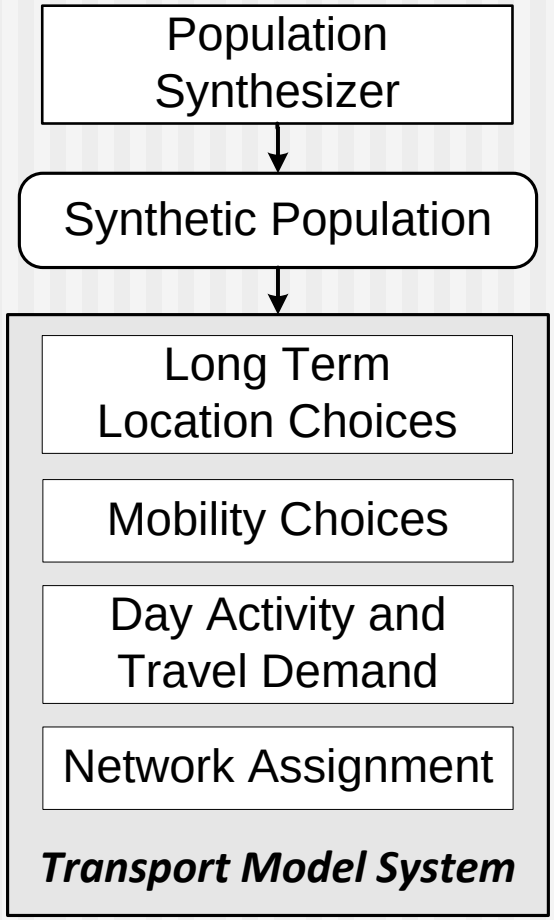
Population Synthesis Challenges

Activity-Based Modelling Symposium

Research Centre for Integrated Transport and Innovation (rCITI)
UNSW, Sydney, Australia
March 10, 2014

John L Bowman, Ph.D.
John_L_Bowman@alum.mit.edu
JBowman.net

Synthetic population as AB model input



Typical synthetic household

Household			
Household income	150K		
Residence location	11111		
Person	**Adult 1**	**Adult 2**	**Child**
age in years	41	40	12
gender	Male	Female	Female
worker status	FT	PT	nonworker
student status	nonstudent	nonstudent	Student

Synthesizing households for one zone using IPF

1. Detailed distribution

	Small HH	Large HH
Low Inc	100	50
High Inc	50	50

2. Control totals

	Small HH	Large HH	
Low Inc			150
High Inc			150
	200	100	

3. Iterative Proportional Fit

	Small HH	Large HH	
Low Inc	111	39	150
High Inc	89	61	150
	200	100	

4. Draw HH from Microdata sample

(e.g., draw 111 small, low inc HH from zone 1's district)

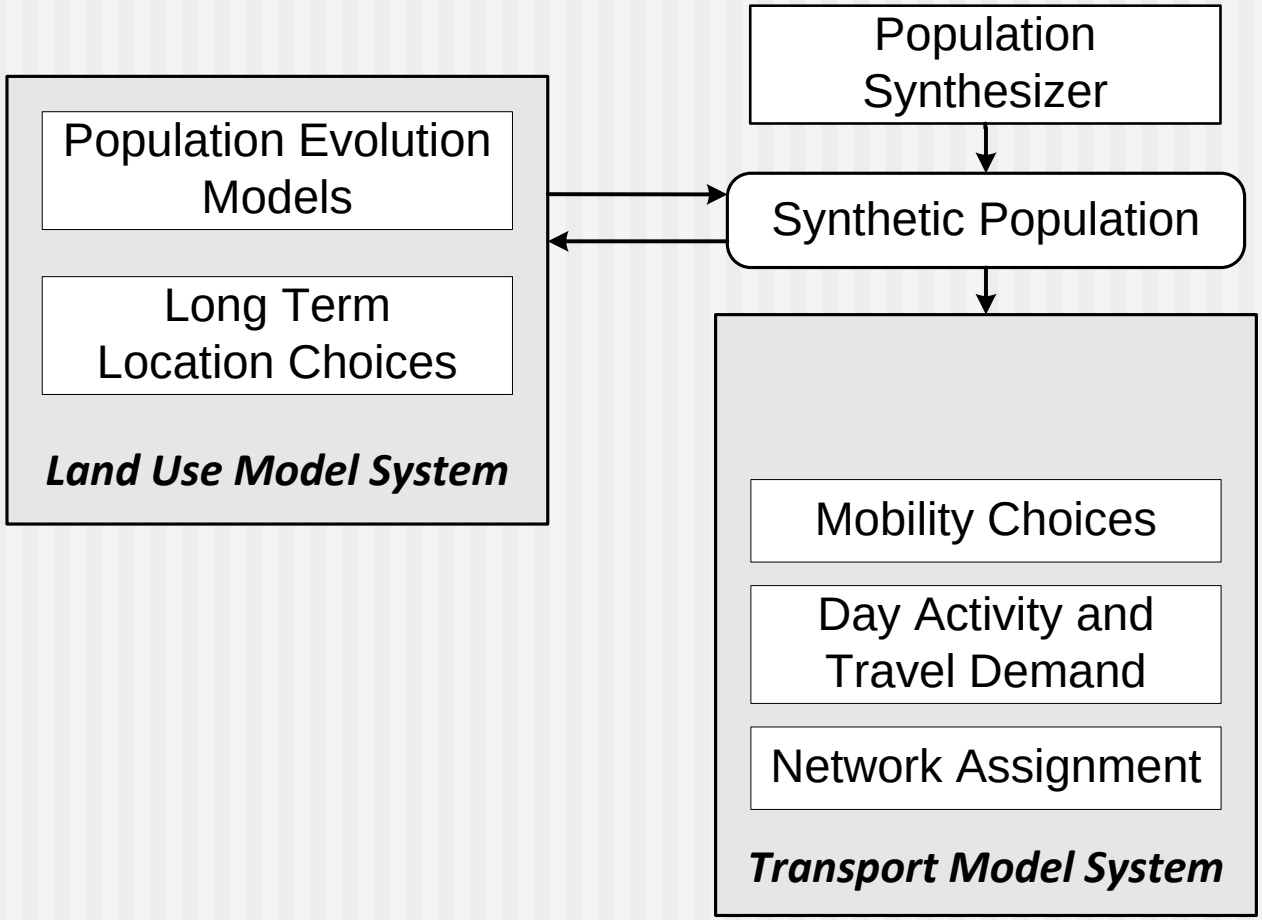
Typical Set of Control Categories for IPF

	Household Categories Defining Cell					
ID	Income	House-holder age	HH Size	Family	Children	Number employed
1	0-20K	15-64 yrs	1	nonfamily	0	0
2	"	"	"	"	"	1
3	"	"	2	nonfamily	0	0
4	"	"	"	"	"	1
5	"	"	"	"	"	2
6	"	"	"	family	0	0
7	"	"	"	"	"	1
8	"	"	"	"	"	2
9	"	"	"	"	1	0
10	"	"	"	"	"	1
11	"	"	"	"	"	2
⋮						
316	100K+	65+ yrs	5+	family	0+	3+

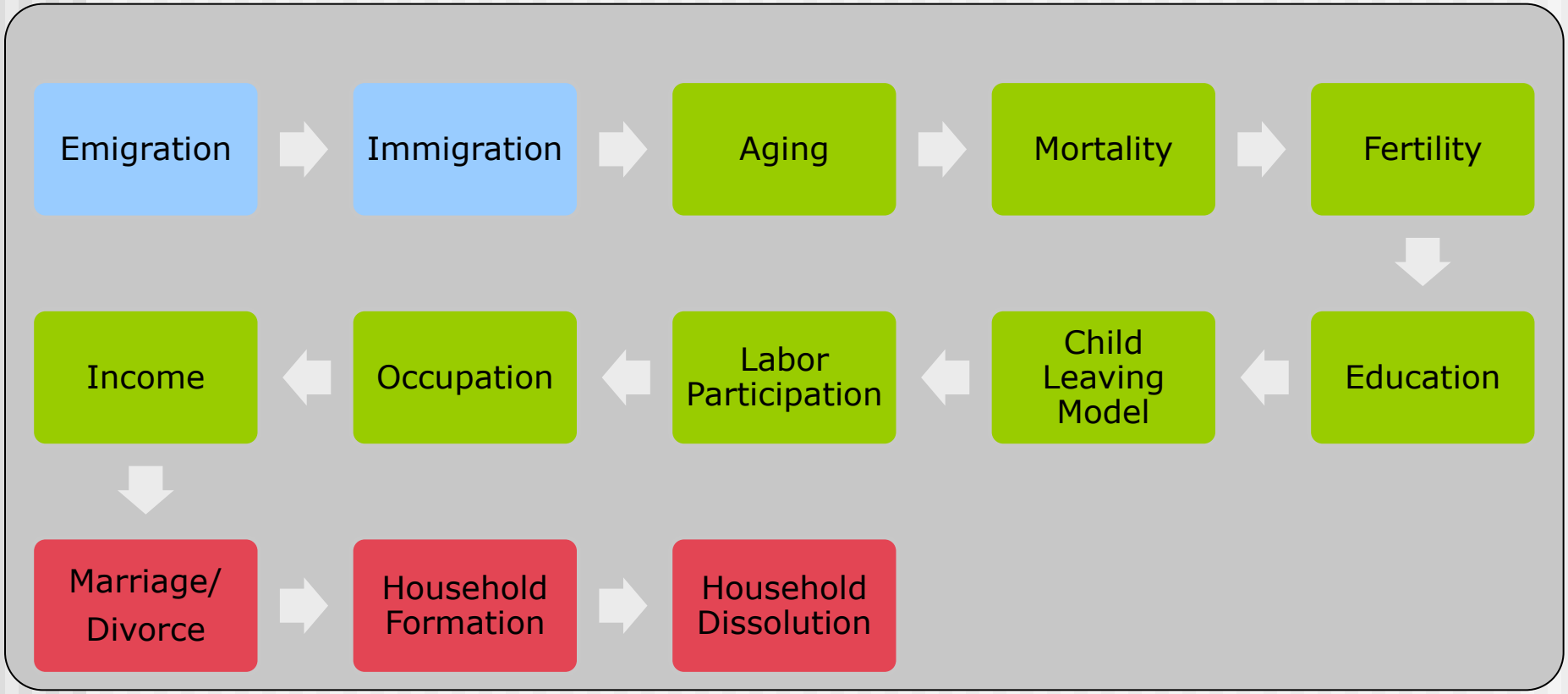
Newer population synthesizers…

- Also control person characteristics
  - Gender
  - Age
- E.g.
  - PopGen (ASU)
  - PopSyn 2 (San Diego)

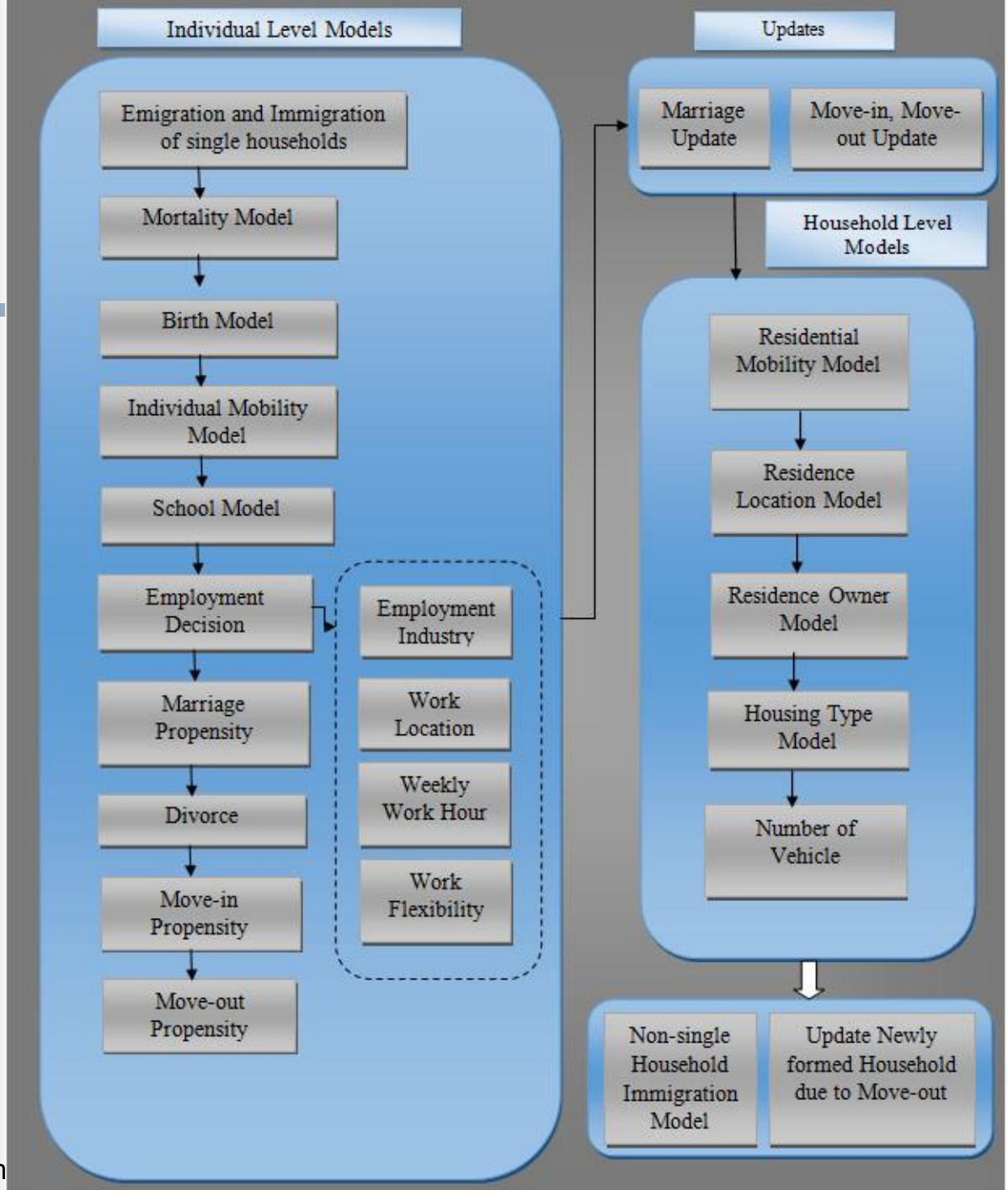
Population evolution approaches



Population Evolution Prototype
Baltimore



UT Austin



Activity-Based Model System

Challenge is producing a good population for future years

- AB models work best with
  - fine spatial resolution
  - many household and person characteristics
- Demographic forecasters can't really provide this for population synthesis
- Future synthetic populations look much like current population except for a few variables controlled at an aggregate geographic level

Do population evolution models provide an answer?

- Data for estimating and validating them is scarce
  - high resolution data on dynamics of population characteristics
- If data were available would the future track past trends?