

“Gis a job”: What use
geographical information systems
in spatial economics?

Henry Overman (LSE and SERC)

Introduction

- GIS widely used in business, government and academia
- Clearly “useful”, but how useful to me?
 - Help deal with data
 - Avoid arbitrary discretization
 - New sources of data
 - Innovative instrumental variables
 - Cutting edge advances and the deep problems of spatial economics?

Defining GIS

- Something that does the following for spatial data:
 - Inputting
 - Storing
 - Managing
 - Analysing
 - Mapping

Inputting

- Two main sources of data
 - Ground survey (a particular location)
 - Census (distribution across locations)
- Remote sensed data
 - Aerial photography
 - Satellite
 - Google views

Inputting (cont)

- Outsource
 - Analogue to digital via digitizing/scanning
 - Processing of remote sensed
- Understand sources of measurement error
 - GIS specific (lake not closed, roads don't cross)
 - Very general (missing out a line of data)
 - Social scientists think differently about (the *consequences* of) measurement error
 - E.g. Land use data and urban sprawl
 - GPS, survey data and measured versus perceived distance

Storing/Managing

- Raster
 - Values to each cell on a regular grid
- Vector
 - Values to irregular polygons plus coordinates
- You can overlay (move between them)
- Understand sources of measurement error
 - GIS specific
 - Non-overlapping polygon boundary
 - Mis-alignment of grids

Storing/Managing

- Rules for merging layers that are recorded for different (non-nested) spatial units
 - Depends on underlying spatial distribution
 - Continuous surface data (e.g. kriging)
 - Statistical properties of matched survey data

Analysis

- GIS quick at analysing spatial data
- Identify observations by characteristics and location
 - Study of land use
 - Hedonics of house prices
- Measurement
 - Length of lines
 - Perimeter or area of polygons
 - Distance between things (physical/network/social)
 - Shape

Analysis (cont)

- Shape
 - [Sinuosity?]; Back gardens and their impact on house prices (Cheshire and Sheppard)
- Distance
 - Plus absolute barriers and movement costs (discrete/continuous) → least cost path
 - Widely used in transport economics
 - Donaldson: impact of India's rail network
 - Faber: least cost path plus spanning algorithm to get least cost network as instrument for actual roads (in the spirit of Baum-Snow, Michaels, Duranton/Turner)

Analysis (cont)

- Distance to other features
 - Hedonics
- Distance to other observations (spatial arrangement)
 - Spatial point pattern analysis
 - Increased availability of geo-referenced data → more use in spatial economics
 - Good algorithms for complex one-off calculations (e.g. nearest neighbour)
 - Bootstrapping and significance values → non-GIS?
 - Increasingly convinced sunk costs worth incurring for one-off calculations

Analysis (cont)

- Buffers
 - Burchfield et al: urban fringe overlying aquifers
 - Duranton et al: all firms within given distance of jurisdictional boundaries
- Avoids arbitrary discretisation
 - Arbitrary wrt to what?
 - Significant variation in unobserved conditions within spatial units?

Mapping

- Spatial equivalent of “plot the data”
- Problems
 - Non GIS specific (how to bin data for choropleth maps)
 - GIS specific (how to overlay features at the same location)

Increasing our understanding

- Reduces measurement error
 - Particularly wrt location of observations
 - Ignored up to now because not a lot could do about it; set to change with geo-referencing?
- Arbitrary discretization
 - Two common solutions for spatial units
 - Identify “functional units”
 - Present results at different spatial scales
 - New solutions
 - Overlay of non-nested spatial units
 - Work in continuous space (but constraints in GE?)

Increasing our understanding

- New sources and types of data
 - Climate (Rapport 6000 meteorological stations to 2x2km grid, covering 20 variables)
 - Terrain (Nunn and Puga ruggedness using 1x1km *global* elevation data)
- As sources of exogenous variation
 - Streams and public schools (Hoxby)
 - Gradients and Dams (Duflo and Pande)
 - Agglomeration economies
 - Building height and geology (Rosenthal and Strange)
 - Terroir (Combes et al)
- Treatment effects literature
 - Timing and *exact* location of sex offenders (Linden and Rockoff)

What about the cutting edge?

- The deep issues of empirical spatial economics
 - Role of theory in structuring empirical analysis; achieving identification
 - Treatment effects; IV or researcher generated sources of exogenous variation

The GIS bread and butter diet

- International Cartographic Association
Workshop on Geospatial Analysis and Modelling
 - Cellular automata and agent based modelling;
 - Visual analytical tools for environ/urban systems;
 - Analysis of human movement data;
 - Spatio-temporal data mining;
 - Hierarchies, scaling and fractals;
 - Modeling vehicle dynamics and crowd behaviour;
 - Patterns of human spatial behavior and migration;
 - Urban high-resolution morphology;
 - Small world modeling and spatial interactions.
- Synergies physics/biology side regional science
- Will we be sticking to our bread and butter diet?