A GIS-based Tool for Forecasting the Travel Demands of Demographic Groups within California – An Optimal Resource Allocation Tool

October 30, 2007

Kostas Goulias

1st PATH-UCTC Conference

Research by UCSB & UCI

Outline

- Project Objectives
- Parallel tracks
  - Statewide analysis
    - Efficiency
    - Equity
  - Individual analysis

- Data and methods used
- Inequality vs Optimality
- Next steps

Key Ideas (the zonal approach)

Consider a geographic area:

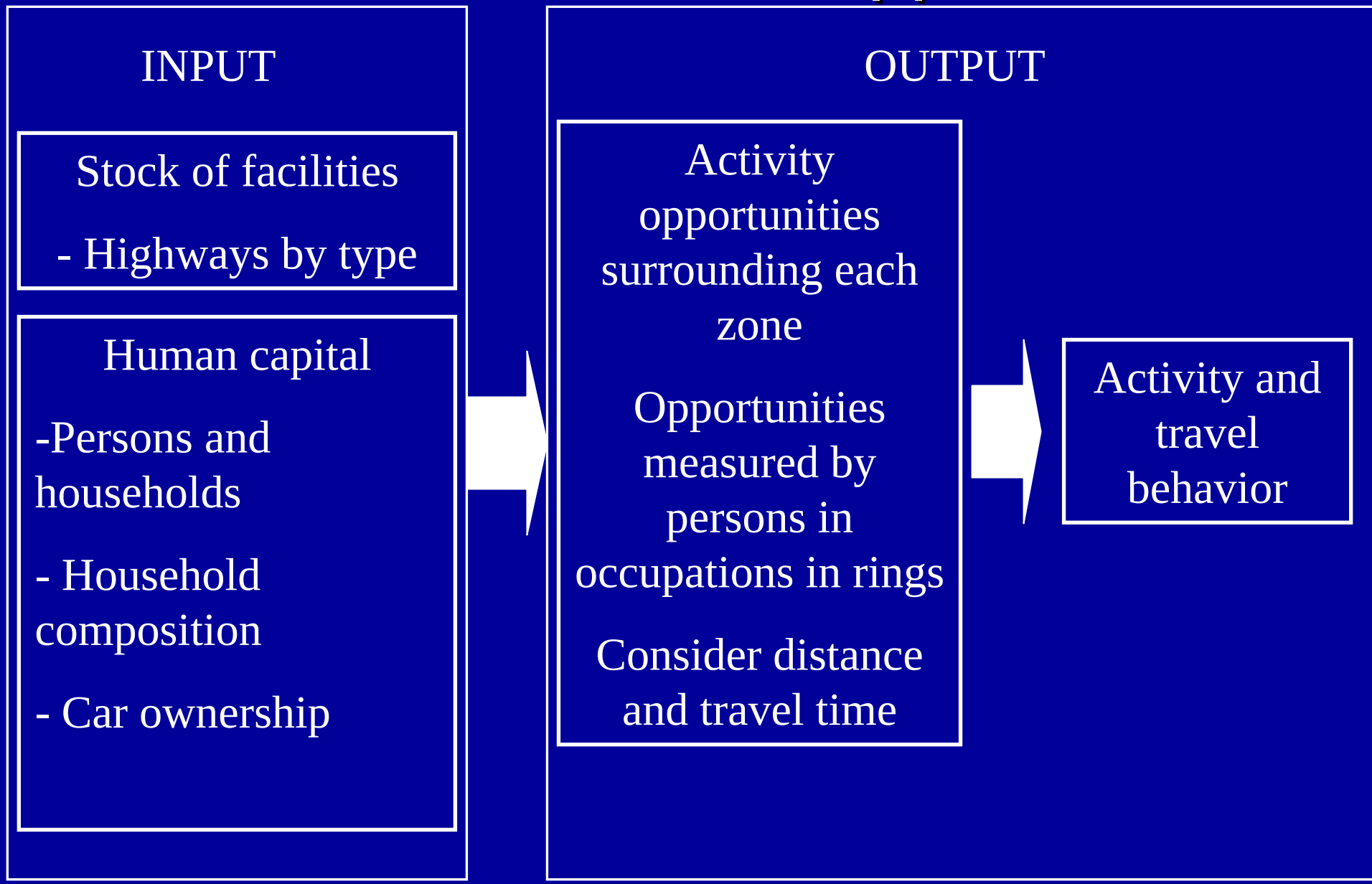
- Identify and assess the level of service offered by the transportation system (accessibility and mobility)
- Build an inventory of available facilities in GIS
- Identify the stream of investments and assess its impact on the economy and social welfare for the specific geographic area
- Compare all geographic areas in terms of the ratio between investments and level of service offered

Key Ideas (the person approach)

Consider a person in a household:

- Identify and assess the level of service offered by the transportation system (accessibility and mobility) for the dwelling unit of the household
- Build an e-inventory of available facilities in GIS
- Identify the stream of investments and assess its impact on the behavior and social benefit for the specific dwelling unit location
- Build geographic summaries of households at a suitable scale (e.g., TAZ, TRACT, City)
- Compare all geographic areas in terms of the ratio between investments and level of service experienced and social benefit

Schematic for the zonal approach



Potential application areas

- Economic development (Gross regional product)
- Distributional justice
- Social welfare
  - Access to opportunities
  - Access to jobs
- Activity and travel behavior
  - Time allocation (activity & travel)
  - Vehicle kilometers of travel
  - Energy consumption

Our study

- Social “welfare” efficiency of investment
  - (Zonal approach) Study the ratio of outputs over inputs and find geographic areas that maximize this ratio using suitable methods
  - (Person approach) Develop a “user benefit” measure (e.g., amount of time in activities over travel time) and study this ratio as a function of inputs using Structural Equations Model (SEM)
  - Create a GIS tool to build maps of benefits vs. investments (for zonal and person approaches)

Focus on Zonal Approach Today

Pragmatic considerations (1)

- Recording of investment streams and their impact on the transportation system does not exist (yet)
- We have their footprint as a detailed network of links and nodes
- We also have indicators of services available (number of persons by industry at Census blocks/block groups)
- We also have resident characteristics at similar geographies

Pragmatic considerations (2)

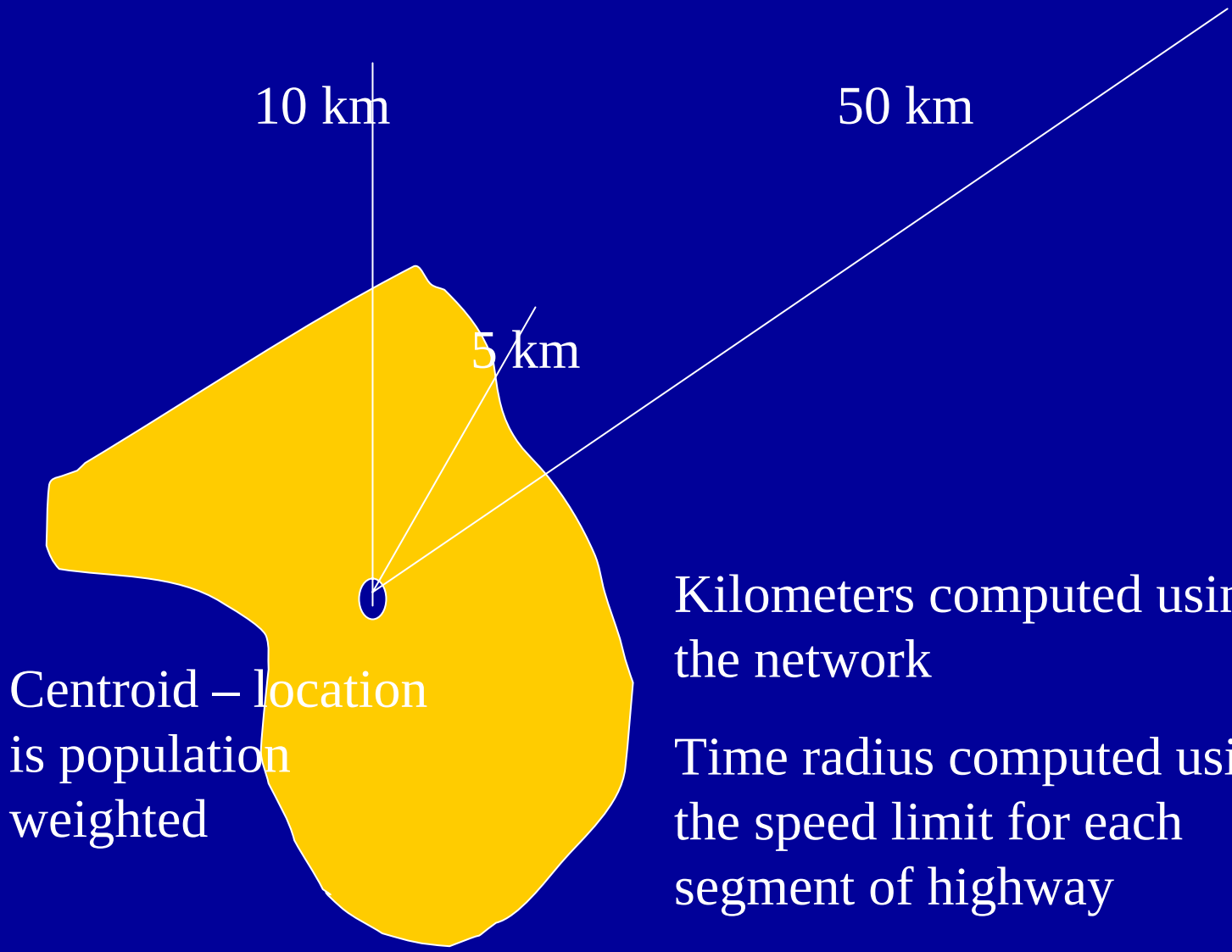
- We cannot do all this using the ideal detail (land parcels and individuals)
- Proceed in 2 parallel tracks
  - First order approximation for the entire state of California
  - Develop a small pilot study using more detailed data for Santa Barbara county
- Compare the two and decide what to do next
- Produce a tool that can be used immediately after the project completion

Steps

- Used CTPP data and developed accessibility measures – number of persons by industry reachable within a given travel time
- Analyzed the data and developed geographic classes
- Dropped CTPP as a network source and replaced it with a TeleAtlas (GDT) network
- Used CENSUS occupation data in addition to CTPP

Assemble stock and service inventory

- INPUT = TeleAtlas Highways+
  - Total lane kilometers of highways within 5km, 10km, 50km, 20 minutes, 40 minutes
  - Primary limited access
  - Primary not limited access
  - Secondary and connecting
  - Local, neighborhood, and rural
  - Special character
  - All other



Centroid – location is population weighted

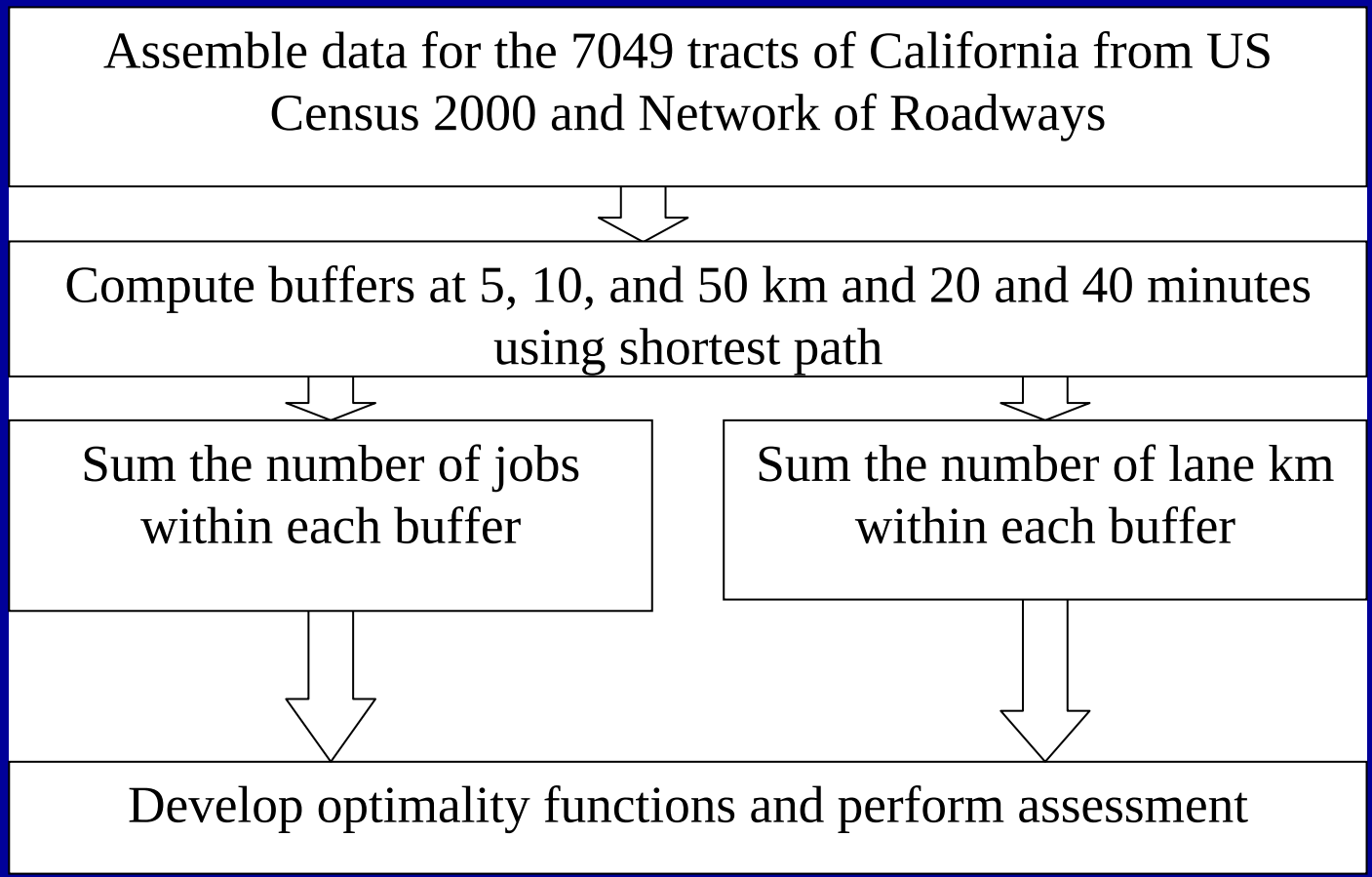
Kilometers computed using the network

Time radius computed using the speed limit for each segment of highway

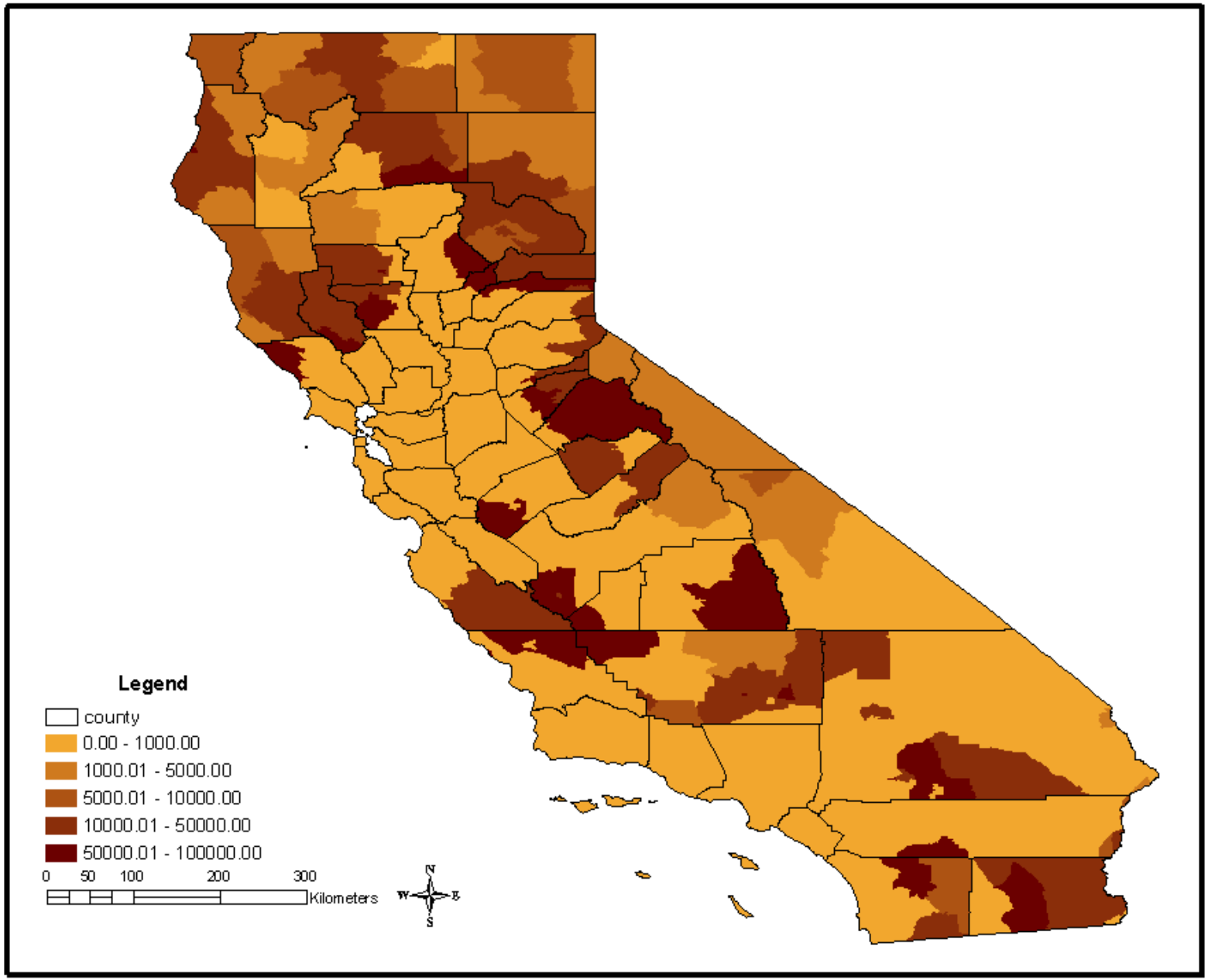
Assemble data for the output (benefit)

- OUTPUT = Number of activity opportunities by distance and travel time that can be reached by each origin (25 variables)
  - Total number of persons that can be reached within 5km, 10km, 50km, 20 minutes, 40 minutes
  - Industries and occupations
  - Occupations
    - Retail
    - Health
    - Services
    - Manufacturing
    - All other

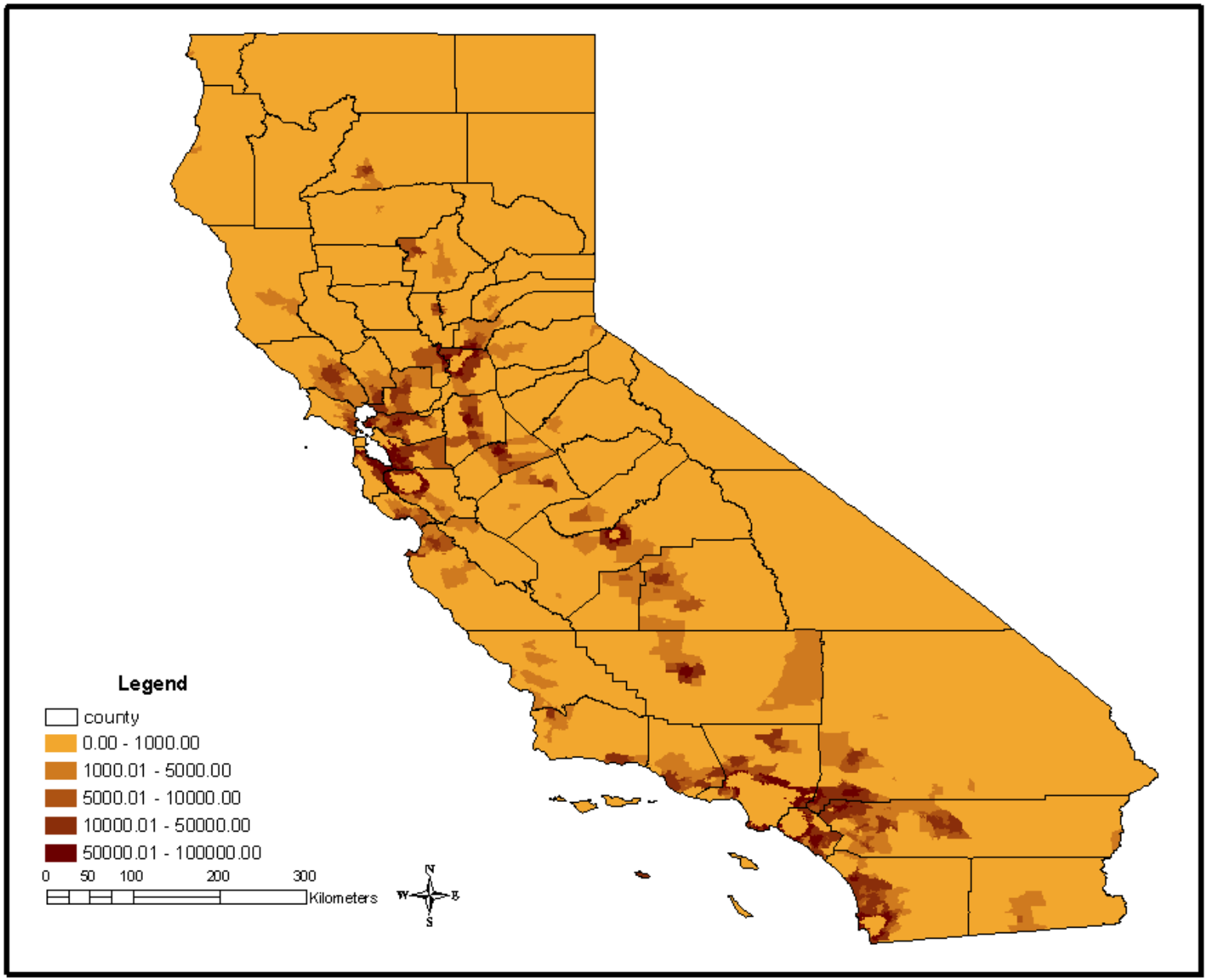
Computation Schema of the Study



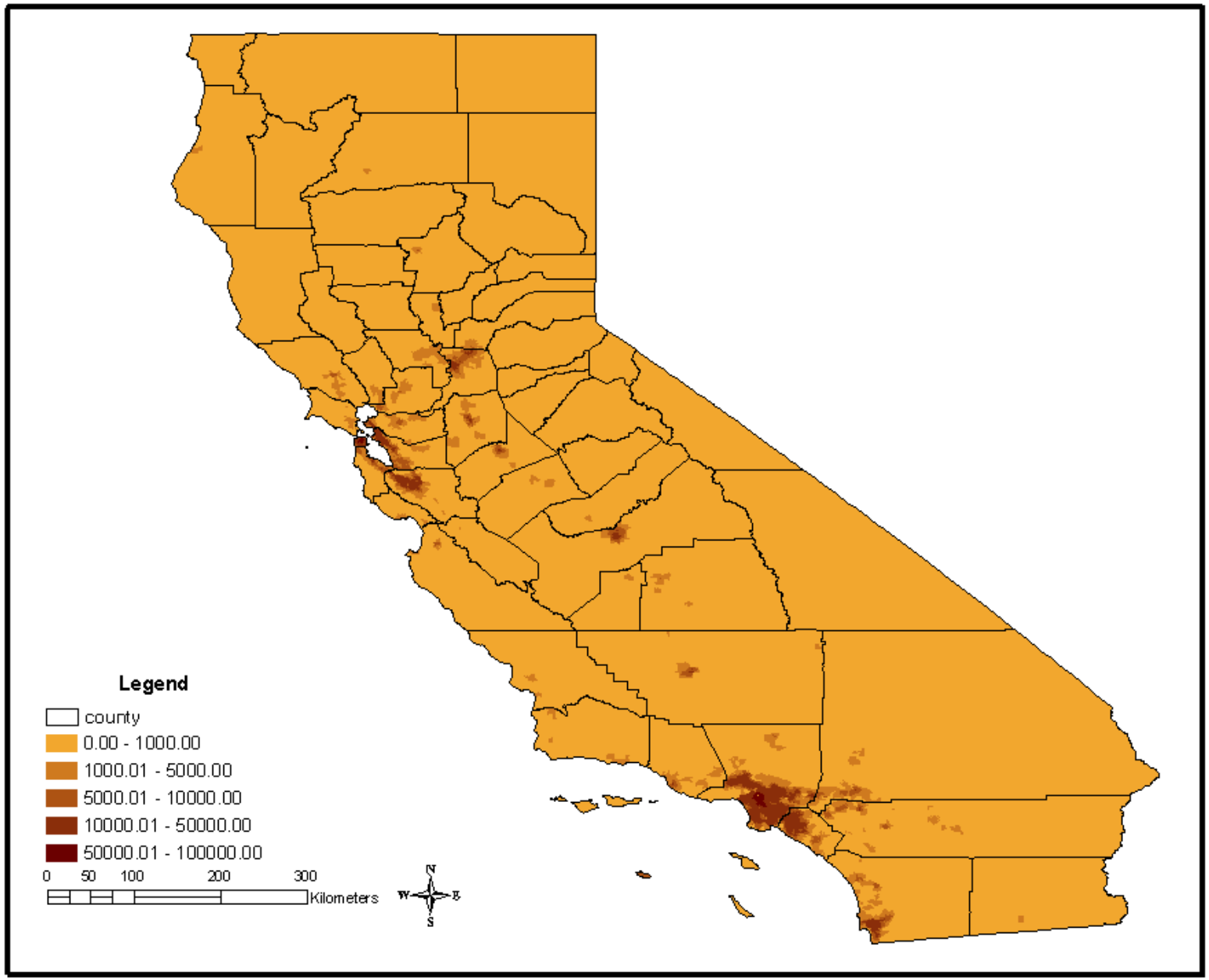
CA: TOTAL LANE KILOMETERS within 50km



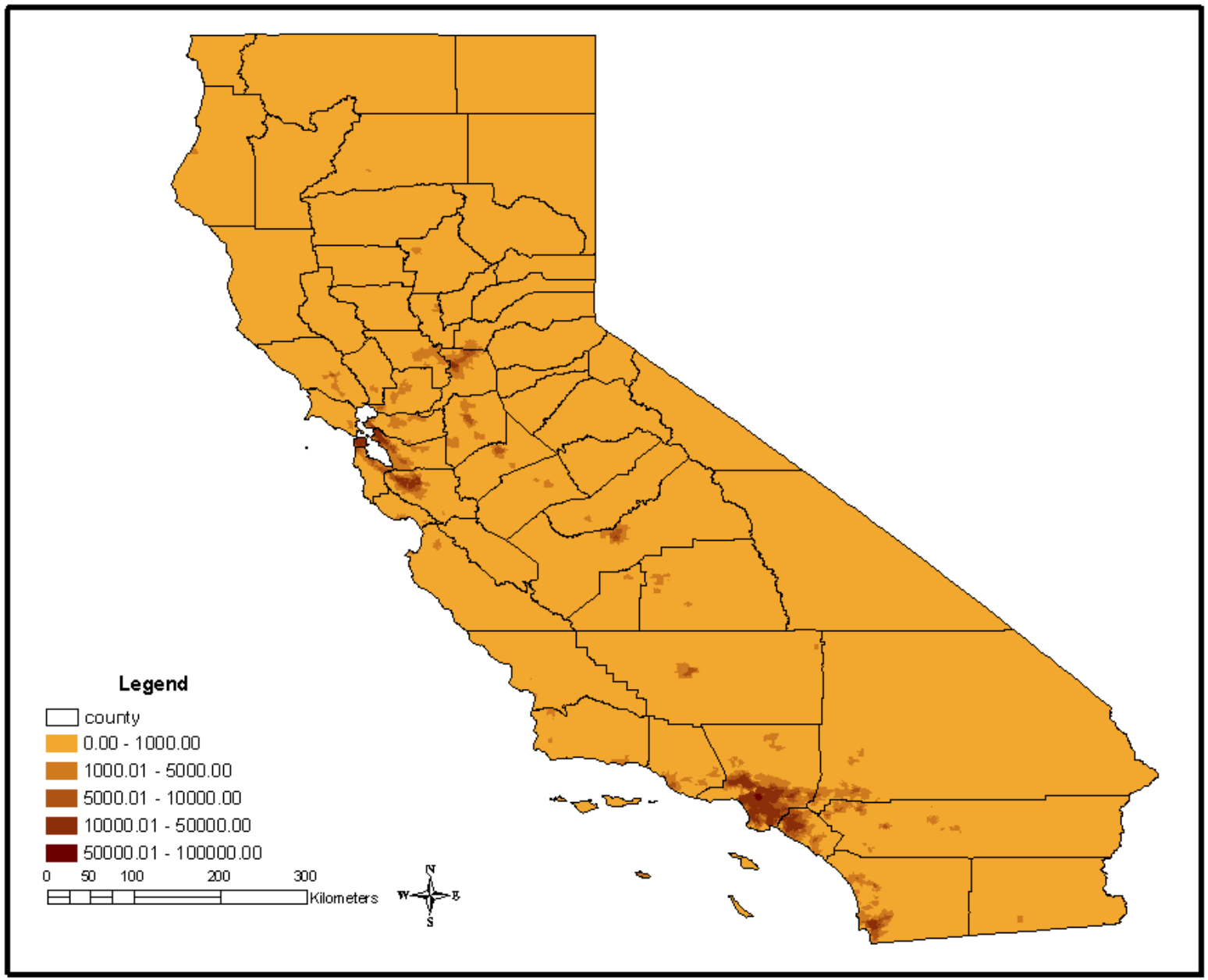
CA: TOTAL LANE KILOMETERS within 10km



CA: TOTAL LANE KILOMETERS within 5km

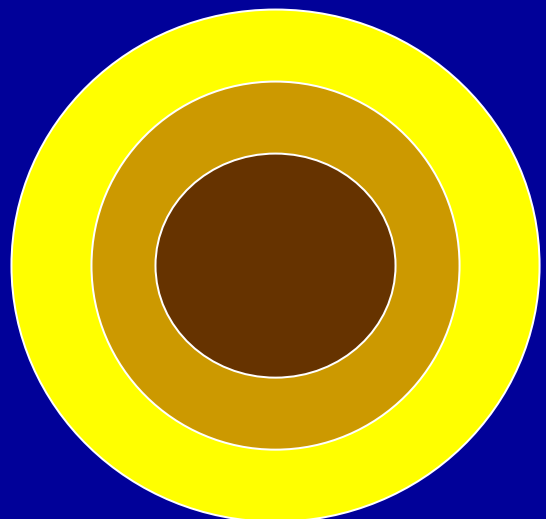


CA: LOCAL, NEIGHBORHOOD, and RURAL ROAD within 5km



Combine different measures

- Use Factor Analysis
  - Three factors derived from 25 variables
  - Represent three concentric rings
  - Approximately represent around each centroid:
    - Core
    - Middle ring
    - Outer ring



Summary Measures of Access

Three principal components extracted from 25 output variables and their scores

	Components		
	Outer Ring Access	**Middle Ring Access**	**Core Access**
Number of Workers in Retail (20 to 40 min)	**0.945**	0.276	0.139
Number of Workers in Services (20 to 40 min)	**0.941**	0.250	0.128
Number of Workers in Other (20 to 40 min)	**0.941**	0.275	0.150
Number of Workers in Manufacturing (20 to 40 min)	**0.939**	0.245	0.130
Number of Workers in Health (20 to 40 min)	**0.936**	0.287	0.140
Number of Workers in Retail (10 to 50 km)	**0.927**	0.330	0.159
Number of Workers in Manufacturing (10 to 50 km)	**0.926**	0.311	0.129
Number of Workers in Other (10 to 50 km)	**0.925**	0.329	0.157
Number of Workers in Services (10 to 50 km)	**0.924**	0.326	0.163
Number of Workers in Health (10 to 50 km)	**0.919**	0.343	0.169
Number of Workers in Manufacturing (0 to 20 min)	**0.665**	0.625	0.265

Three principal components extracted from 25 output variables and their scores

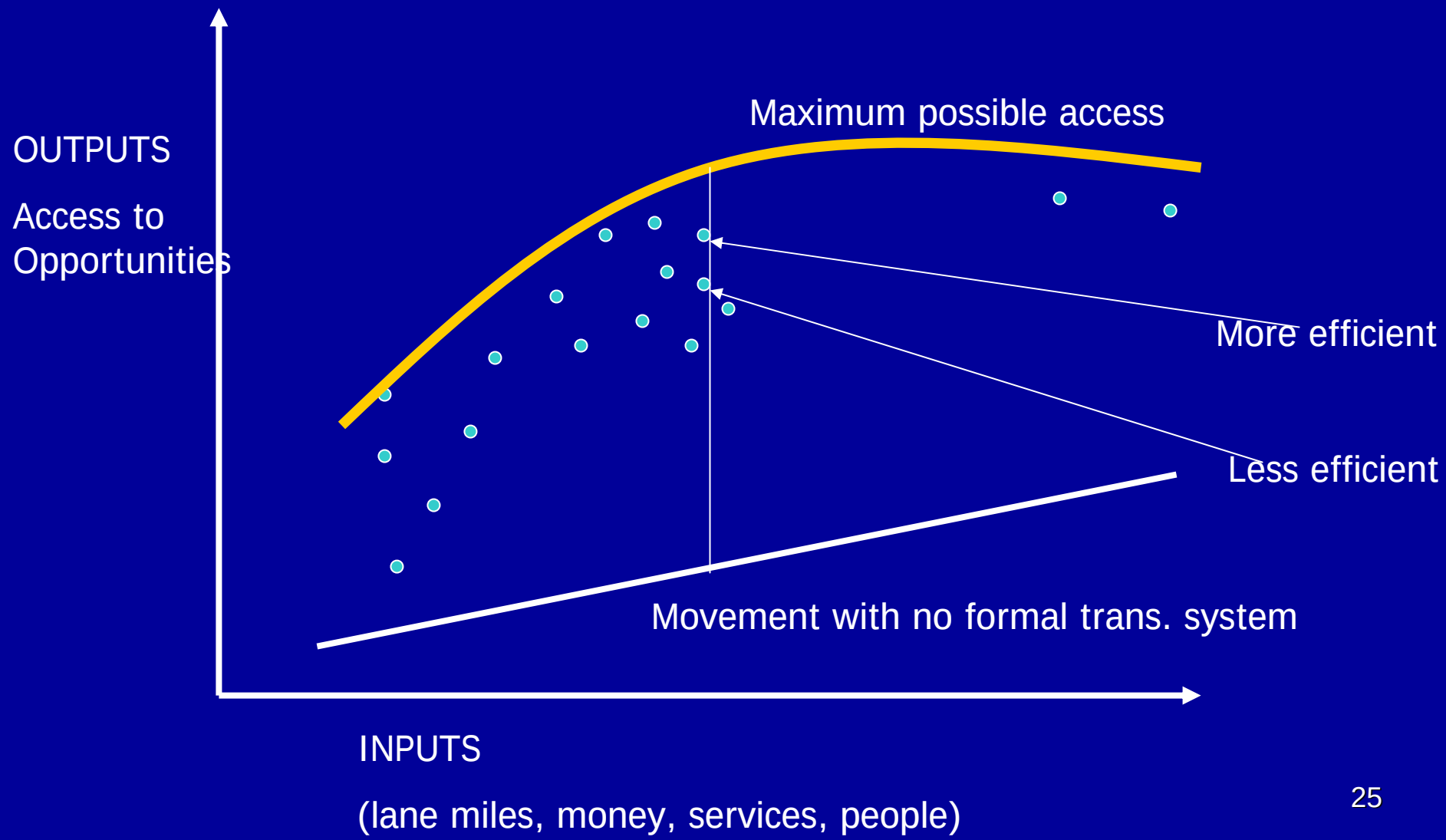
	Outer Ring Access	Middle Ring Access	Core Access
Number of Workers in Services (5 to 10 km)	0.234	**0.878**	0.296
Number of Workers in Retail (5 to 10 km)	0.322	**0.868**	0.275
Number of Workers in Other (5 to 10 km)	0.380	**0.841**	0.289
Number of Workers in Health (5 to 10 km)	0.267	**0.817**	0.350
Number of Manufacturing in Services (5 to 10 km)	0.438	**0.766**	0.220
Number of Workers in Services (0 to 20 minutes)	0.504	**0.703**	0.430
Number of Workers in Health (0 to 20 minutes)	0.532	**0.688**	0.421
Number of Workers in Retail (0 to 20 minutes)	0.585	**0.680**	0.389
Number of Workers in Other (0 to 20 minutes)	0.605	**0.672**	0.345

Three principal components extracted from 25 output variables and their scores

	Outer Ring Access	Middle Ring Access	Core Access
Number of Workers in Services (0 to 5 km)	0.071	0.198	**0.955**
Number of Workers in Retail (0 to 5 km)	0.139	0.226	**0.942**
Number of Workers in Other (0 to 5 km)	0.190	0.325	**0.871**
Number of Workers in Health (0 to 5 km)	0.075	0.308	**0.839**
Number of Workers in Manufacturing (0 to 5 km)	0.289	0.354	**0.699**

Efficiency Analysis (Stochastic Frontiers)

“Optimality”


OUTPUTS
Access to Opportunities
Maximum possible access
More efficient
Less efficient
Movement with no formal trans. system
INPUTS
(lane miles, money, services, people)

$$y_i = \alpha + x'_i \beta + v_i - u_i$$

Access Variable Factor (core, middle, outer)

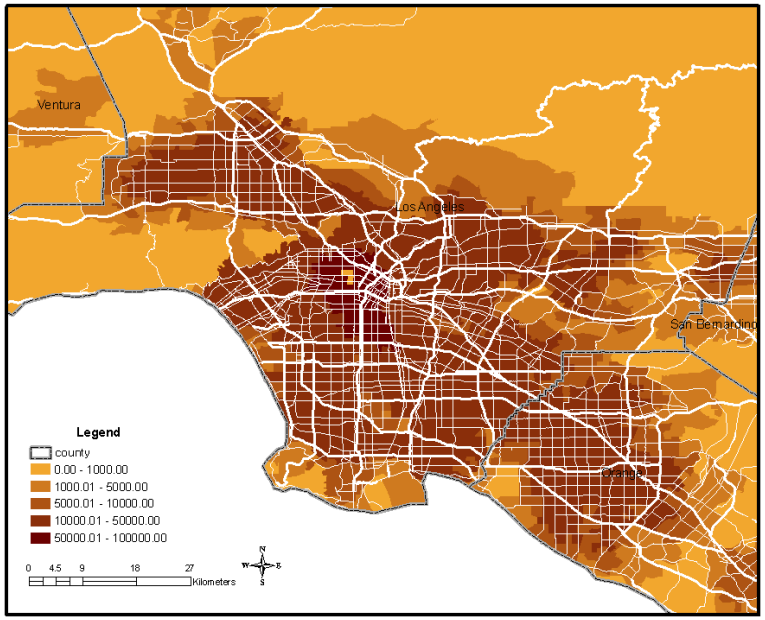
Highways Available & Other vars

Coefficients to Estimate

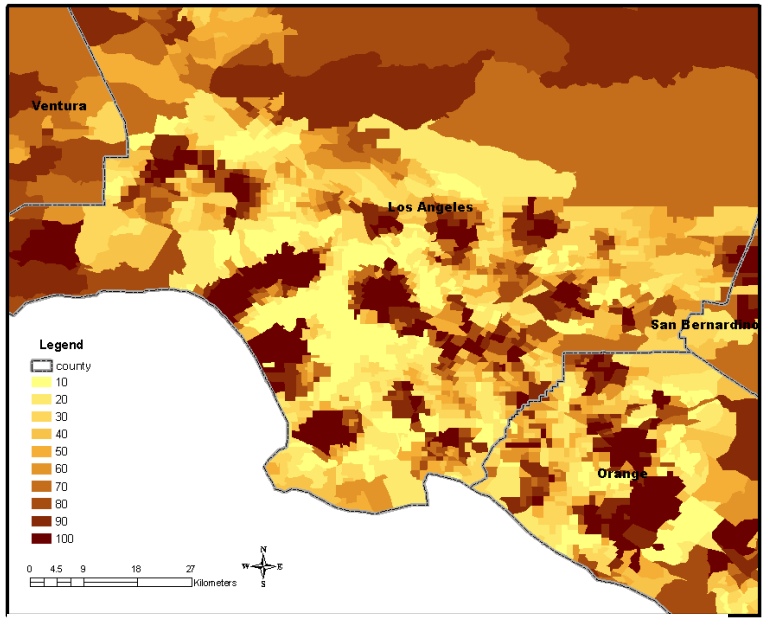
Measurement Error

Measure of Efficiency

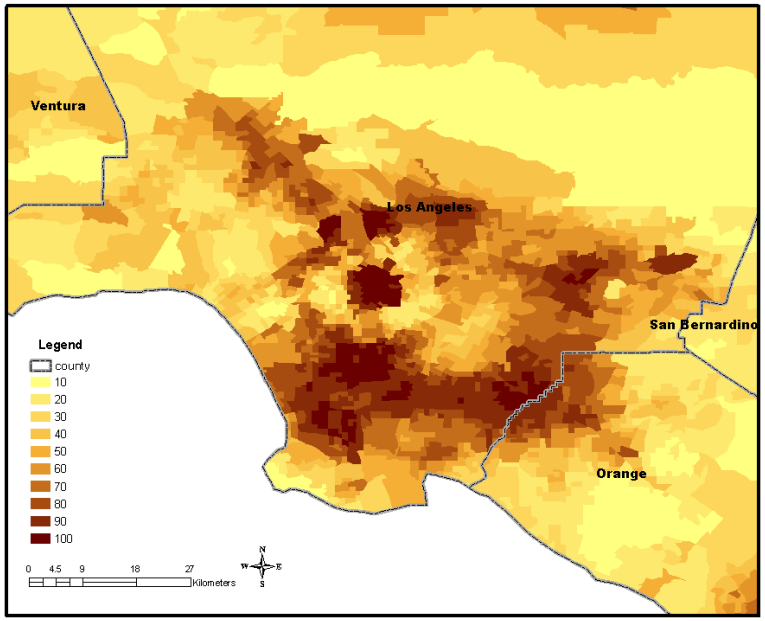
LA: TOTAL LANE KILOMETERS within 5km



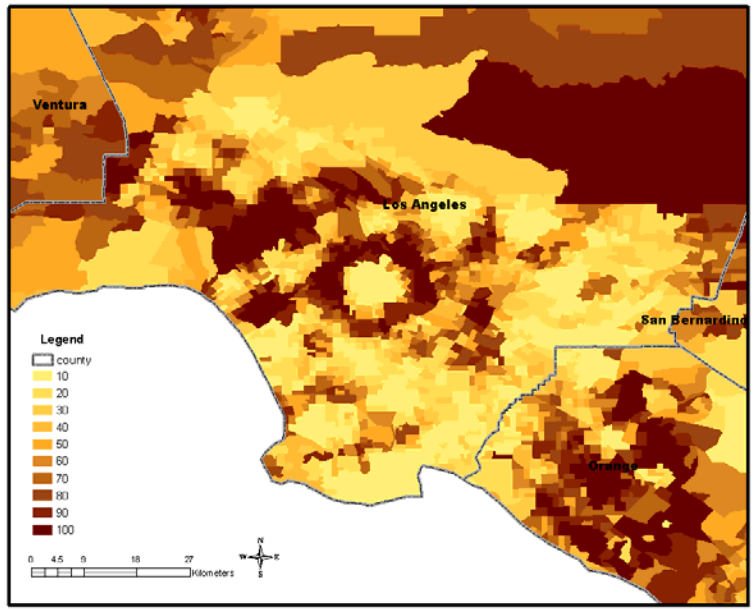
Core Efficiency



Outer Ring Efficiency



Middle Ring Efficiency



	Core	Middle Ring	Outer Ring
Median Efficiency Value	81%	92.4%	83.8%
Tenth Lower Percentile	62%	83%	72%

EFFICIENCY IS DIFFERENT FOR EACH TYPE OF ACCESSIBILITY MEASURED

Some Findings

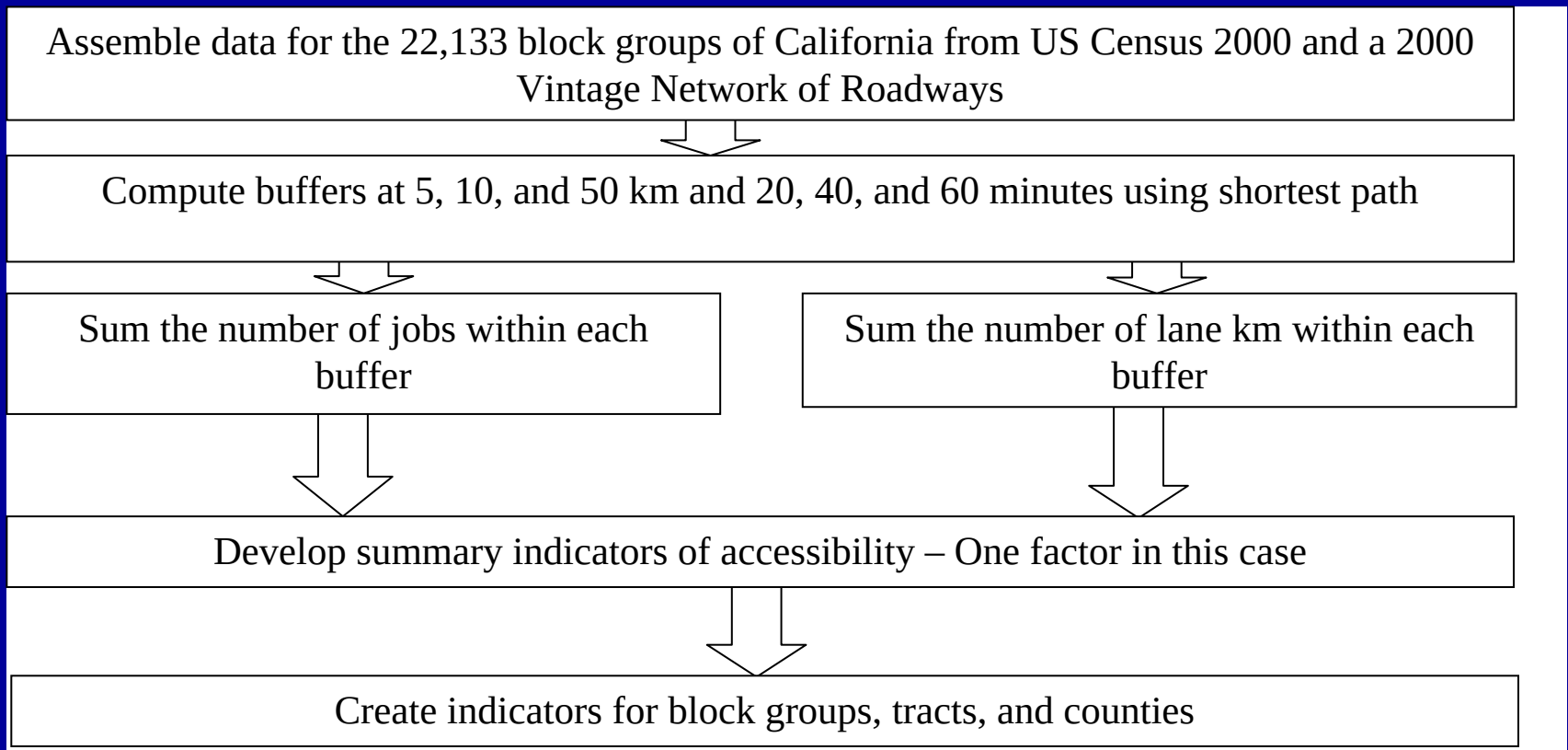
- The presence of primary roadways has a strong positive impact on access.
- For core access the secondary roadways seem to have a much higher impact and merit attention for investment.
- Efficiency in the transformation of roadways to access depends on the residents of each tract.
- Efficiency also depends on the measurement of access (outer ring vs. middle ring).
- More detailed analysis possibly using much smaller geographical areas is needed to pinpoint differences.

(In)Equality

Objectives

- Inequality is very often limited to a few disadvantaged population segments
- We need a Global measure with local connections
- It should show over or under provision of accessibility when compared to other localities and the amount of residents
- Account for strong spatial correlation among accessibility indicators - due to the connectivity of trans. network and business location practices
- Prefer a fractal indicator that can be decomposed all the way to a parcel of land and social segments
- Not happy with work on Census tracts – go one step further down to block groups

Computational Schema



The Theil Index

$$T = \sum_{i=1}^{22,133} \frac{a_i}{A} \log \left(\frac{\frac{a_i}{A}}{\frac{n_i}{N}} \right)$$

The Theil Index

Locations are the 22,133 block groups in CALIFORNIA

Accessibility indicator for location i

$$T = \sum_{i=1}^{22,133} \frac{a_i}{A} \log \left(\frac{\frac{a_i}{A}}{\frac{n_i}{N}} \right)$$

Persons at location i

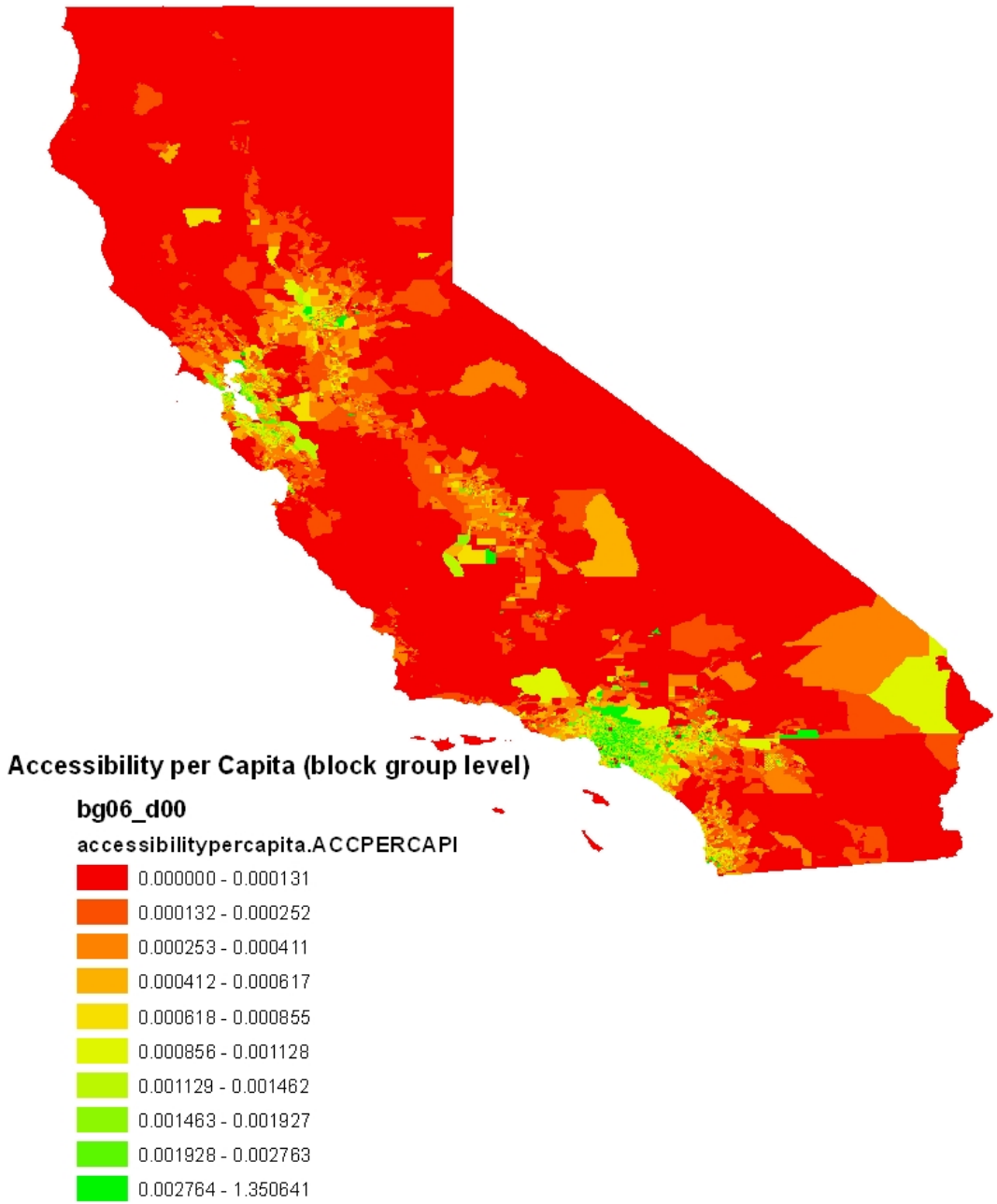
Total accessibility in CALIFORNIA

Total population in CALIFORNIA

The Theil Index

- Each component in the Theil index is a **weighted measure** of the mismatch between its accessibility share and its population share.
- We'll focus on each term of the sum, which we name contribution of the block group to the Theil index, or Theil contribution.

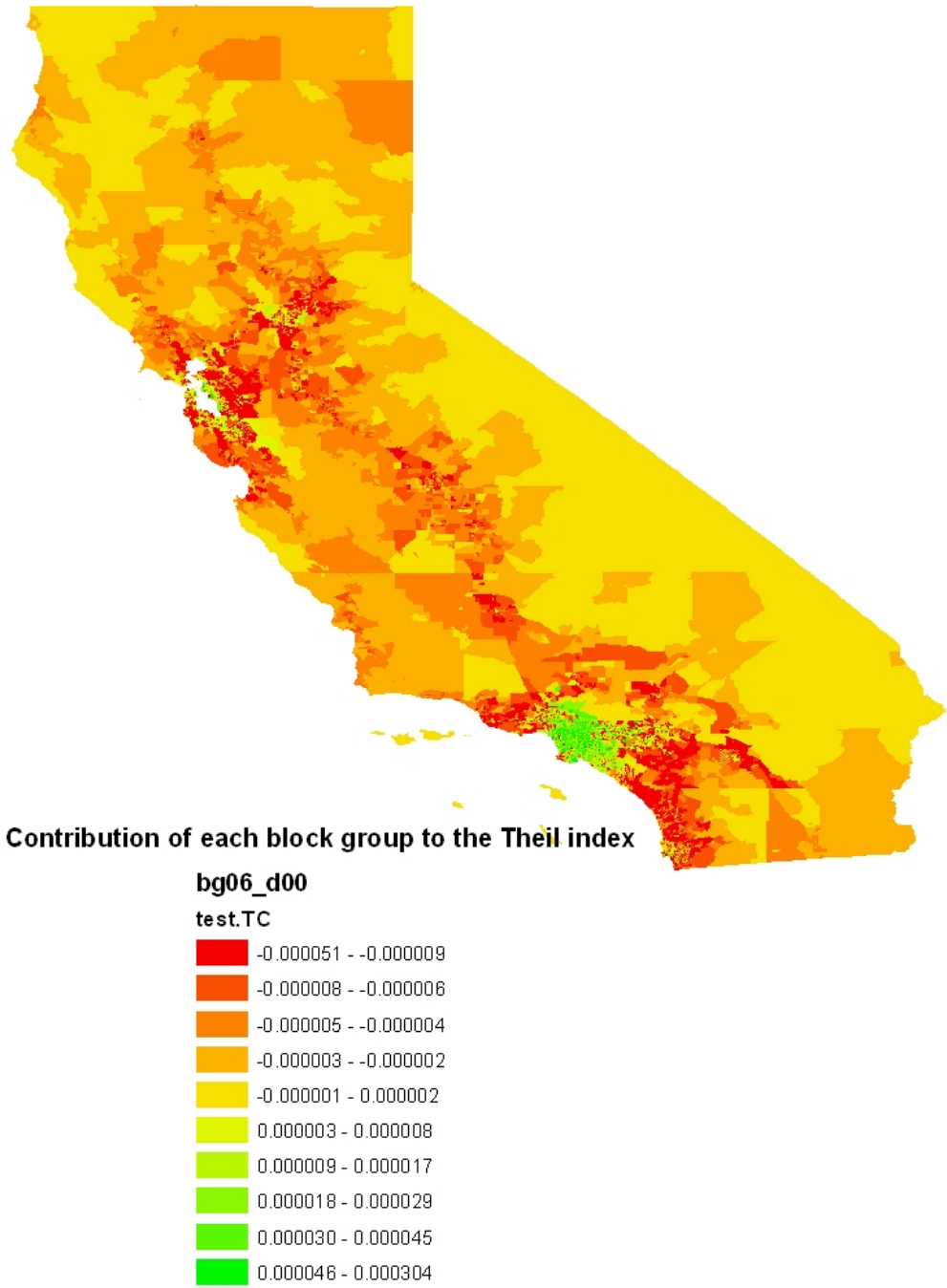
ai/ni


Accessibility per Capita (block group level)
bg06_d00
accessibilitypercapita.ACCPERCAPI
0.000000 - 0.000131
0.000132 - 0.000252
0.000253 - 0.000411
0.000412 - 0.000617
0.000618 - 0.000855
0.000856 - 0.001128
0.001129 - 0.001462
0.001463 - 0.001927
0.001928 - 0.002763
0.002764 - 1.350641

Theil index contribution of each locality

Positive (green) = better than the average locality (block group)

Negative (yellow-red) = worse than average locality (block group)



But block groups are connected into a network!

Also, people live in one locality and work in another

=>

Need to go up and down geographic aggregation and look at within region distribution

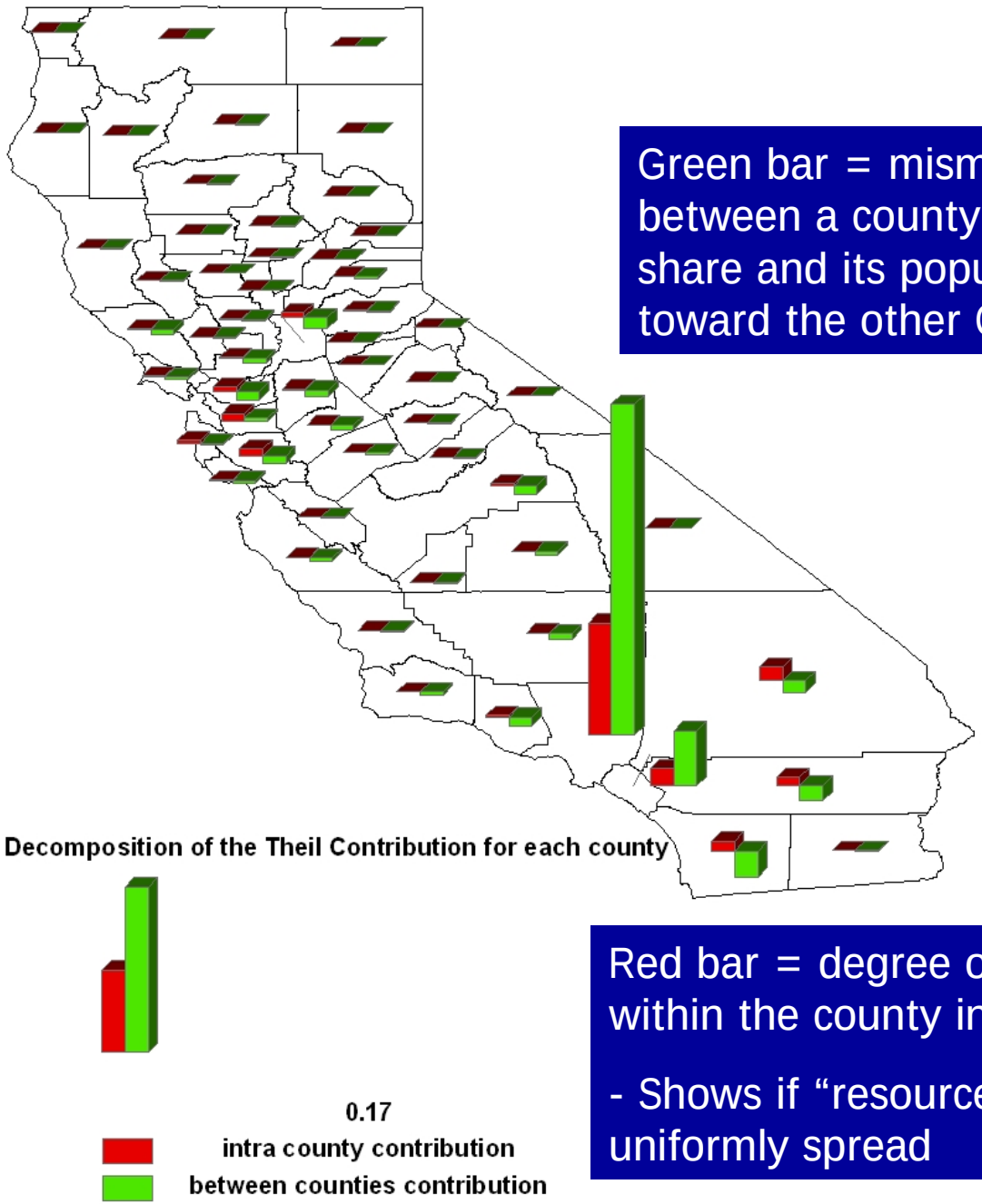
Fractal Nature of Theil index

$$T_{CA} = \sum_{i=1}^{NC} \frac{a_i}{A_{CA}} . \log \left(\frac{\frac{a_i}{A_{CA}}}{\frac{n_i}{N_{CA}}} \right) + \sum_{i=1}^{NC} \frac{a_i}{A_{CA}} . \left[\sum_{\text{Tract } j} \frac{a_j}{A_{\text{Couty } i}} . \log \left(\frac{\frac{a_j}{A_{\text{Couty } i}}}{\frac{n_j}{N_{County\ i}}} \right) + \sum_{\text{Tract } j} \frac{a_j}{A_{\text{Couty } i}} . \left(\sum_{\text{Blockgroup } k} \frac{a_k}{A_{\text{Tract } j}} . \log \left(\frac{\frac{a_k}{A_{\text{Tract } j}}}{\frac{n_k}{N_{\text{Tract } k}}} \right) \right) \right]$$

Relative Contribution to Total Index Between Counties

Relative Contribution to Total Index Within Each County

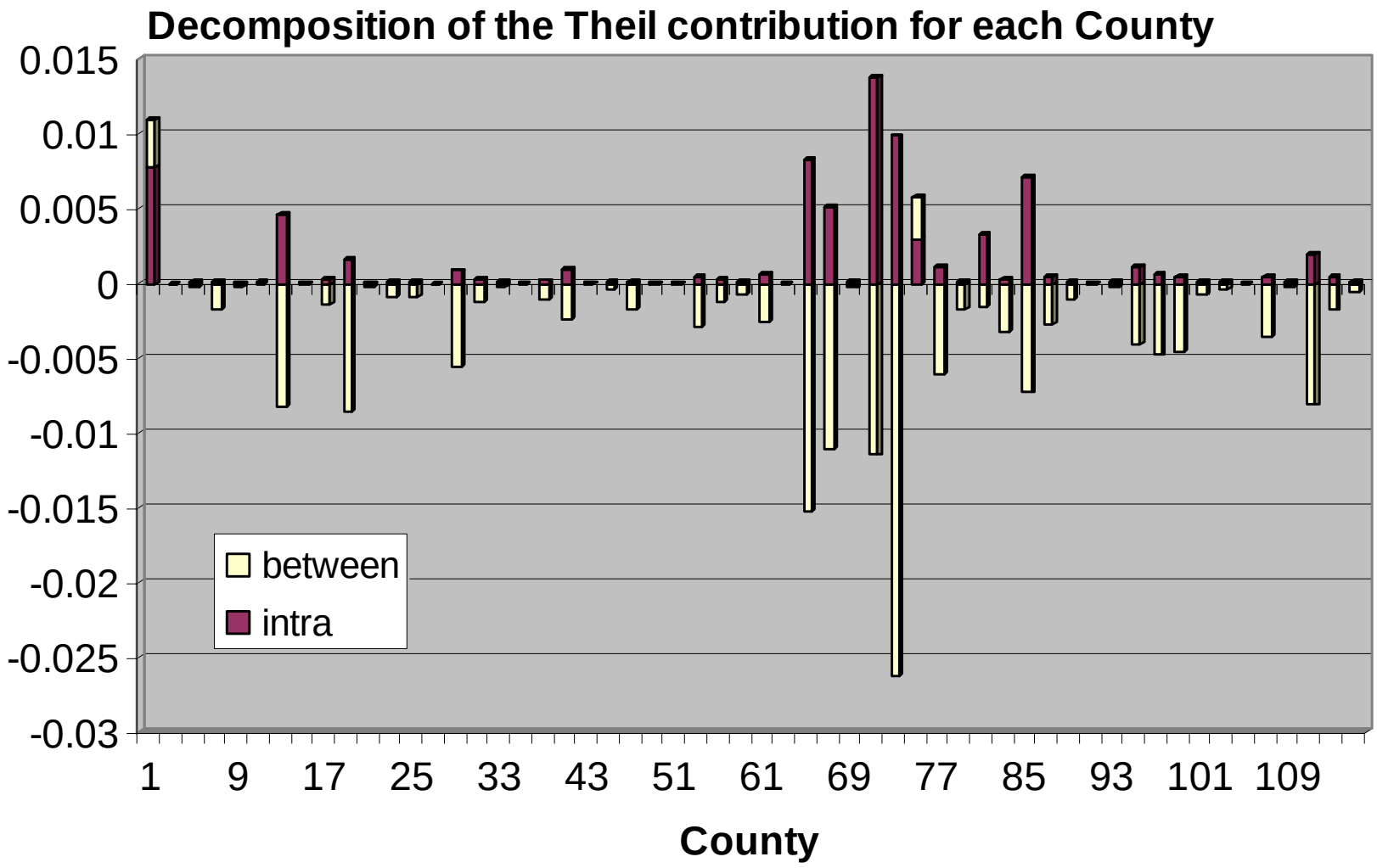
LA & Orange above average by far



Green bar = mismatch that exists between a county's accessibility share and its population share toward the other Counties

Red bar = degree of uniformity within the county in accessibility

- Shows if "resources" are spatially uniformly spread



Graph without LA and Orange

Findings

- Spatial structure of inequality and its distribution through different geographic levels can be done
- Theil Index helps us to identify areas that need the most transport infrastructure for an egalitarian development of the State
- It also allows to study regions that have allocated their investments (infrastructure in this case) targeting within an area homogeneous development
- It also allows to target investments at a more local and focused way while at the same time controlling for the impact to everybody else

Are we done ?

- Definitely NO!
- Congestion effects need careful examination – included implicitly but not well
- Continue more work on individuals and households behavior – function of multiple indicators of service available at different geographic scales
- Continue work on service locations instead of number of persons employed in sectors (health, retail, leisure etc)
- Start collecting the data and building the databases to map $$$$ to facilities to Level of Service
- Where is public transportation?
- Do all of the above longitudinally (time and space)

Thank you and Questions?

Kostas Goulias (805) 893-4190
goulias@geog.ucsb.edu