THE EFFECTS OF SOCIO-ECONOMIC & TRANSPORTATION ACCESSIBILITY ON AREA-LEVEL DIABETES COUNTS: A LATENT-VARIABLE STRUCTURAL EQUATIONS MODEL APPROACH

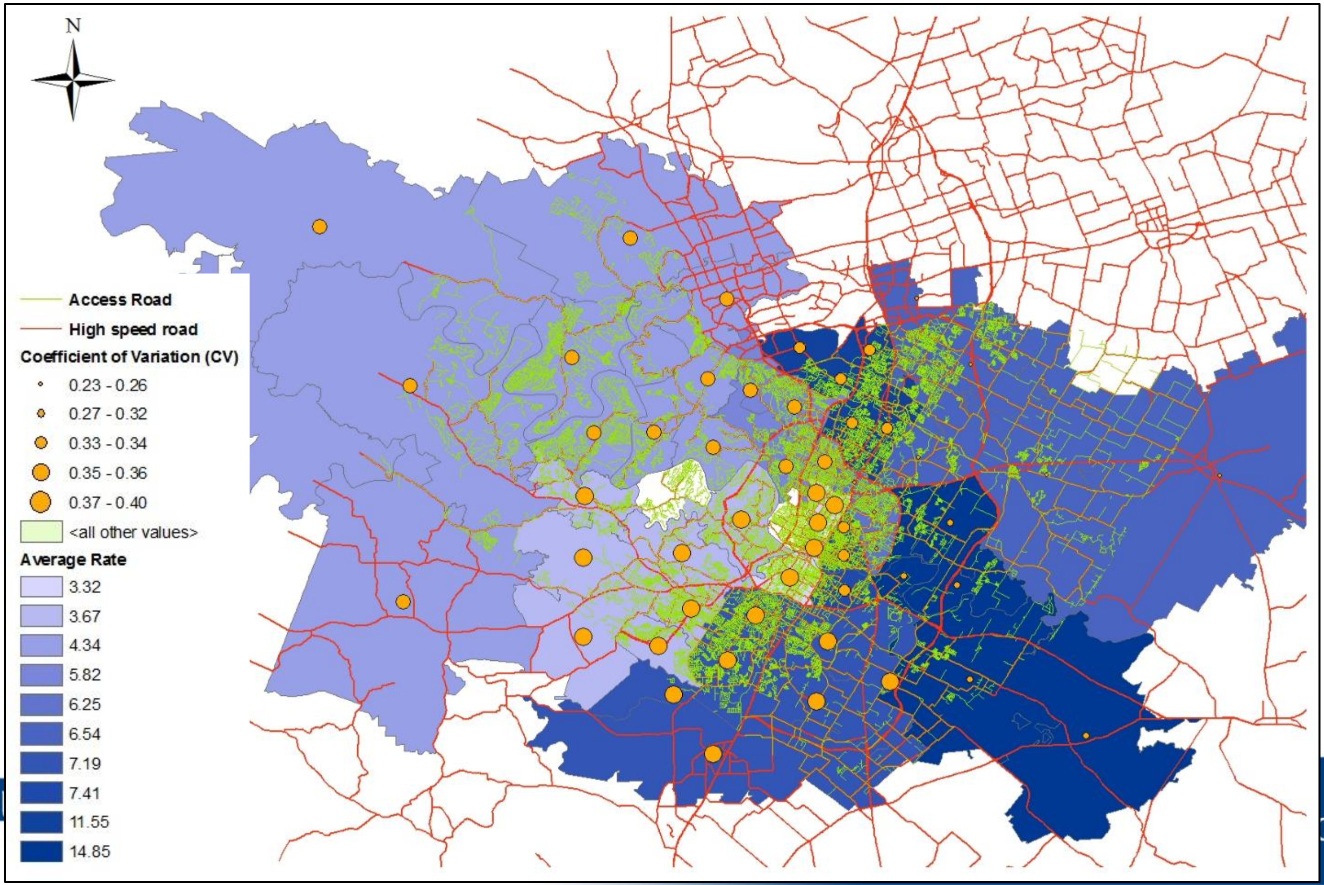
Yiyi Wang, Ph.D.
Civil Engineering Department
Montana State University

Motivation

- A sedentary lifestyle was estimated to cost **300,000** lives & **$51.6 billion** of medical treatment per year in the U.S.
- **60 to 80 percent** of the world's population does not meet the **physical activity** levels recommended by WHO, and even people in industrialized countries lead **inactive** lifestyles.
- The goal of this work is to develop a **generalized** model to test if & how transportation accessibility affects **diabetes** rates, while controlling for **self-selection bias** and other confounding effects.

Data Sets (1)

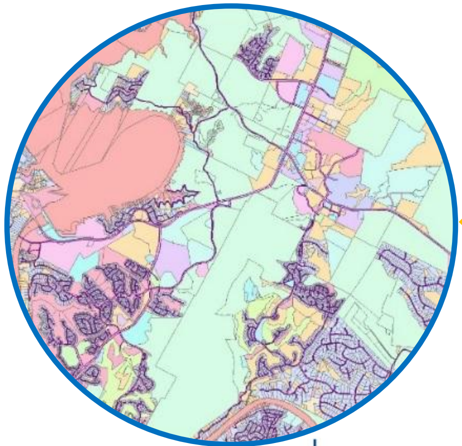
- Age-adjusted diabetes prevalence data were obtained from the Behavioral Risk Factor Surveillance System Survey (BRFSS) for Travis County, Texas.



Data Sets (2)

ACS + TCAD

- Population density (persons per sq. mile)
- Median household income
- Land market value*

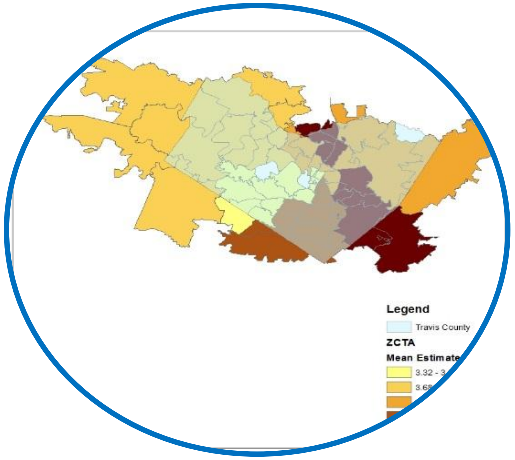


- Land use entropy
- Multi- (or Single-) family w/in ½ mile of bike lanes (or parks)%
- Sidewalk density; Job/POP..

Data Sets (3)

Vehicle miles traveled (VMT) density;

Local road density



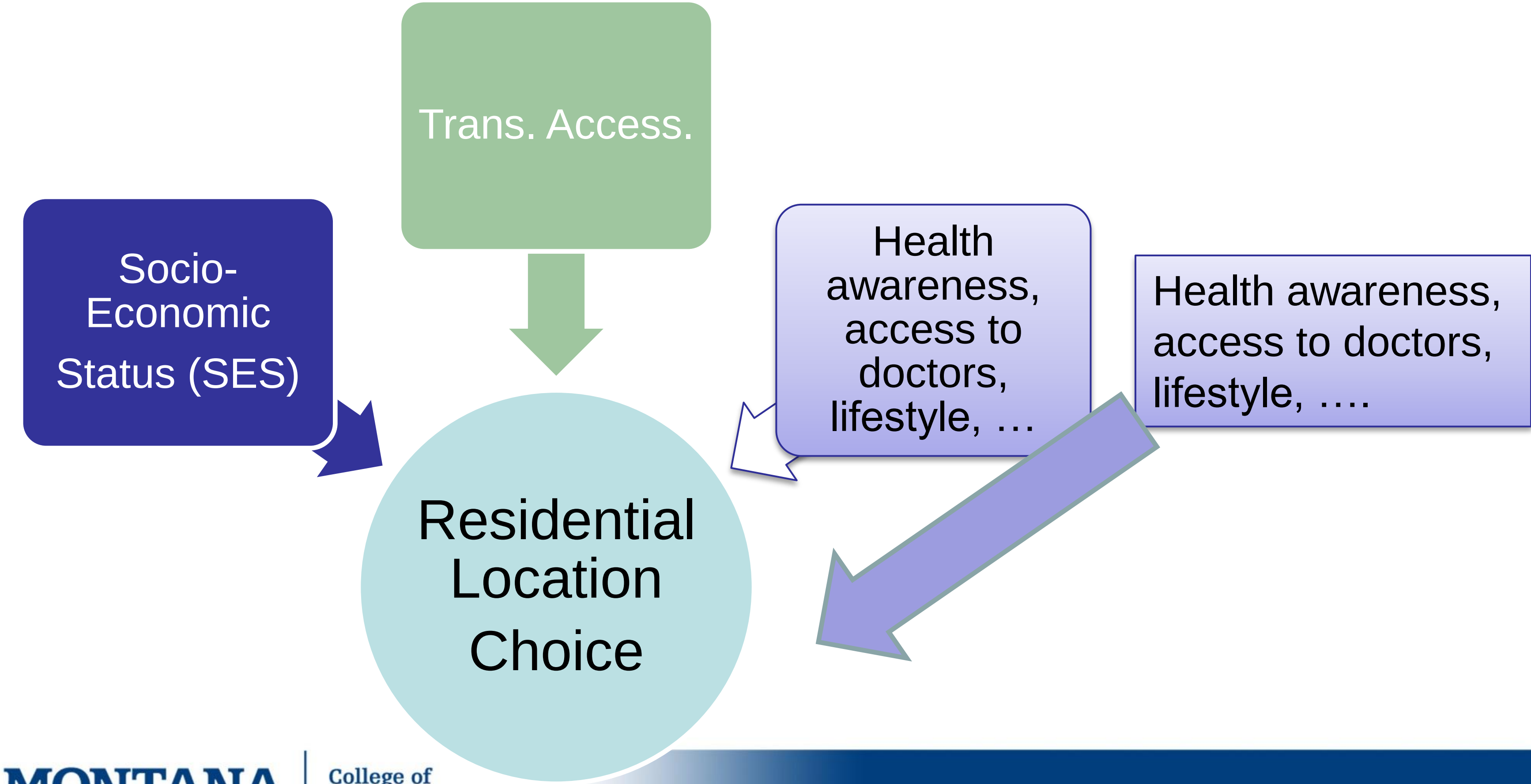
Age-adjusted diabetes counts

TABLE 1 Summary Statistics (No. of Obs. = 606)

	Mean	Std. Dev.	Min	Max
Population Density (persons per sq. mile)	3,200	3,990	0.849	30,718
Land-Use Entropy	0.490	0.250	0.00	1.00
Job-Population Ratio	30.79	198.2	0.00	3132.00
Sidewalk Density	18.38	15.93	0.00	58.54
VMT Density	1.26E+05	1.59E+05	0.00	1.01E+06
Load Road Density	5.99E+04	4.29E+04	0.00	2.06E+05
Multi-Family w/in ½ mile Bike%	0.48	0.44	0.00	1.00
Single-Family w/in ½ mile Bike%	0.44	0.48	0.00	1.00
Multi-Family w/in ½ mile Parks%	0.02	0.04	0.00	0.36
Single-Family w/in ½ mile Parks%	0.18	0.31	0.00	2.84
Land Market Value ($1,000 per sq. mile)	2,297	4,885	0.00	5.14E+04
Median Household Income ($1,000)	70.23	42.80	0.00	233.13
Age-adjusted Diabetes Counts (Response Variable)	119.95	182.51	0.00	1291.00

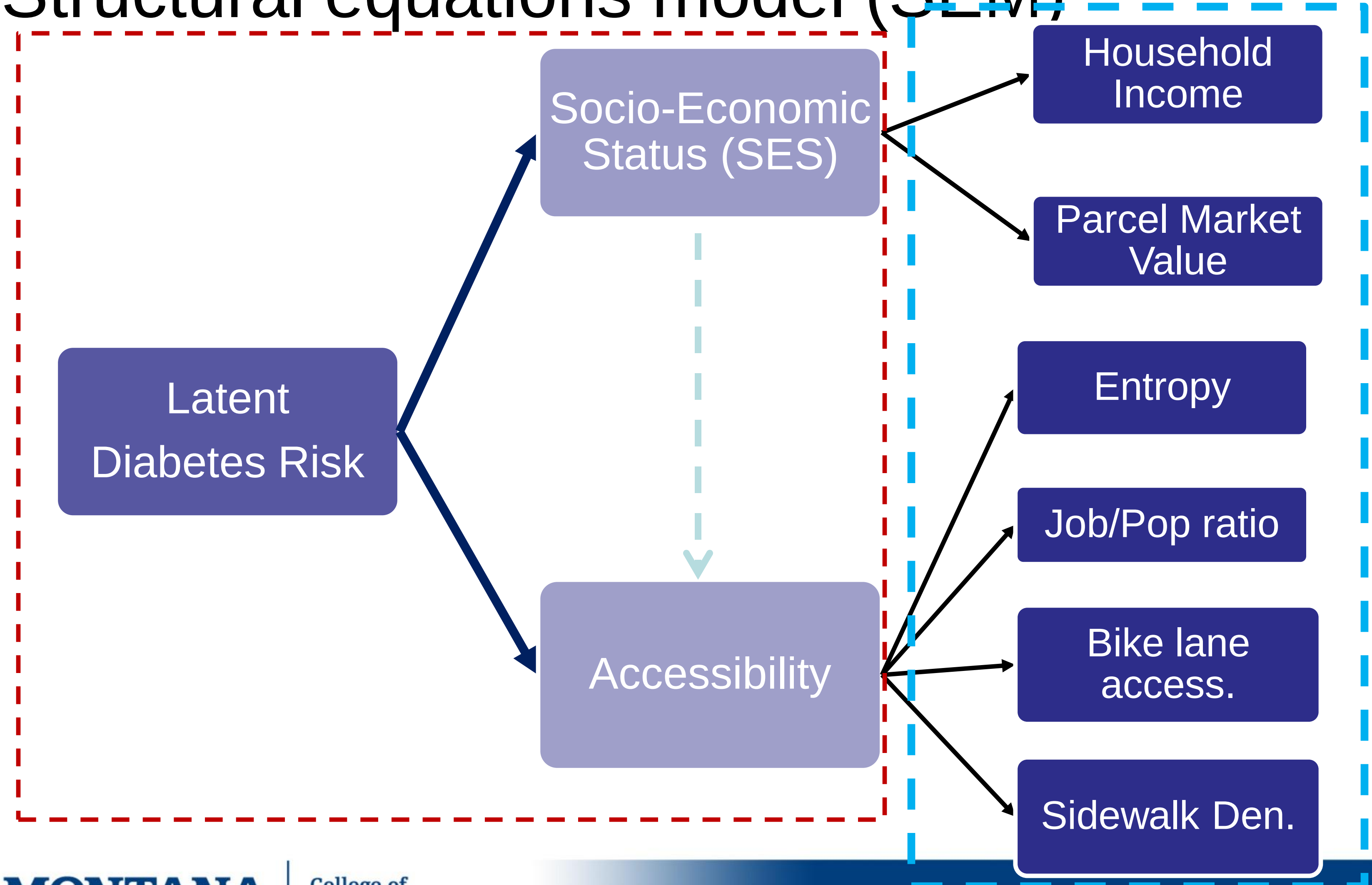
Methodology

- Residential self-selection



Methodology(2)

- Structural equations model (SEM)



Methodology (3)

❑ Phase I

$$D_i \sim \text{Poisson}\ (Z_i)$$

$$Z_i = POP_i \cdot \exp(f_{3,i} + v_i)$$

- *D* = the observed diabetes count at each TAZ
- Z = diabetes rates
- exp( ) = the exponential function
- f_3 = the underlying (latent) diabetes risk factor
- *v* = a random term unique to each TAZ, assumed to follow a normal distribution with area-specific variance, $N(0, \sigma_i^2)$.

Methodology (4)

❑ Phase II

$$\begin{cases} f_3 = \beta_0 + \beta_1 f_1 + \beta_2 f_2 + \varepsilon \\ f_1 = \beta_4 f_2 + \delta \end{cases}$$

- β_1, β_2 = measure the signs and magnitude of the effects of accessibility and SES on the logarithm of diabetes risk.
- β_4 explains, to some extent, the degree of residential self-selection.

❑ Phase III – Measurement models

Results

- Structural model

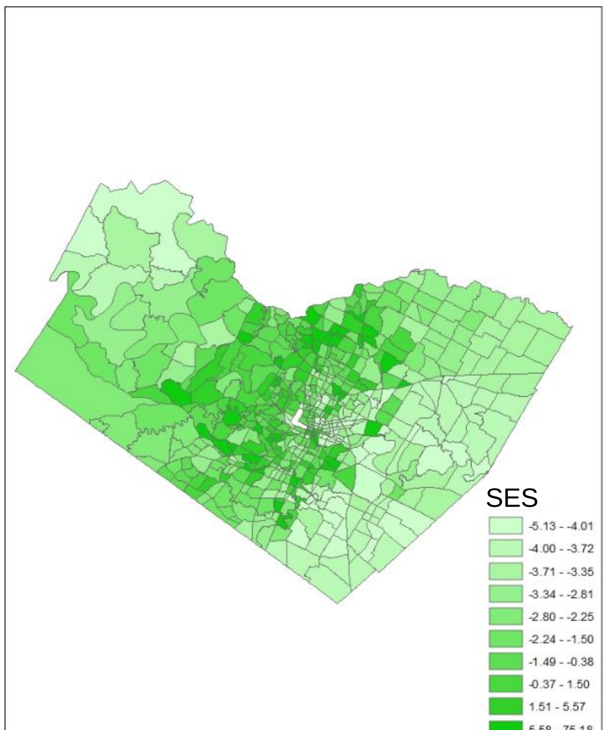
Variable Name	Node	Mean (t stats)	Marginal Effect*	Variance Terms	
Constant	β_0	-3.124 (-35.08)	—	Node	Mean (t stats)
Accessibility	β_1	-1.233 (-7.190)	0.291	τ_{e2}	4.091 (3.252)
SES	β_2	-0.838 (-7.490)	0.433	τ_{e1}	0.332 (2.162)
Self-selection	β_3	0.642 (6.528)	—	τ_{e3}	34.34 (0.725)

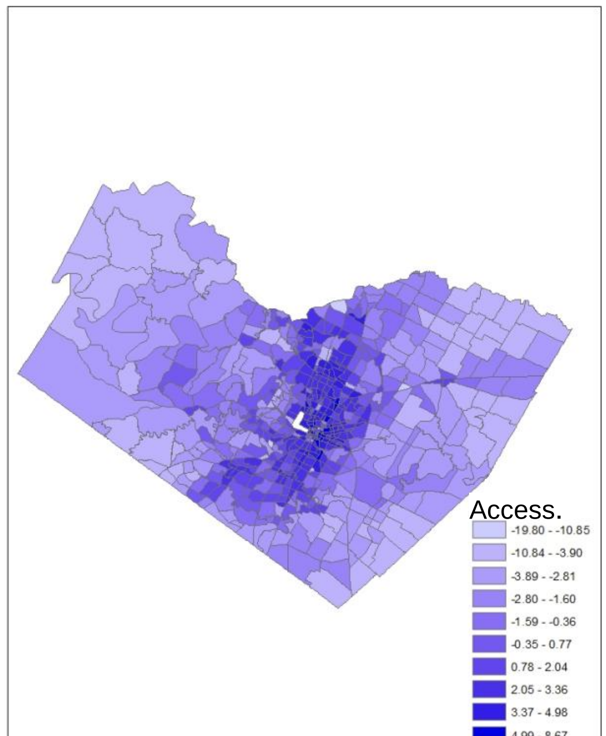
- Measurement models

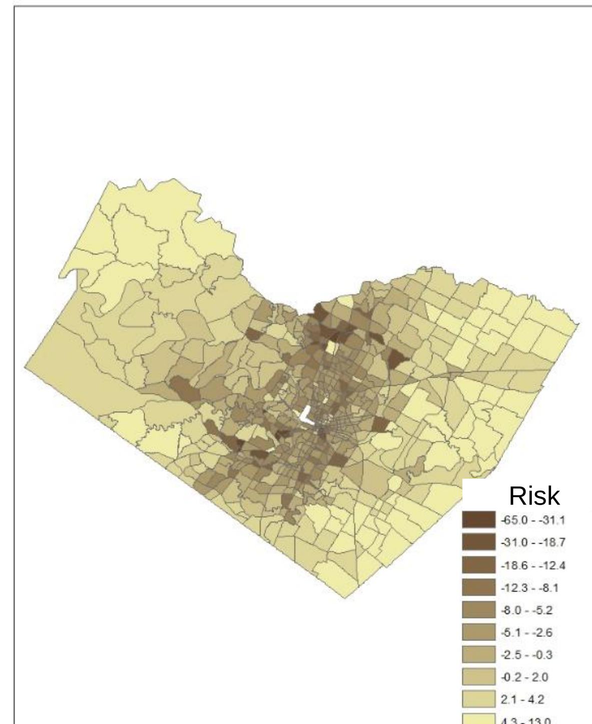
	Variable Name	Node	Mean (t stats)	Elasticity	Variance Terms	
Accessibility	**Land-use entropy**	λ_1	0.608 (3.565)	-0.280	τ_1	6.403 (16.36)
	job pop ratio	λ_2	1.071 (1.890)	0.012	τ_2	0.005 (17.24)
	VMT Den	λ_3	1.104 (0.348)	0.034	τ_3	5.94E-12 (0.059)
	AccessRd Den	λ_4	0.948 (0.297)	-0.045	τ_4	3.56E-11 (0.356)
	BikeSingFam %	λ_5	0.390 (3.548)	-1.03E-15	τ_5	4.990 (16.80)
	BikeMulFam%	λ_6	1.194 (3.572)	-0.049	τ_6	10.3 (1.263)
	Pop. Density	λ_7	1.000 (--)	—	τ_7	3.44E-07 (17.03)
SES	**Market Value**	λ_8	7.424 (3.729)	-0.366	τ_8	0.001 (15.93)
	Med. HH. Income	λ_9	1.000 (--)	—	τ_9	4.20E-06 (17.16)
	"—" not applicable				τ_v	9.807 (1.293)

Results (2)

- Latent Factors for SES, Accessibility, and Diabetes Risk







Conclusions

- A latent-variable approach for discrete responses is developed to gauge the effects of SES and neighborhood accessibility on diabetes rates.
- Higher accessibility correlates with lower diabetes risk in statistically significant ways (t stats = -7.19), so does higher socio-economic status (t stats = -7.49).
- The structural model also identifies significant self-selection effect. Within the training data, people with a higher socio-economic status (SES) tend to live in more accessible neighborhoods.

Conclusions (cont'd)

- Visual displays of latent scores can be used as a proactive way to identify sub-optimal areas.
- Wealth is correlated with lower diabetes risk. A one percent increase in average market land value is associated with a 0.366% reduction in diabetes risk, holding everything else constant.
- Land-use balance and proximity to bike lanes contribute to greater accessibility in practically and statistically significant ways. Providing biking facilities near multi-family developments (i.e., high density) correlates with greater health benefits than providing these opportunities near single-family developments (i.e., low density).

Limitations

- Attitude, values, preference …
- Nutrition, diet …
- Data resolutions of diabetes estimates

Thank you!

Questions & suggestions?