

# Comparison of Social Capital Indicators from Position Generators and Name Generators in Modeling Activity Selection

---

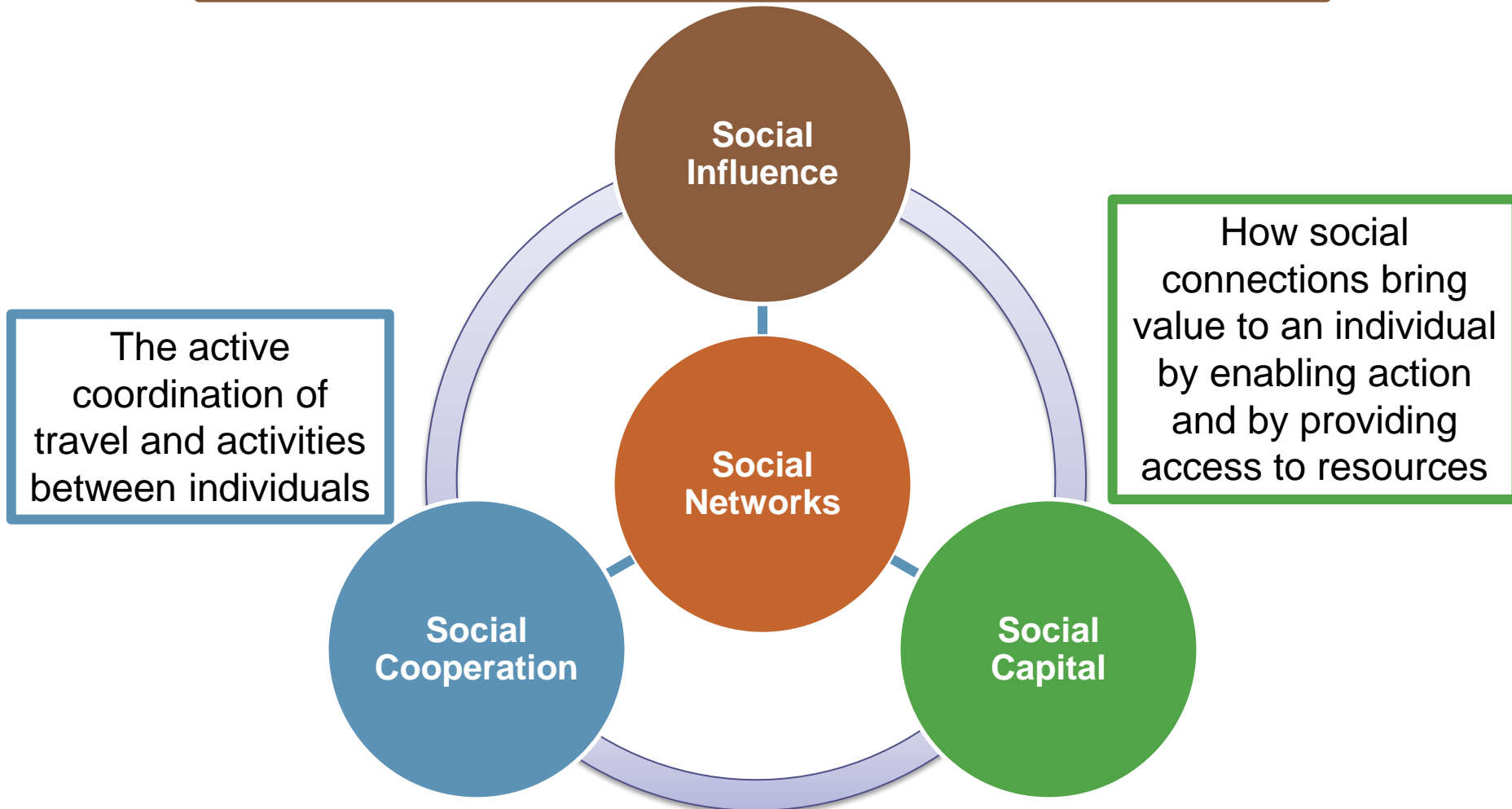
Michael Maness

University of Maryland  
Frontiers in Transportation  
July 24, 2015

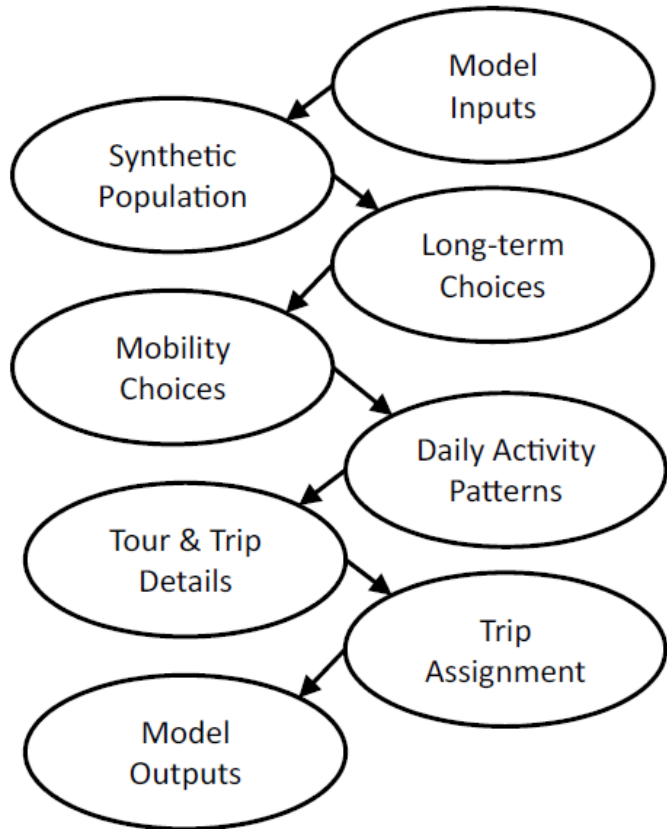


# Social Interactions in Activity-Travel

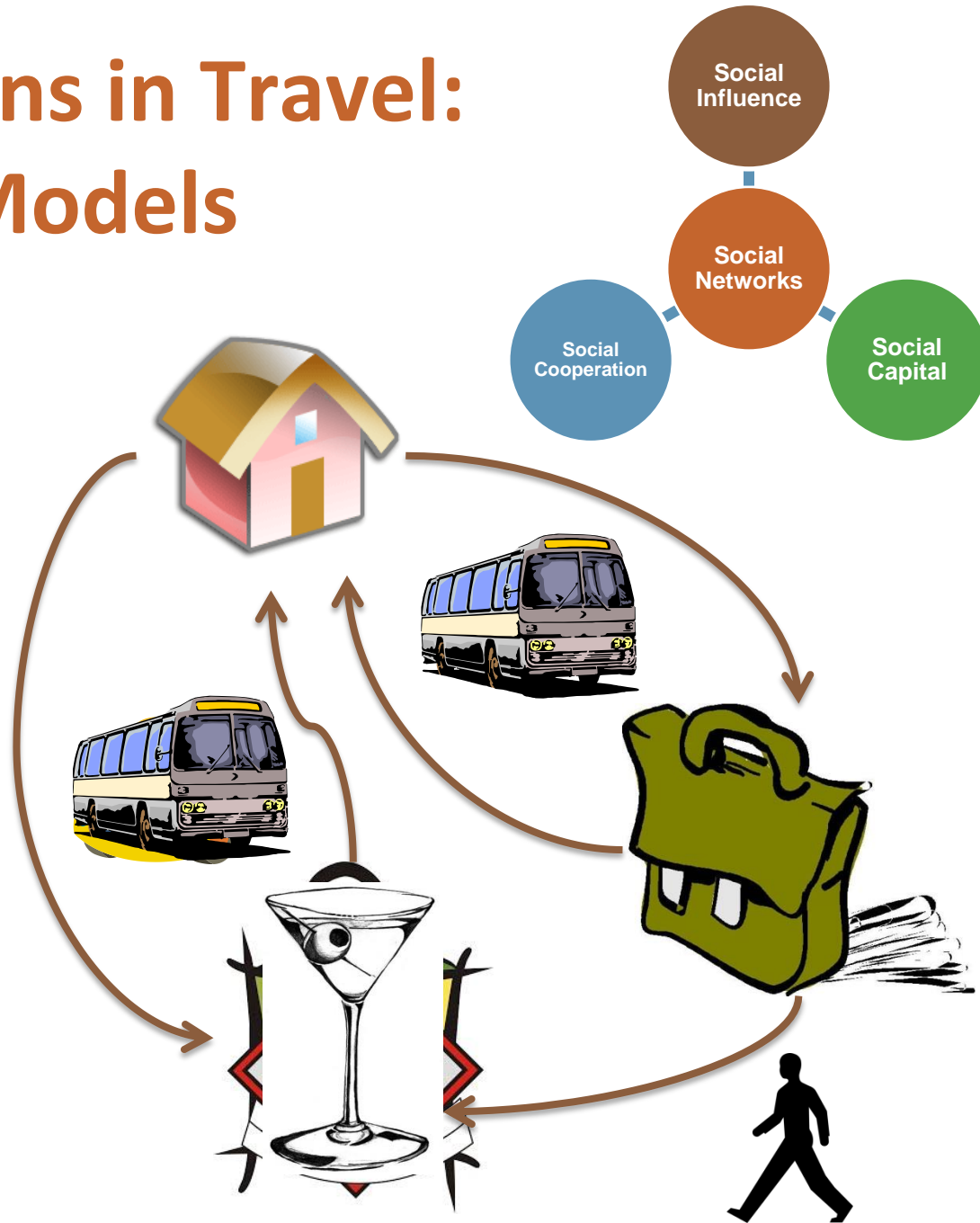
An individual's decision making process is altered by the actions, behaviors, attitudes, and beliefs of others



# Social Interactions in Travel: Activity-Based Models



Example Process of an  
Activity-Based Model  
(Castiglione et al. 2015)



# Measuring social capital



- Various methods exist in the social sciences
  - Global Name Generators
  - Interpersonal Name Generators
  - Position Generators
  - Resource Generators

# Connected Lives Study (Carrasco et al. 2008)



- “people whom you discuss important matters with, or regularly keep in touch with, or are there for you if you need help” (p. 966) [*very close contacts*]
- people who are “more than just casual acquaintances, but not very close” (p. 966) [*somewhat close contacts*]

# Sadri et al. (2015)



- “From time to time, most people discuss important matters with other people. Looking back over the last one month -- who are the people with whom you discussed matters important to you?”

# Kowald and Axhausen (2012)



- “Please list the people with whom you make plans to spend free time (examples: errands, sports, club or organized activities, cultural events, cooking together or going out to eat, taking holidays or excursions together)” and
- “If there are other people with whom you discuss important problems, please list them here” (p. 1089)

# Frei and Axhausen (2007)



- “Please indicate persons with whom you discuss important problems, with whom you stay in regular contact or whom you can ask for help” (p. 194) and
- “Please indicate additional people with whom you undertake leisure activities” (p.195)



# Pike (2014)



- “any five people who have been in your social circle over the past six months” (p.76)
- “the five contacts you have had the most frequent regular interaction with over the past six months” (p. 77)
- “five people in your social circle, with whom you spoke about transportation in the past six months” (p. 77)

# Social Capital in the Modeling of Activity and Travel Behavior



- Measuring social connections & obtaining indicators
  - Name Generator & Interpreter (e.g. Carrasco & Cid-Aguyao 2012, Sadri et al. 2015, Tilahun & Li 2015)
- What about alternative techniques?
  - Position Generator (Lin & Dumin 1986)
  - Useful for measuring *Diversity, Reachability & Resource Access* of an individual's social network
- This study aims to compare the performance of:
  - (1) Name Generator, (2) Name Interpreter, (3) Position Generator

# Data – Personal Networks and Community Study



Time Frame	July and August 2008
Target Population	Noninstitutionalized adults living in the United States, aged 18 and older
Sampling Frame	Households with landline phones and individuals with cellular phones
Sample Design	Random digit dialing of landline and cellular phones
Sample Size	2,512 adults
Response Rate	21% (landline), 22% (cellular phone)
Use of Interviewer	Interviewer administered
Mode of Administration	Phone interview
Computer Assistance	None by respondents
Reporting Unit	One person aged 18 or older per household reports for him/herself and the entire household (landline), one person aged 18 or older reports for him/herself (cellular phone)
Time Dimension	Cross-sectional survey
Frequency	One two-month phase of collecting responses
Levels of Observation	Person, Household
Note: Statistics come from Hampton et al. (2009) and Hampton (2011)	

# Name Generator



- *From time to time, most people discuss important matters with other people. Looking back over the last six months — who are the people with whom you discussed matters that are important to you?*
- *Now let's think about people you know in another way. Looking back over the last six months, who are the people especially significant in your life?*

# Position Generator



- *... I am going to ask about types of jobs and whether people you know hold such jobs. These people include your relatives, friends and acquaintances. Do you happen to know someone who is...*

- |                            |                              |  |
|----------------------------|------------------------------|--|
| 1. a nurse                 | 9. a bookkeeper              | 16. a CEO of large company               |
| 2. a farmer                | 10. a production manager     | 17. a writer                             |
| 3. a lawyer                | 11. an operator in a factory | 18. an admin. assistant of large company |
| 4. a middle school teacher | 12. a computer programmer    | 19. a security guard                     |
| 5. a full-time babysitter  | 13. a taxi driver            | 20. a receptionist                       |
| 6. a janitor               | 14. a professor              | 21. a Congressman                        |
| 7. a personnel manager     | 15. a policeman              | 22. a hotel bell boy                     |
| 8. a hair dresser          |                              |  |

# Social Capital Indicators



- Name Generator
  - Core Network Size\*
  - Tie Dispersion – Discussions, Significant
- Name Interpreter
  - Homophily – Gender, Race, Politics
  - Propinquity
  - Alter Attributes – Family, Social Media Friend
- Position Generator
  - Extensity / Network Diversity (Occupational)

# Activity Participation Model Formulation\*



$$y_{na}^* = \beta_a x_n + \gamma_a z(G_n) + \varepsilon_{na}$$

$$y_{na} = \begin{cases} 1 & \text{if } y_{na}^* \geq 0 \\ 0 & \text{if } y_{na}^* < 0 \end{cases}$$

- $z(G_n) \equiv$  network indicators for the ego network  $G_n$  from individual  $n$
- $x_n \equiv$  individual and household-level characteristics for individual  $n$
- $y_{na} \equiv$  choice indicator of participation in activity type  $a$  for individual  $n$
- $\beta_a, \gamma_a \equiv$  model parameters for activity type  $a$
- $\varepsilon_{na} \sim N(0,1) \equiv$  error term, IID by activity type and individual

\*Hampton et al. (2009) use the same model form but the only indicator they use is core network size. Additionally, their model used data about media use and internet activities, which are unlikely to be measured in an activity-travel survey.



# Analysis Technique

- Model variable decided before analysis begins
- Explanatory Analysis
  - Name generator & interpreter
  - Position generator
- Predictive Analysis – Repeated hold-out cross-validation
  - Naïve Model
  - Nonsocial Model
  - Name generator
  - Name generator & interpreter
  - Position generator
  - Combination model



# Mean Prediction Rate and Count- $R^2$ for Activity Participation Models



Model	Coffee Shop	Place of worship	Library	Fast-Food Restaurant	Other Restaurant	Community Center	Park	Bar
Naïve Model	199.8 [0.00]	215.2 [0.00]	243.8 [0.00]	259.0 [0.00]	273.3 [0.00]	309.4 [0.00]	233.5 [0.00]	273.7 [0.00]
Non-social Model	226.5 [0.15]	236.3 [0.13]	253.4 [0.07]	256.8 [-0.02]	279.6 [0.06]	309.0 [-0.01]	252.9 [0.13]	278.3 [0.04]
Simple Core Network (Name Generator)	229.4 [0.16]	238.9 [0.14]	254.2 [0.07]	255.7 [-0.02]	280.1 [0.06]	309.4 [0.00]	256.2 [0.15]	278.2 [0.04]
Core Network + Attributes (Name Interpreter)	229.1 [0.16]	236.5 [0.13]	253.1 [0.07]	255.4 [-0.03]	279.6 [0.06]	309.2 [0.00]	258.2 [0.17]	276.2 [0.02]
Network Diversity (Position Generator)	234.7 [0.19]	249.4 [0.21]	255.1 [0.08]	259.2 [0.00]	283.4 [0.09]	308.2 [-0.02]	260.4 [0.18]	278.1 [0.04]
Combination: Core Network + Diversity	232.0 [0.18]	248.2 [0.20]	254.7 [0.08]	257.9 [-0.01]	283.2 [0.09]	308.0 [-0.02]	263.3 [0.20]	276.7 [0.03]

Note: In each cell, top number is the mean number of correct predictions on the test set.

In each cell, the number in square brackets is the Adjusted Count- $R^2$  for the correct predictions versus the naïve model.

Shaded cells represent models in which the mean Adjusted Count- $R^2 \geq 0$  at the 99% level.

# Exploratory Model (Name Generator & Interpreter Data)



Activity Types

		Coffee Shop	Place of worship	Library	Fast-Food Restaurant	Other Restaurant	Community Center	Park	Bar
Parameter									
Homophily	Heterophily: Gender					-0.16**			0.18**
	Heterophily: Politics		-0.11*			0.12**			
	Heterophily: Race			0.16*					
	Heterophily: Gender (Small Net)							0.14**	
	Heterophily: Politics (Small Net)								
Alter Attributes	Heterophily: Race (Small Net)			0.13**			0.17**	0.17**	
	Proportion: Family			0.18**					
	Proportion: Social Media	0.25**		0.27**			0.58*	0.40*	
Propinquity	Proportion: In Home					0.63**			
	Proportion: < 5 Miles								
	Proportion: 5-50 Miles								
	Proportion: 50-100 Miles								
Core Network Size	Proportion > 100 miles								
	Number of Alters	0.06**	0.06*	0.07*		0.09*	0.05*	0.07*	
	Proportion: Significant								-0.31*
	Proportion: Important		0.27*		0.21*			0.27*	
<i>Model Statistics:</i>									
	Log-Likelihood	-1209	-1172	-1104	-1120	-981	-854	-1101	-950
	Nonsocial Log-likelihood	-1224	-1188	-1124	-1127	-1003	-871	-1132	-962

Note: Blank cells are parameter that were estimated but not significant at 95% level (\*) or 90% level (\*\*)

# Exploratory Model (Position Generator Data)



## Activity Types

Parameter	Coffee Shop	Place of worship	Library	Fast-Food Restaurant	Other Restaurant	Community Center	Park	Bar
Network Diversity	0.04*	0.07*	0.04*	0.03*	0.04*	0.07*	0.06*	0.03*
Network Diversity > 18	0.47*							
Network Diversity < 3					-0.28*		-0.23**	
<i>Model Statistics:</i>								
Log-likelihood	-1192	-1132	-1101	-1109	-972	-825	-1075	-951
Nonsocial Log-likelihood	-1224	-1188	-1124	-1127	-1003	-871	-1132	-962

Note: Blank cells are parameter that were estimated but not significant at 95% level (\*) or 90% level (\*\*)

# Predictive Model Results



Non-social model generally better than naïve model

Name interpreter indicators typically predicted no better than nonsocial

Models with Extensity indicator predicted better than non-social models

Excluding name interpreter did not improve predictive strength

Models with Extensity indicator generally predicted better overall

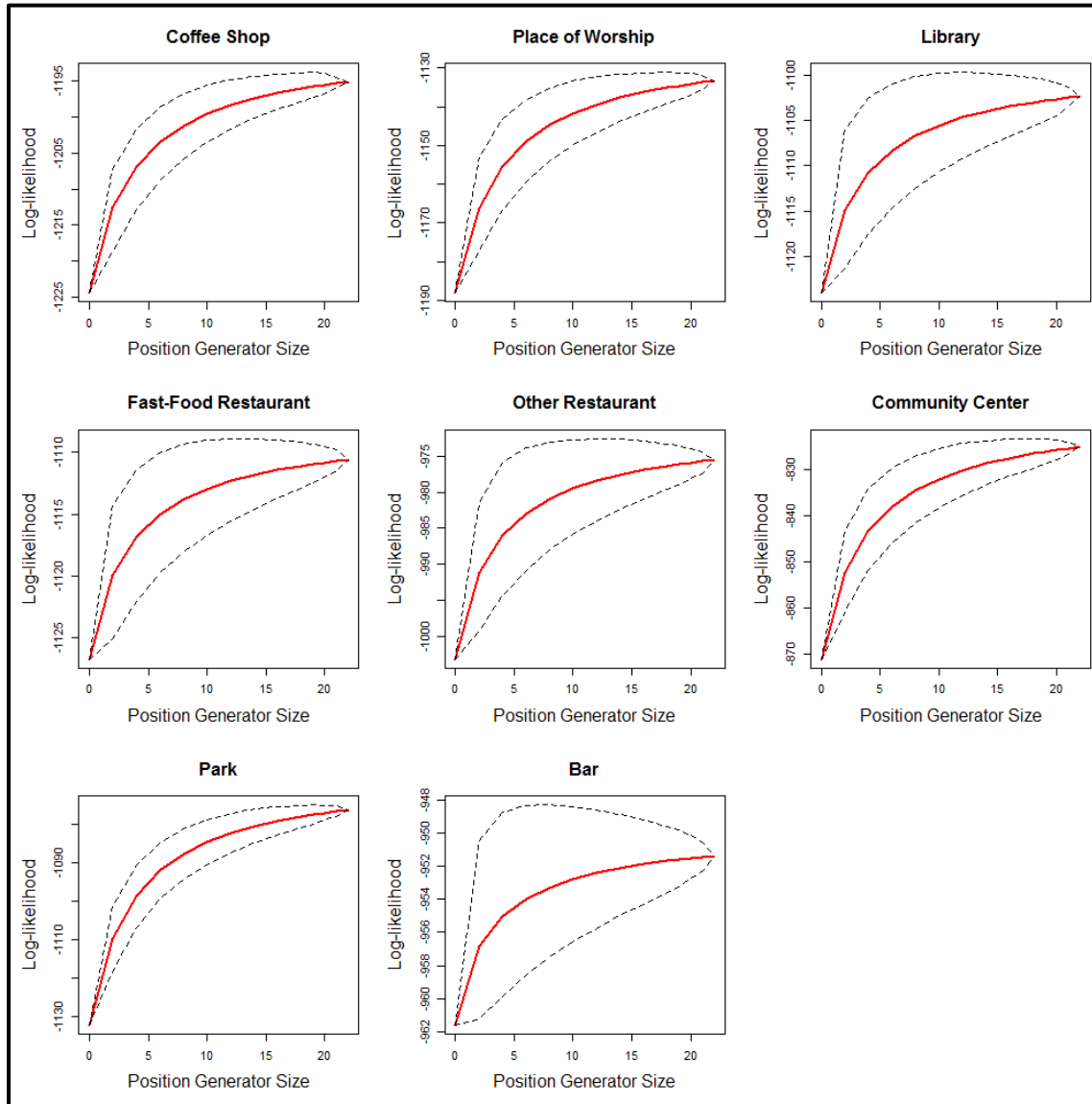
Hypothesis Test Conditions	Coffee Shop	Place of worship	Library	Fast-Food Restaurant	Other Restaurant	Community Center	Park	Bar
$H_0: \mu_{nonsocial} = \mu_{naive}$ $H_1: \mu_{nonsocial} > \mu_{naive}$	Reject (<0.01)	Reject (<0.01)	Reject (<0.01)	Accept (0.999)	Reject (<0.01)	Accept (0.725)	Reject (<0.01)	Reject (<0.01)
$H_0: \mu_{ngni} = \mu_{nonsocial}$ $H_1: \mu_{ngni} > \mu_{nonsocial}$	Reject (<0.01)	Accept (0.417)	Accept (0.636)	Accept (0.976)	Accept (0.508)	Accept (0.349)	Reject (<0.01)	Accept (0.998)
$H_0: \mu_{pg} = \mu_{nonsocial}$ $H_1: \mu_{pg} > \mu_{nonsocial}$	Reject (<0.01)	Reject (<0.01)	Reject (<0.01)	Reject (<0.01)	Reject (<0.01)	Accept (0.881)	Reject (<0.01)	Accept (0.631)
$H_0: \mu_{combo} = \mu_{nonsocial}$ $H_1: \mu_{combo} > \mu_{nonsocial}$	Reject (<0.01)	Reject (<0.01)	Reject (0.037)	Accept (0.064)	Reject (<0.01)	Accept (0.940)	Reject (<0.01)	Accept (0.986)
$H_0: \mu_{ng} = \mu_{ngni}$ $H_1: \mu_{ng} > \mu_{ngni}$	Accept (0.354)	Reject (<0.01)	Accept (0.065)	Accept (0.326)	Accept (0.258)	Accept (0.604)	Accept (0.997)	Reject (0.003)
$H_0: \mu_{pg} = \mu_{ngni}$ $H_1: \mu_{pg} > \mu_{ngni}$	Reject (<0.01)	Reject (<0.01)	Reject (<0.01)	Reject (<0.01)	Reject (<0.01)	Accept (0.942)	Reject (<0.01)	Reject (<0.01)
$H_0: \mu_{pg} = \mu_{ng}$ $H_1: \mu_{pg} > \mu_{ng}$	Reject (<0.01)	Reject (<0.01)	Accept (0.109)	Reject (<0.01)	Reject (<0.01)	Accept (0.905)	Reject (<0.01)	Accept (0.593)
$H_0: \mu_{combo} = \mu_{ngni}$ $H_1: \mu_{combo} > \mu_{ngni}$	Reject (<0.01)	Reject (<0.01)	Reject (0.017)	Reject (<0.01)	Reject (<0.01)	Accept (0.974)	Reject (<0.01)	Accept (0.240)
$H_0: \mu_{combo} = \mu_{pg}$ $H_1: \mu_{combo} > \mu_{pg}$	Accept (1.000)	Accept (0.952)	Accept (0.551)	Accept (0.965)	Accept (0.624)	Accept (0.655)	Reject (<0.01)	Accept (0.969)

# Robustness of Position Generator



- Extensity was measured through an indicator
  - *How much is the model dependent on question design?*
- Prior Research
  - Verhaeghe et al. (2013) – multiple occupational lists
  - Hällsten et al. (2015) – impact of removing each occupation & sets of 10 occupations
- Sensitivity Analysis of Occupational List Length
  - *Model Fit*: Log-likelihood
  - *Parameter Estimate*:  $\text{MSE}(\gamma_{extensity})$
  - *Parameter Ratios*:  $\text{MSE}(\gamma^a / \gamma_{extensity})$

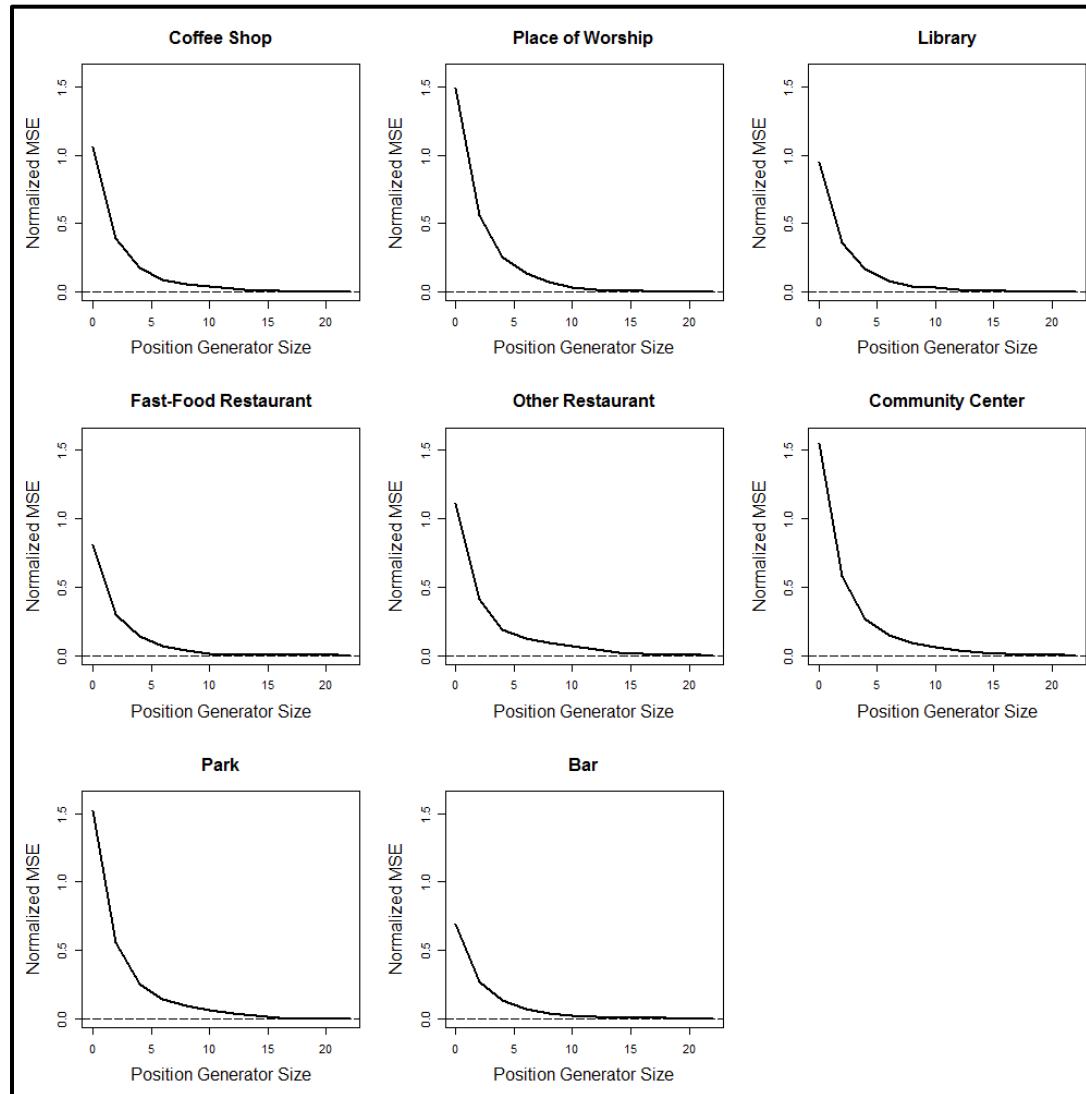
# Results: Model Fit



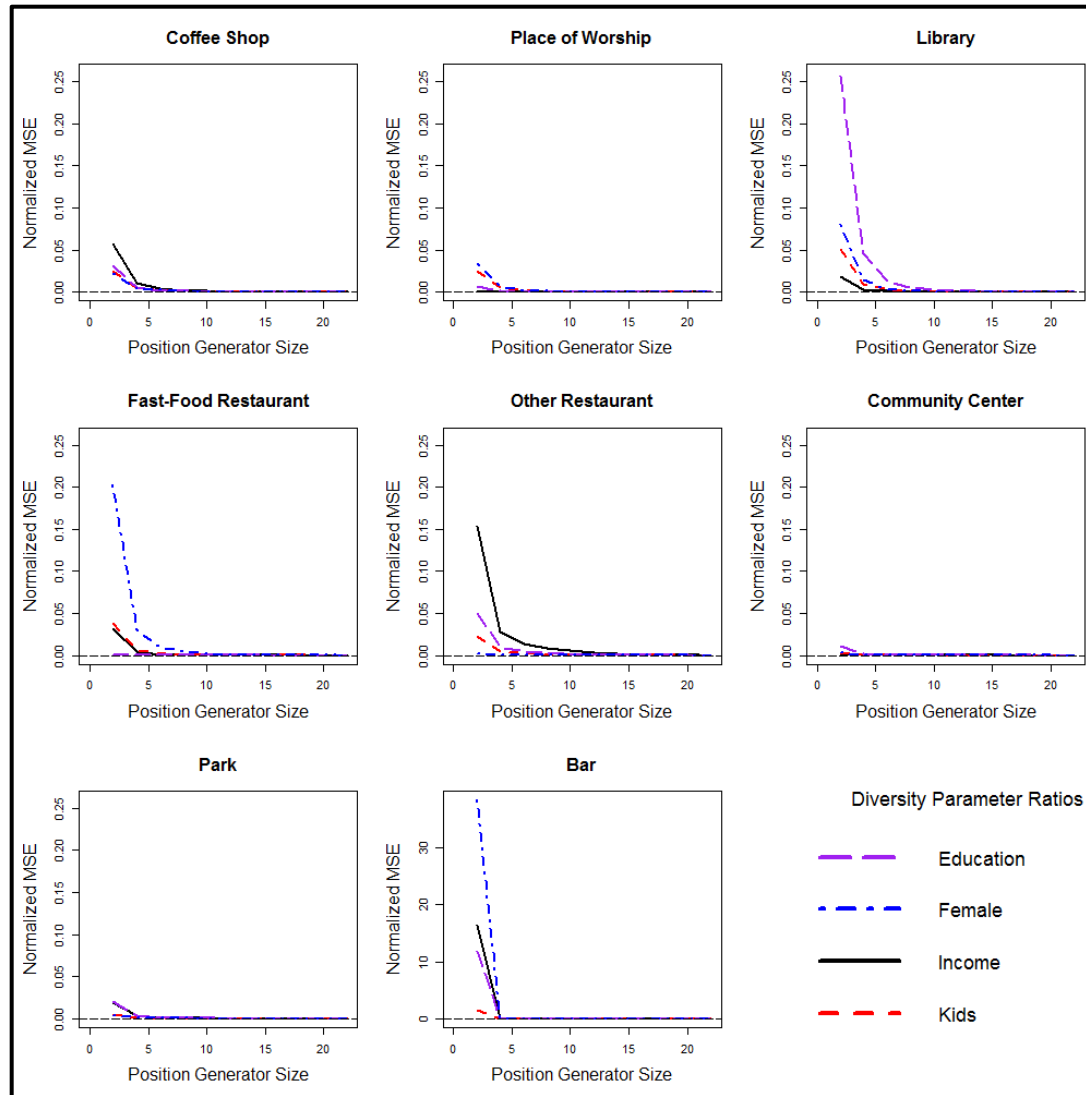
## Legend

- 95<sup>th</sup> Percentile
- Median
- 5<sup>th</sup> Percentile

# Results: Parameter Estimates



# Results: Parameter Ratios





# Summary



- Supports prior research on:
  - Correlation of social capital and activity participation (Tilahun and Li 2015, Sadri et al. 2015)
  - Robustness of extensity indicator (Verhaeghe et al. 2013), Hällsten et al. 2015)
- Extensity / network diversity valid for explaining & predicting activity selection
  - Extend to occupational prestige
- Position generator provides new tool for activity diary & travel survey: content-free, structure-centric
  - Does it need to be tailored to transport & activity?
- Direction of extensity & activity participation unclear

# Future Research Directions



- Social Capital and Activity-Travel
  - Activity Diary Design
  - Combining position and name generator data
  - New indicators, incorporate resource generators