predictive accuracy. While their model did not explicitly include attitudes, subsequent work has included them (e.g., Liao, Farber, and Ewing, 2015). With the dataset developed by this project, we, too, can explicitly include attitudes rather than assume they can be captured by choices and sociodemographics.

In our first analysis, we will estimate joint models of residential choices and travel behavior. These will take a variety of forms, beginning by replicating the work of Salon (2009) and Salon (2015) using datasets that include attitudes. Salon (2009) estimated a joint multinomial logit model of residential location, car ownership, and commute mode, while Salon (2015) estimated a joint multinomial logit/regression model of neighborhood type choice and vehicle miles traveled. Both prior projects focused on identifying sources of heterogeneity in the estimated relationships, but were somewhat limited because attitudinal variables were not available. The goal of the proposed analysis will be to gain an improved understanding of how attitudes affect residential choices, how they affect travel choices, and how these choices relate to each other when a robust set of attitudes are factored in.

In our second analysis, we will make use of detailed property characteristics together with homeowner sociodemographics and attitudinal factors to estimate a hedonic model of housing prices. The vast majority of hedonic studies include only property and neighborhood characteristics, and do not include any information about the people who chose to purchase the property. Kestens et al. (2005) is a rare exception, with hedonic models that allow estimates to vary according to sociodemographic characteristics of the household. We are not aware of any existing hedonic models that interact attitudinal factors with property characteristics.

The rich datasets of attitudes, location attributes, and travel behavior produced by this project will be made available to other TOMNET researchers for use in other projects as well.

3. Proposed Methodology and Data

The proposed work is to directly merge data from multiple sources to create a fused dataset that includes attitudes, household vehicle ownership and use information, property characteristics for the home, and neighborhood characteristics. If data are missing for certain households, we will use multiple imputation methods to create complete datasets.

We will begin with observations in three attitudinal surveys previously collected by TOMNET team members: the ASU Travel Survey (2012), the Northern California Multitasking Survey (2011), and the California Millennials Dataset (2015). We will then use home location and vehicle make, model, and model year information provided in each survey to merge these observations with odometer-based vehicle miles traveled (VMT) estimates, and neighborhood and property characteristics information.

The VMT estimates for the California surveys will be derived from the California Department of Motor Vehicles registration database together with its Smog Check program emissions test database, administered by the California Bureau of Automotive Repair. The VMT estimates for the Arizona survey will come from the equivalent state agencies and databases in this state: the Arizona Motor Vehicles Division of the state's Department of Transportation, and the Arizona Department of Environmental Quality, which administers the state's vehicle emissions testing program. Note that we are not yet certain that we will have access to the Arizona data required to estimate VMT. However, even without this data for Arizona, we will still produce a dataset useful for modeling residential choice and daily travel demand.

Neighborhood characteristics will come from the EPA's Smart Location database and from the US Census. Home property characteristics will come from the relevant county's assessor or from a company that aggregates county assessor data in California, depending on the county. In either case,

property characteristics will include variables such as number of bedrooms, home size, lot size, recording date, assessed value, and in many cases, the sale price.

The link that will connect these data sources will be the home addresses of the attitudinal survey respondents. Some survey respondents provided exact home addresses, while others provided cross streets near their homes. For those who provided cross streets, we will use vehicle registration information to identify their home addresses. We expect that this will be possible for most surveyed households by matching the reported make, model, and model year of their vehicles to vehicle registration addresses near the cross streets that they reported in the survey. In order to preserve privacy, once the datasets have been merged, exact home addresses for these respondents will be removed.

4. Work Plan (Project Tasks)

Task 1: Institutional Review Board Approval
This task will consist of creating a project plan and submitting it to ASU's Institutional Review Board (IRB). IRB approval is needed for this project as it will involve the personal information of survey participants from three existing attitudinal surveys. Because the datasets already exist and were collected by TOMNET team members, we are hopeful that the IRB approval process will be swift.

Task 2: Use California DMV registration data to identify home addresses for respondents to the Northern California Multitasking Survey and the California Millennials Dataset
In this task, we will use vehicle registration addresses together with matched vehicle characteristics and near-home cross streets to pinpoint actual home addresses for attitudinal survey respondents. These addresses will be used only so that property and neighborhood characteristics can be added to the dataset. After the merge is completed, the home addresses will be dropped to preserve respondent privacy.

This and all tasks using California DMV registration data require partnership with our colleague Jeffrey Williams at the University of California, Davis who has access to these data. Salon has partnered with Williams to access these data for two previous projects (Cook et al., 2015; Salon et al., under review), and Williams has expressed willingness to partner again for the purposes of this project.

Task 3: Obtain assessor property characteristics for home addresses in the Northern California Multitasking Survey and the California Millennials Dataset
This task requires exactly matching addresses from the attitudinal surveys with assessor information about the properties. Matching addresses can be extremely difficult due to text formatting concerns, so this process will be time consuming, and may include the development of county-specific tools in order to interface with the datasets and tools available in each county.

For some California counties (e.g., Los Angeles, Ventura, San Francisco), the assessor property characteristics database is available to the research team at low or no cost. For most California counties, however, the assessor information is available only for a fee from a private company ($2/record, plus monthly access fee). We have included this expense in our project budget.

Task 4: Use Smog Check and registration odometer readings to estimate VMT for all vehicles in each surveyed household in California
For this task, we will use a time series of Smog Check and registration odometer readings to estimate annual vehicle miles traveled (VMT) for each vehicle in the California surveyed households. Odometer readings and dates are recorded during emissions tests and when vehicles are sold. Creating time series of readings using unique vehicle identification numbers (VINs) will allow us to estimate annual number of miles that each vehicle was driven. Salon has experience with this task, as she estimated VMT this way for two previous projects (Cook et al., 2015; Salon et al., under review).

Task 5: Fuse Maricopa County assessor property characteristics information with ASU Travel Survey dataset

The research team has access to the assessor property characteristics data for Maricopa County. The home locations of respondents to the ASU Travel Survey have been geocoded. If geocoded home locations correspond directly to home parcels, merging the data will be straightforward. In some cases, however, geocoded home locations may correspond to cross streets rather than actual addresses. In these cases, we will use average nearby parcel characteristics to represent the home that the respondent actually lives in. Many neighborhoods in the Phoenix area are homogeneous in terms of parcel size and home characteristics, in part because they were built by a single developer. Thus, average nearby parcel characteristics should be a good approximation of actual home parcel characteristics.

Task 6: If we can gain access, fuse ASU Travel Survey information with AZ MVD and AZDEQ data
This task requires exactly the same steps as Task 4, once the data are in-hand. Gaining access to these data requires a partnership with a colleague at ASU who has access, which Salon is currently negotiating.

Task 7: Clean fused dataset and impute missing data
The final task required to produce three complete datasets will be data checking/cleaning and missing data imputation. To check the data, we will look at the distributions of each included variable, and examine any clear outliers individually to make sure there appear to be no obvious errors that might substantially sway analysis results. For missing data imputation, we plan to use the method of multiple imputation (Rubin, 1987).

5. Project Schedule

Task	2017					2018						
	Aug	Sept	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June	July
1	▦	▦										
2		▦	▦	▦	▦							
3	▦	▦	▦	▦								
4				▦	▦	▦						
5						▦	▦	▦				
6								▦	▦	▦		
7									▦	▦	▦	▦

6. Relevance to the Center Theme/Mission
By creating enhanced travel behavior-residential location-attitudinal datasets, the proposed project is central to TOMNET's mission of "advancing data ... to reflect role of attitudes, perceptions, values, and preferences in activity-travel behavior and mobility choice models".

7. Anticipated Outcomes and Deliverables
The main deliverables of this project are three final datasets that integrate information on attitudes, travel behavior and residential location, each using a different attitudinal survey as its base. We plan to both use

these data for our own future analysis, and to share them with other TOMNET partners for future research. Beyond TOMNET, we hope to eventually be able to share versions of these datasets more broadly -- after stripping them of information that could individually identify respondents.

More generally, the anticipated outcome of this project is an improved understanding of how attitudes affect residential location choice, as detailed in section 2 of this proposal.

8. Research Team and Management Plan

The research team for this project will include Matthew Wigginton Conway and Deborah Salon. Conway is beginning his PhD studies in the School of Geographical Sciences and Urban Planning at ASU in August 2017, and Salon is both a faculty member in the School and Associate Director of TOMNET.

Conway has experience in transportation planning, accessibility modeling, and computer programming. Salon has experience in travel behavior research and analysis. Both team members have substantial experience working intensively with varied transportation-related datasets, including the assembly of fused datasets such as this work plan will develop.

Salon and Conway will meet regularly to ensure timely delivery of the project.

9. Technology Transfer Plan

In future years, this project will result in at least two peer-reviewed journal articles; one on the joint modeling of location choice and travel demand, and one on the variation of house valuations due to attitudes. If funds permit, these articles will be published as open access and/or deposited into the ASU Digital Repository (https://repository.asu.edu) to facilitate application and reuse outside of the academic realm.

10. Workforce Development and Outreach Plan

One of the main team members on this project will be Conway, who will be a PhD student in the School of Geographical Sciences and Urban Planning at ASU. Conducting this data fusion project and subsequent data analysis will provide useful training for him, and may become part of his PhD dissertation work. So that other graduate students in SGSUP may benefit from the project, we will share progress with a transportation research working group that meets weekly during the academic year. Once the data are analyzed, we plan to present findings at major transportation and geography conferences.

11. References

Bagley, M.N., & Mokhtarian, P.L. (1999). The Role of Lifestyle and Attitudinal Characteristics in Residential Neighborhood Choice. In *Transportation and Traffic Theory*, edited by Avishai Ceder, 1–15. https://escholarship.org/uc/item/22g7w9pp, accessed 7 May 2017.

Cook, J.A., Sanchirico, J.N., Salon, D., & Williams, J. (2015). Empirical distributions of vehicle use and fuel efficiency across space: Implications of asymmetry for measuring policy incidence. *Transportation Research Part A: Policy and Practice* 78, 187-199.

Fu, X., Bhat, C.R., Pendyala, R.M., Vadlamani, S., & Garikapati, V.M. (2014). Understanding the Multiple Dimensions of Residential Choice, July. The University of Texas at Austin. https://repositories.lib.utexas.edu/bitstream/handle/2152/26195/MultipleDimensionsofResidentialChoiceICLV.pdf?sequence=2&isAllowed=y, accessed 7 May 2017.

Kestens, Y., Thériault, M. & Des Rosiers, F. (2005). Heterogeneity in Hedonic Modelling of House Prices: Looking at Buyers' Household Profiles. *Journal of Geographical Systems*, 8(1), 61–96. doi:10.1007/s10109-005-0011-8.

Liao, F.H., Farber, S., & Ewing, R. (2015). Compact development and preference heterogeneity in residential location choice behaviour: A latent class analysis. *Urban Studies,* 52(2), 314-337.

Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley.

Salon, D. (2009). Neighborhoods, cars, and commuting in New York City: A discrete choice approach. *Transportation Research Part A: Policy and Practice*, *43*(2), 180-196.

Salon, D. (2015). Heterogeneity in the relationship between the built environment and driving: Focus on neighborhood type and travel purpose. *Research in Transportation Economics*, *52*, 34-45.

Salon, D., Blumenberg, E., Thomas, T., & Williams, J. Transportation and housing affordability in Los Angeles County. *Under review.*

Walker, J.L., & Li, J. (2007). Latent Lifestyle Preferences and Household Location Decisions. *Journal of Geographical Systems,* 9 (1): 77–101. doi:10.1007/s10109-006-0030-0.

Zhang, J. (2014). Revisiting residential self-selection issues: A life-oriented approach. *Journal of Transport and Land Use, 7*(3), 29-45. doi:10.5198/jtlu.v7i3.460.

12. Budget Including Non-Federal Matching Funds

Institution: ASU
Project Title: Assembling Integrated Data Sets for Analyzing Connections between Travel Behavior, Attitudes, and the Built Environment
Principal Investigator: Deborah Salon
Budget Period: 8/1/2017 - 07/31/2018

CATEGORY	Budgeted Amount from Federal Share	Budgeted Amount from Matching Funds	Explanatory Notes; Identify Source of Matching Funds
Faculty Salaries	$4,379.50	$8,319.50	0.5 month summer salary CS: Salon 5% AY effort + 0.5 month summer salary
Other Staff Salaries	$-	$-	
Student Salaries	$8,000.00	$6,250.00	SNR Fellowship, SGSUP
Fringe Benefits	$1,577.47	$2,683.77	
Total Salaries & Benefits	$13,956.97	$17,253.27	
Student Tuition Remission	$4,044.00	$4,044.00	SNR Fellowship, SGSUP
Operating Services and Supplies	$-	$-	
Domestic Travel	$2,000.00	$-	
Other Direct Costs	$-	$-	
Other Direct Costs	$-	$-	
Total Direct Costs	$20,000.97	$21,297.27	
F&A (Indirect) Costs	$9,095.47	$9,834.36	
TOTAL COSTS	$29,096.44	$31,131.63	

Grant Deliverables and Reporting Requirements for UTC Grants (November 2016)
Exhibit F

UTC Project Information	
Project Title	Assembling Integrated Data Sets for Analyzing Connections between Travel Behavior, Attitudes, and the Built Environment
University	Arizona State University
Principal Investigator	Deborah Salon
PI Contact Information	deborah.salon@asu.edu
Funding Source(s) and Amounts Provided (by each agency or organization)	TOMNET: $29,096 ASU, SGSUP: $31,132
Total Project Cost	$60,228
Agency ID or Contract Number	
Start and End Dates	8/1/2017 - 7/31/2018
Brief Description of Research Project	Travel survey datasets used for transportation system planning do not routinely include variables relating to attitudes and preferences, nor do they include much information pertinent to understanding residential location choice. The proposed project will begin to address this data gap by fusing existing datasets that are rich in attitudinal information with vehicle ownership and use, home neighborhood, and individual property characteristics datasets. These combined datasets will allow us to investigate questions about how attitudes influence residential location choice, and how the combination of attitudes and location choice influences travel behavior. The development of fused datasets will be the deliverable from Year 1 of this project. We will propose to analyze these data in Year 2.
Describe Implementation of Research Outcomes (or why not implemented) Place Any Photos Here	TBD
Impacts/Benefits of Implementation (actual, not anticipated)	TBD
Web Links · Reports · Project Website	TBD

DEBORAH SALON, Ph.D.
Assistant Professor, School of Geographical Sciences and Urban Planning,
Arizona State University, Tempe, AZ 85287-3005. Ph: (480) 965-7475; Email:
deborah.salon@asu.edu

EDUCATION
- University of California at Davis, Davis, CA, USA
  - Ph.D., Agricultural and Resource Economics, May 2006
- Carleton College, Northfield, MN
  - B.A., Physics, June 1994

PROFESSIONAL EXPERIENCE
- Arizona State University
  - Assistant Professor, School of Geographical Sciences and Urban Planning, 2014-present
  - Graduate Faculty, School of Sustainability, 2016-present
  - Senior Sustainability Scientist, Global Institute of Sustainability, 2014- present
- University of California, Davis, Institute of Transportation Studies
  - Professional Researcher, 2008-2014
- The Earth Institute at Columbia University
  - Post-Doctoral Fellow, 2006-2008

RELEVANT REFEREED PUBLICATIONS (Total: 17 Refereed Publications)
1. Salon, Deborah. (2015) Heterogeneity in the relationship between the built environment and driving: Focus on neighborhood type and travel purpose. *Research in Transportation Economics*, 52, 34-45.
2. Cook, Jonathan, James Sanchirico, Deborah Salon, and Jeffrey Williams. (2015) Empirical distributions of vehicle use and fuel efficiency across space: Implications of asymmetry for policy analysis. *Transportation Research Part A: Policy and Practice*, 78, 187-199.
3. Salon, Deborah, Marlon Boarnet, Susan Handy,Steven Spears, and Gil Tal. (2012) How do local actions affect VMT? A critical review of the empirical evidence. *Transportation Research Part D* 17(7): 495-508.
4. Salon, Deborah. (2009) Neighborhoods, cars, and commuting in New York City: A discrete choice approach. *Transportation Research Part A: Policy and Practice* 43(2): 180-196.

RELEVANT RESEARCH PROJECTS (Total Sponsored Research: ~ $600,000)
- *A Spatial Analysis of Housing and Transportation Affordability in Los Angeles County*, University of California Transportation Center, 2012-2015
- *Quantifying the effect of local government actions on VMT*, California Air Resources Board, 2010-2014

JOURNAL EDITORIAL ACTIVITIES
- CO- EDITOR OF SPECIAL ISSUE, *RESEARCH IN TRANSPORTATION ECONOMICS* (ELSEVIER), 2015
- EDITORIAL BOARD, JOURNAL OF *TRANSPORTATION GEOGRAPHY* (ELSEVIER), 2016-present
- EDITORIAL BOARD, TRANSPORTATION RESEARCH PART D (ELSEVIER), 2017-present

EDUCATION AND STUDENT ADVISING
- Thesis/Dissertation Major Advisor/Chair: 3 MS (Thesis) students completed, 1 MS (Thesis) student in progress
- Thesis/Dissertation Committee Member: 2 PhD students completed; 5 MS (Thesis) students completed