Household Allocations and Land Development Modeling Components

JD Hunt
JE Abraham

Third Oregon Symposium on
Integrated Land Use and Transport Models

Portland OR, USA
23-25 July 2002

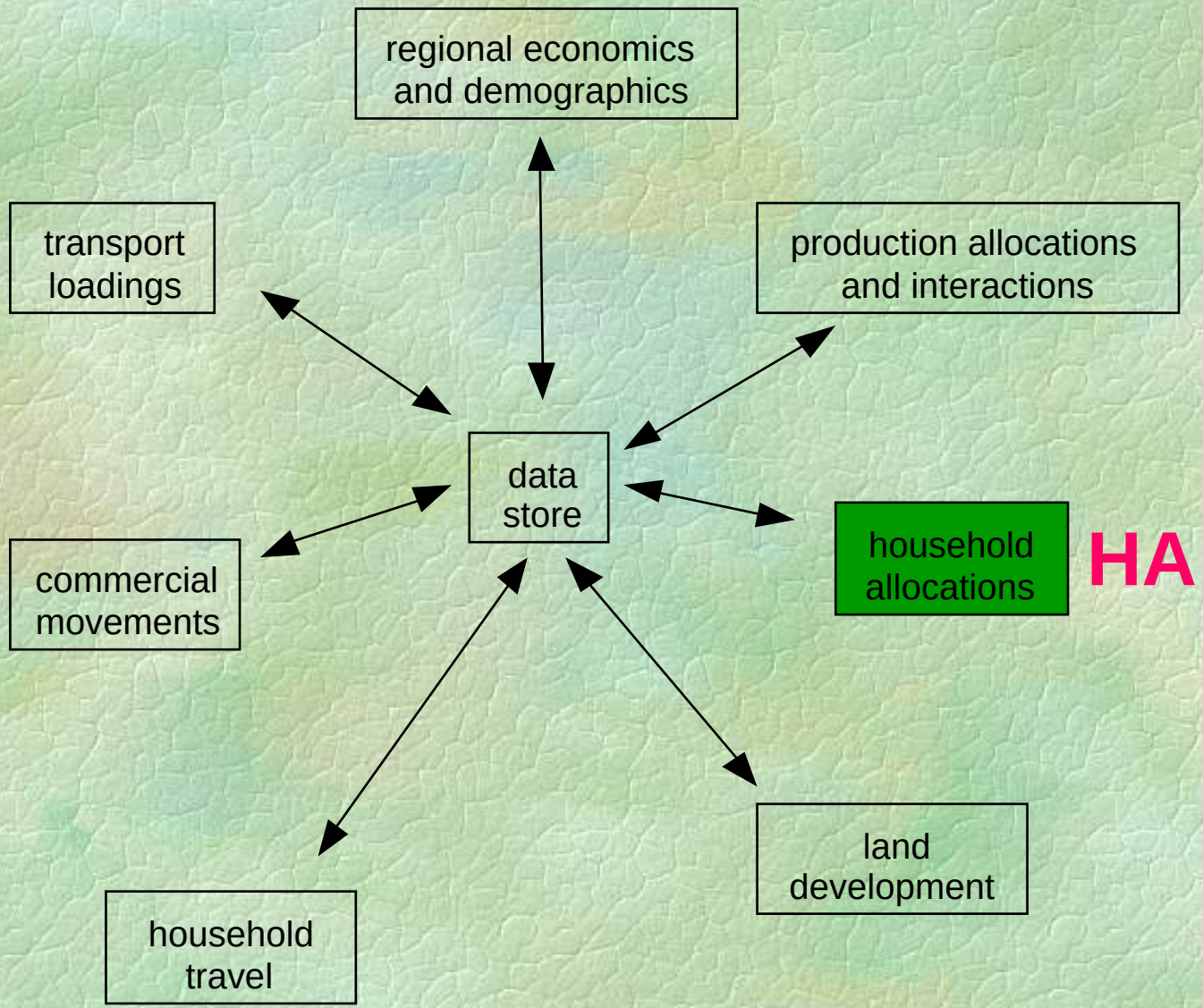
Outline

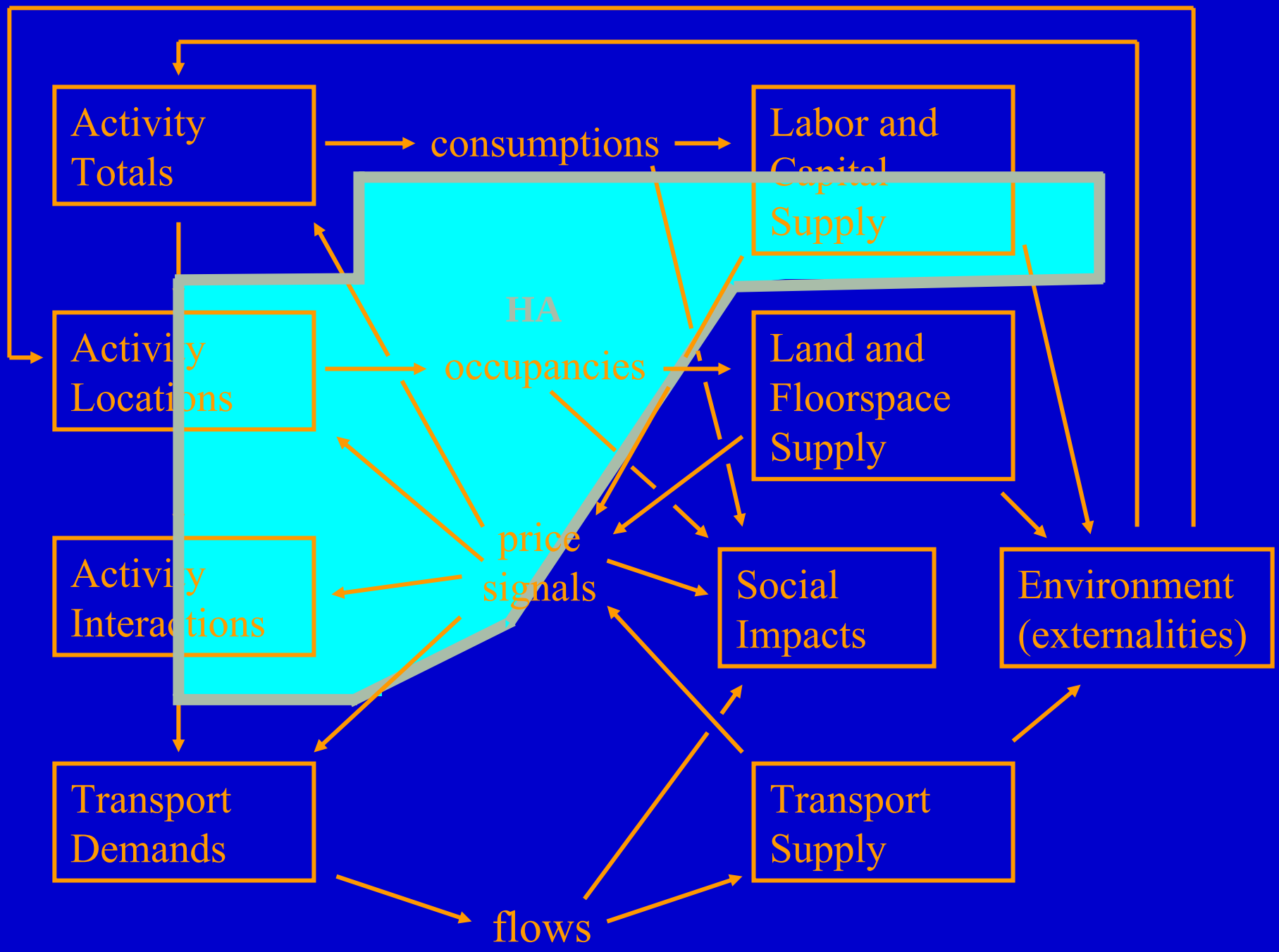
Describe Two Linked Disaggregate Modules

- **HA:** Household Allocations
  - Framework
  - Examples
  - Inputs and Outputs
- **LD:** Land Development
  - Framework
  - Examples
  - Inputs and Outputs
- Joint time series behavior
- Conclusions

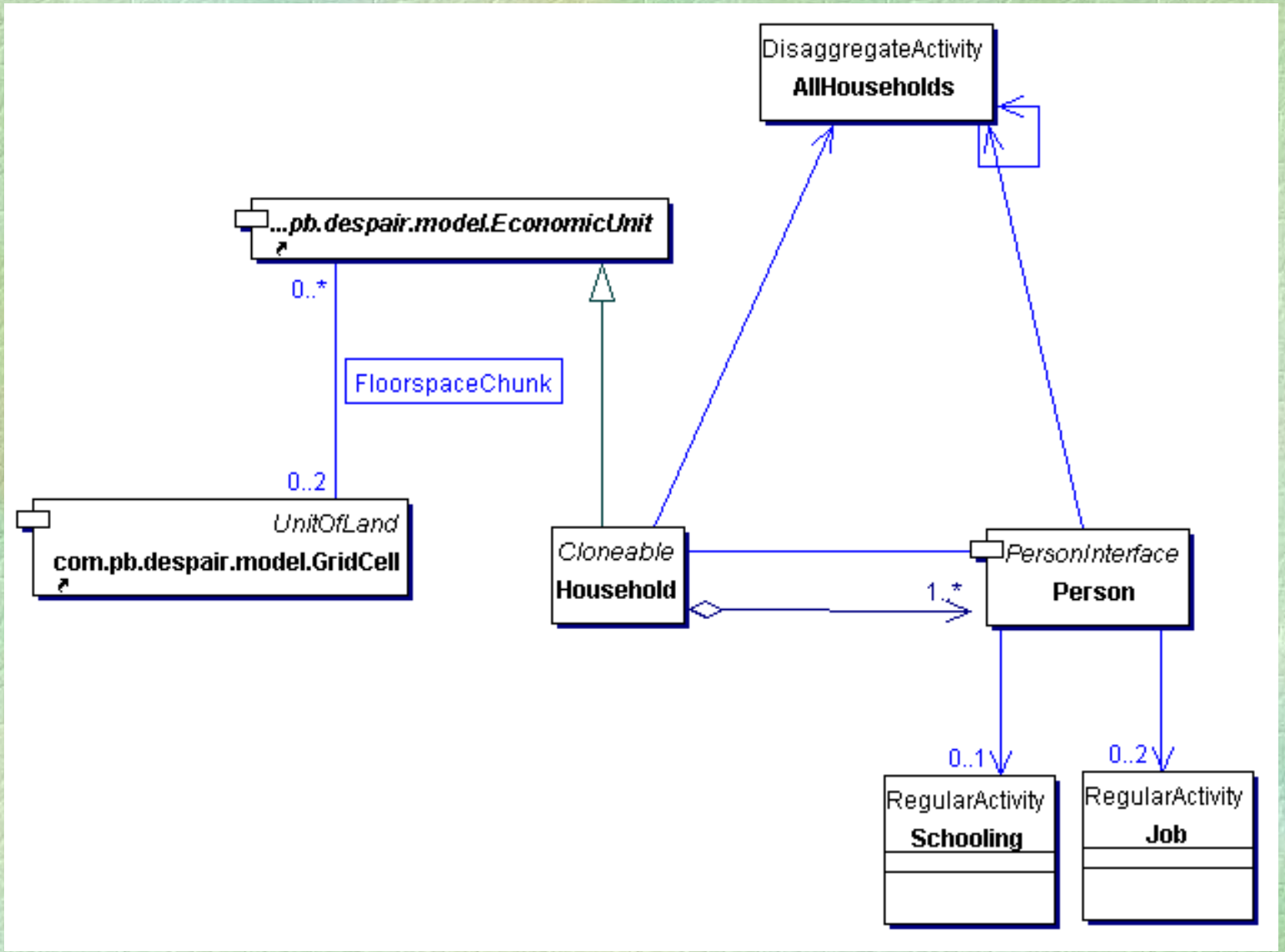
Model Framework; Treatment of Behaviour

Household Allocations




Activity Totals
consumptions
Labor and Capital Supply
HA
Activity Locations
occupancies
Land and Floorspace Supply
price signals
Activity Interactions
Social Impacts
Environment (externalities)
Transport Demands
Transport Supply
flows

HA Framework



Household

- Attributes:
  - household members, income, vehicles
- Methods:
  - births from female household members
  - decide whether to move
  - choose location (zone)
  - choose dwelling type (and associated grid cell) within zone
  - utility of zone
  - space requirements (# of rooms, square feet)
  - move in, move out, evicted

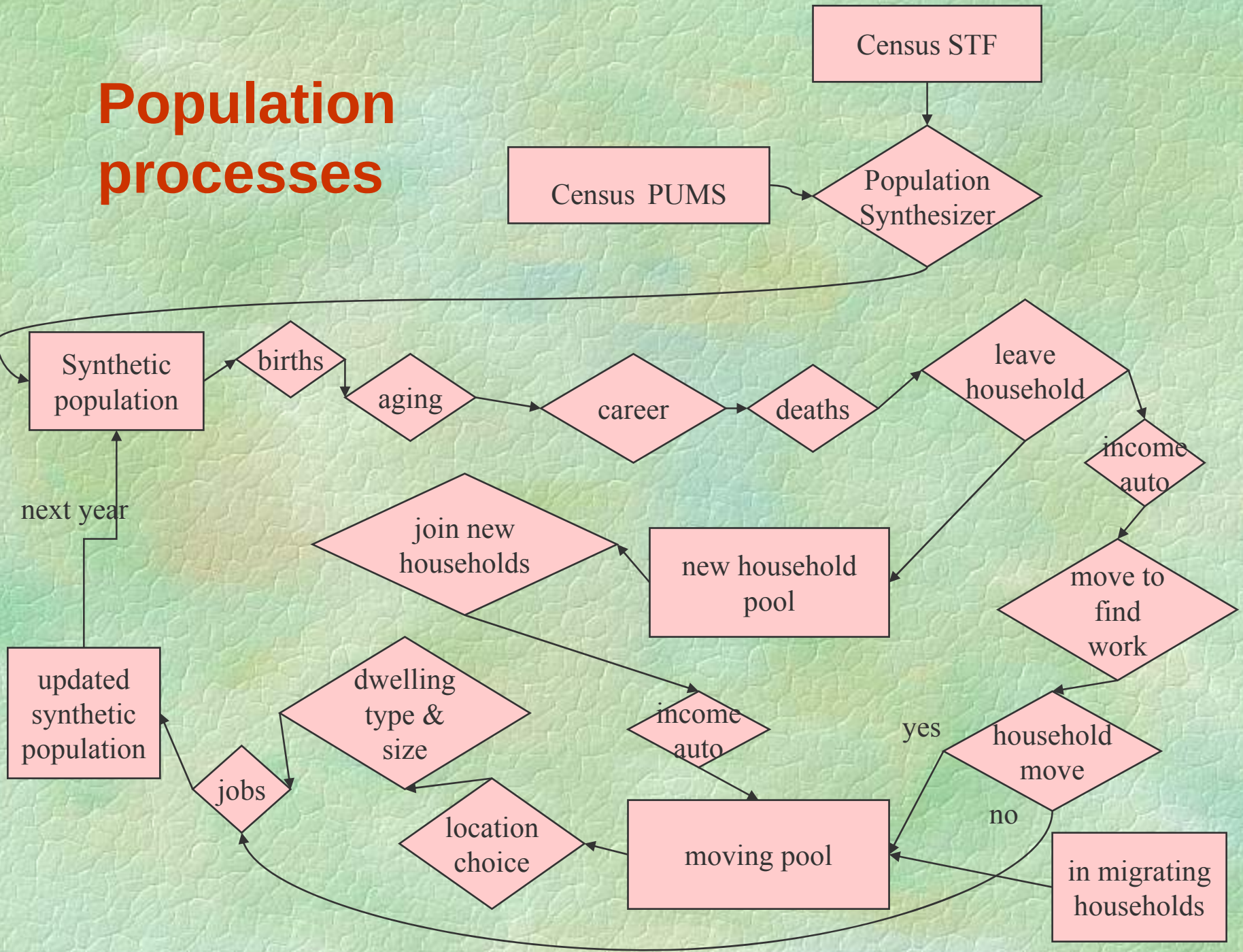
Person

- Attributes:
  - age, gender, occupation, work status, school status, education level
- Methods:
  - death, leaving household
  - occupation/work/school transitions
  - probability of giving birth

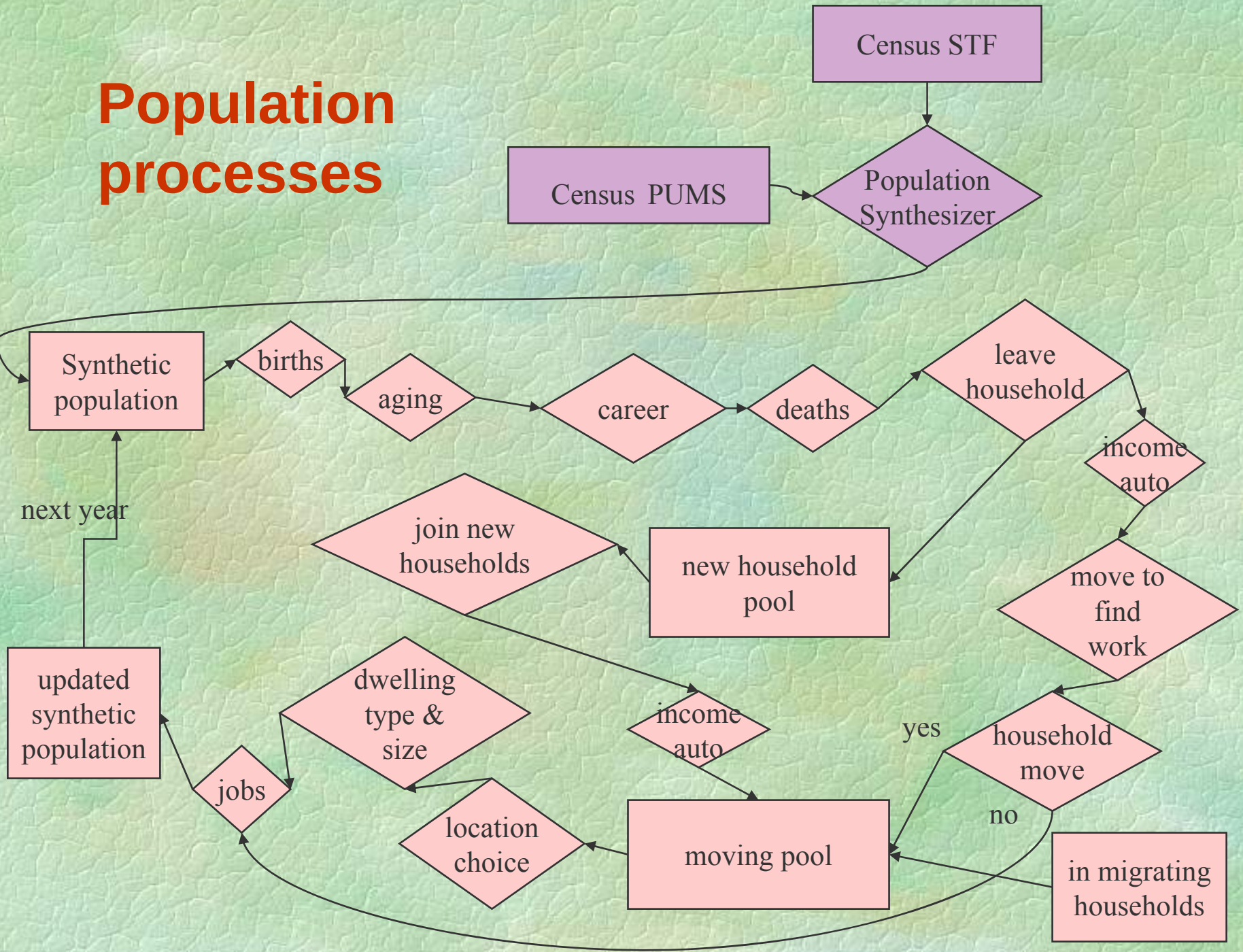
AllHouseholds

- Attributes:
  - household list, moving pool, secondary moving pool, new household pool
- Methods:
  - iterate through households
  - housing market representation
    - price and “size term” update procedures
  - create in migration households
  - create new households out of people in new household pool

Population processes



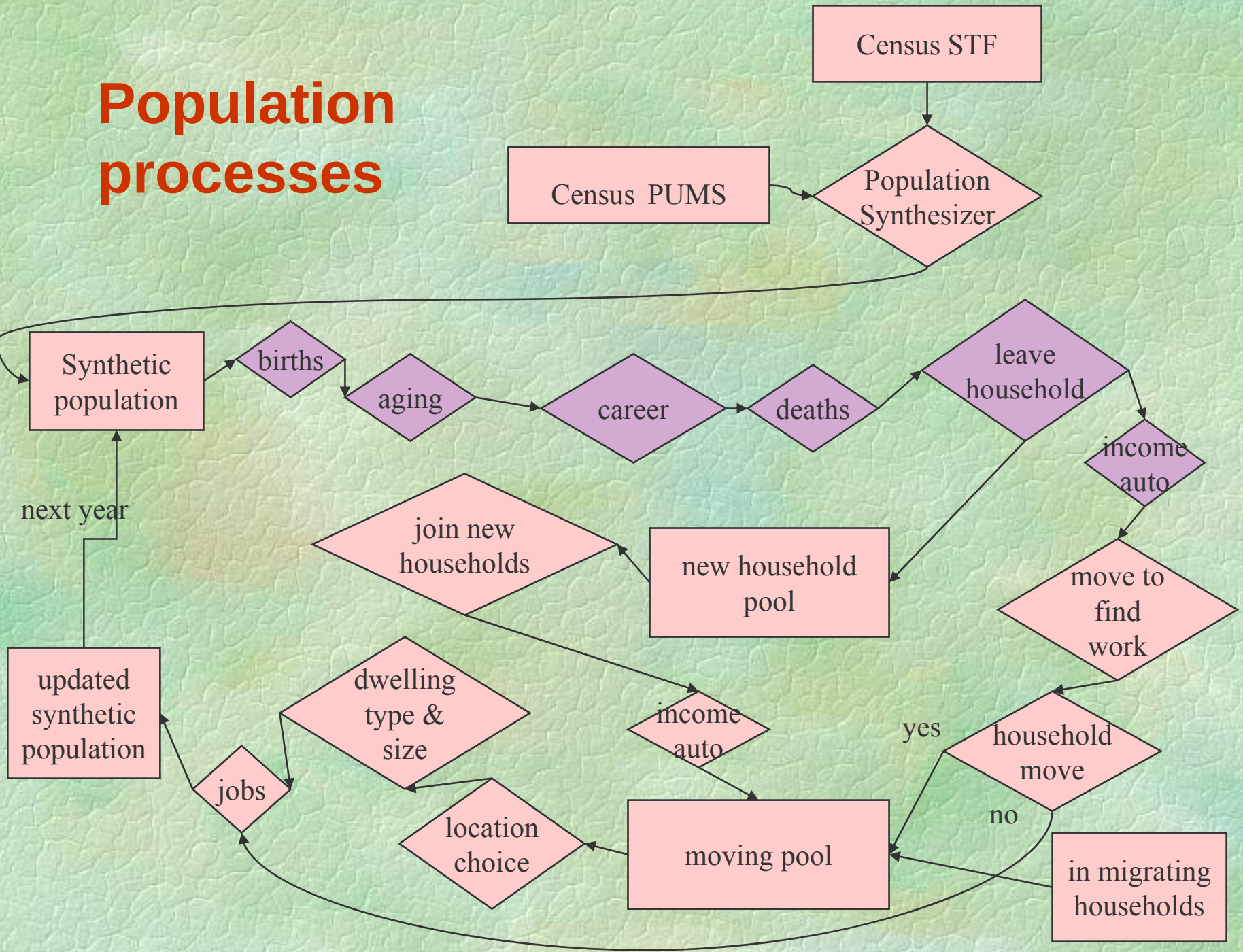
Population processes



Population synthesizer

- Based on two Census data sets
  - STF files describing aggregate attributes
  - PUMS showing individual household records, with exact location removed to protect privacy
- Sample individual records to match aggregate attributes
- Results in a population of individual households, assigned to STF geography, that match STF data

Population processes



Birth, death, career, leaving home

- Birth and death probabilities based on age/gender from census
- Career transition and leaving home based on Panel Study of Income Dynamics
  - tracks households through time
  - has individual income and individual job status
  - can identify whether individuals have left a household

Career transitions

- Logit choice model between each category
- Constants depend on current state
- Age included based on typical life progression
- Years of education
- Wage rates
- Matches with labor categories in PI

Not Working (and not looking for work)
Retired
Unemployed (and looking for work)
K-12 Student
College Student
Managers and Professionals - **1_ManPro**
Health care workers - **1a_Health**
Post Secondary Teachers - **2_PstSec**
Non-P.S. Teachers - **3_OthTchr**
Other Prof and Tech Office - **4_OthP&T**
Retail Sales Workers - **5_RetSls**
Other Retail & Clerical Office - **6_OthR&C**
non-office employment - **7_NonOfc**

Formula for logit career transitions

$$U_{newOcc} = K_{newOcc}^{oldOcc} + \phi_{newOcc}^{ageBand} + \phi_{newOcc}^{kids} + \phi_{newOcc}^{educLevel} + \theta_{newOcc} \cdot wage_{newOc}$$

- Currently working with simpler conditional model, but with more coefficients, eg:

Old Occupation :	Managers & Professionals (1)					
		Variables				
Alternatives	Male	Married	Kids < 5	Age Gp. 1	Age Gp. 2	Age Gp. 3
Stay Same Occp.		-0.870	-0.171	4.926	3.780	3.464
Change Occp.	-0.389	-0.969	-0.165	4.897	3.914	3.504
College	0.102	-2.432	-0.059	1.319	-1.033	
Unempl.	-0.925	-1.334	-0.293	2.402	1.686	
Retired	0.099	-0.181	-2.010			2.339
Housewife	0.000	0.000	0.000	0.000	0.000	0.000

Leaving household probabilities

- Tables of probability of household members leaving, conditional on
  - household size
  - age of oldest household member
  - age of individuals leaving the household
- Moving to a logit choice model, with utility of leaving based on attributes of individual and of household

Income

- Determined annually based on regression
  - PUMS data used for regression
  - error term included
  - error term is not resampled, to have stability in income and spatial distribution of income

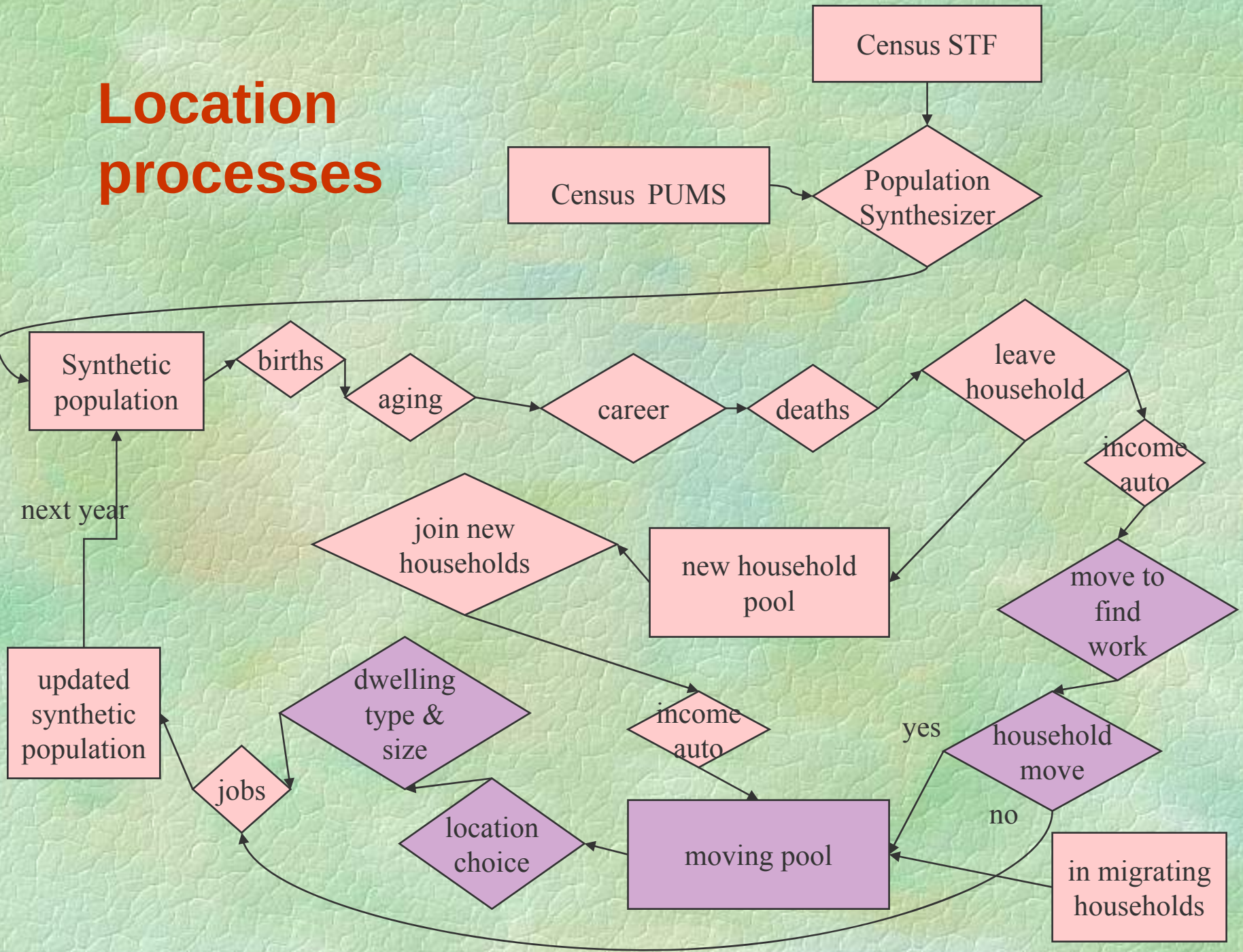
Income regression

Variable	Influence	t value
Intercept	5559.485	21.02
number of 1_ManPro jobs	25757	111.51
number of 1a_Health jobs	24222	65.55
number of 2_PstSec jobs	19712	19.21
number of 3_OthTchr jobs	21228	50.23
number of 4_OthP&T jobs	18392	30.06
number of 5_RetSls jobs	10058	28.77
number of 6_OthR&C jobs	18538	86.95
number of 7_NonOfc jobs	9715.38	66.34
number in military	12741	4.38
number of college students	4555.087	5.68
number of k12 students	1073.673	8.13
no one under 5	-937.5489	-1.42
number of unemployed people	-1828.967	-9.58
number of not working people	-2393.28	-1.1
number of retired people	8115.446	40.27

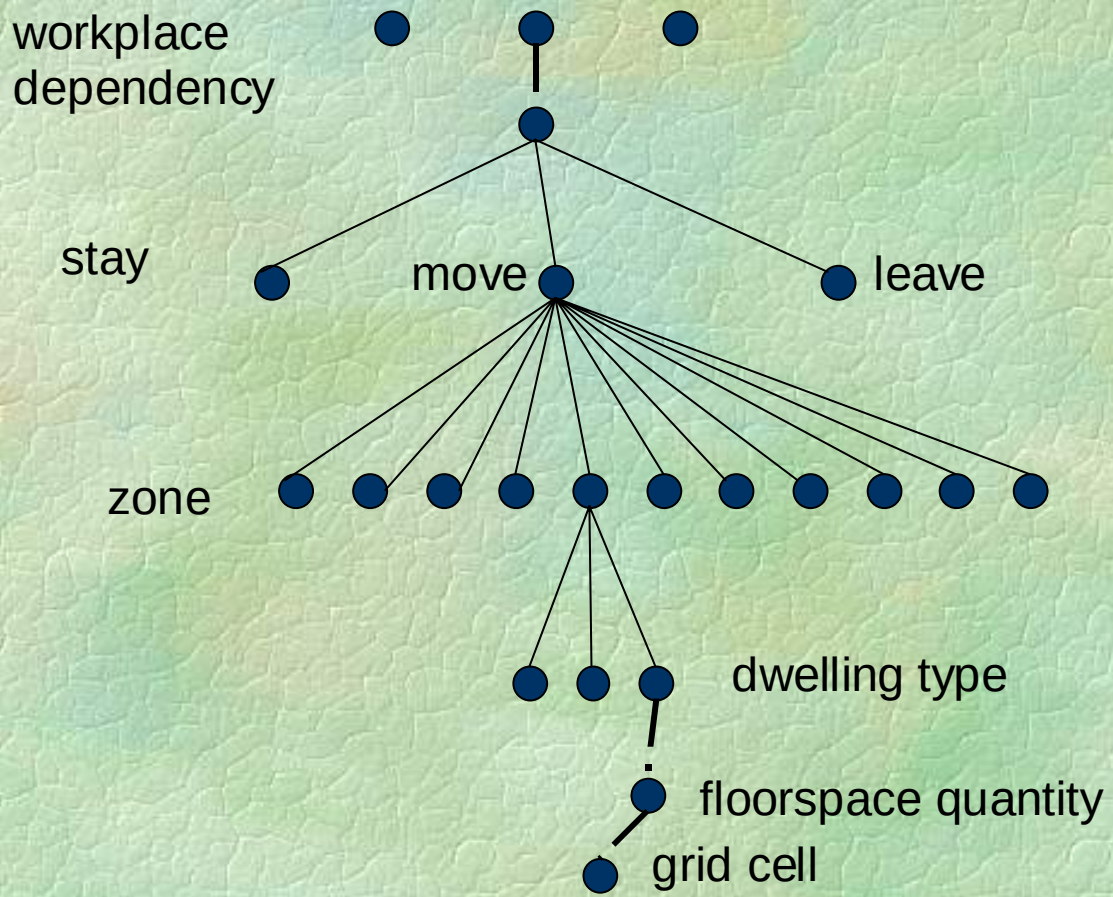
Auto ownership

- Portland Metro auto ownership model

Location processes



Location choice



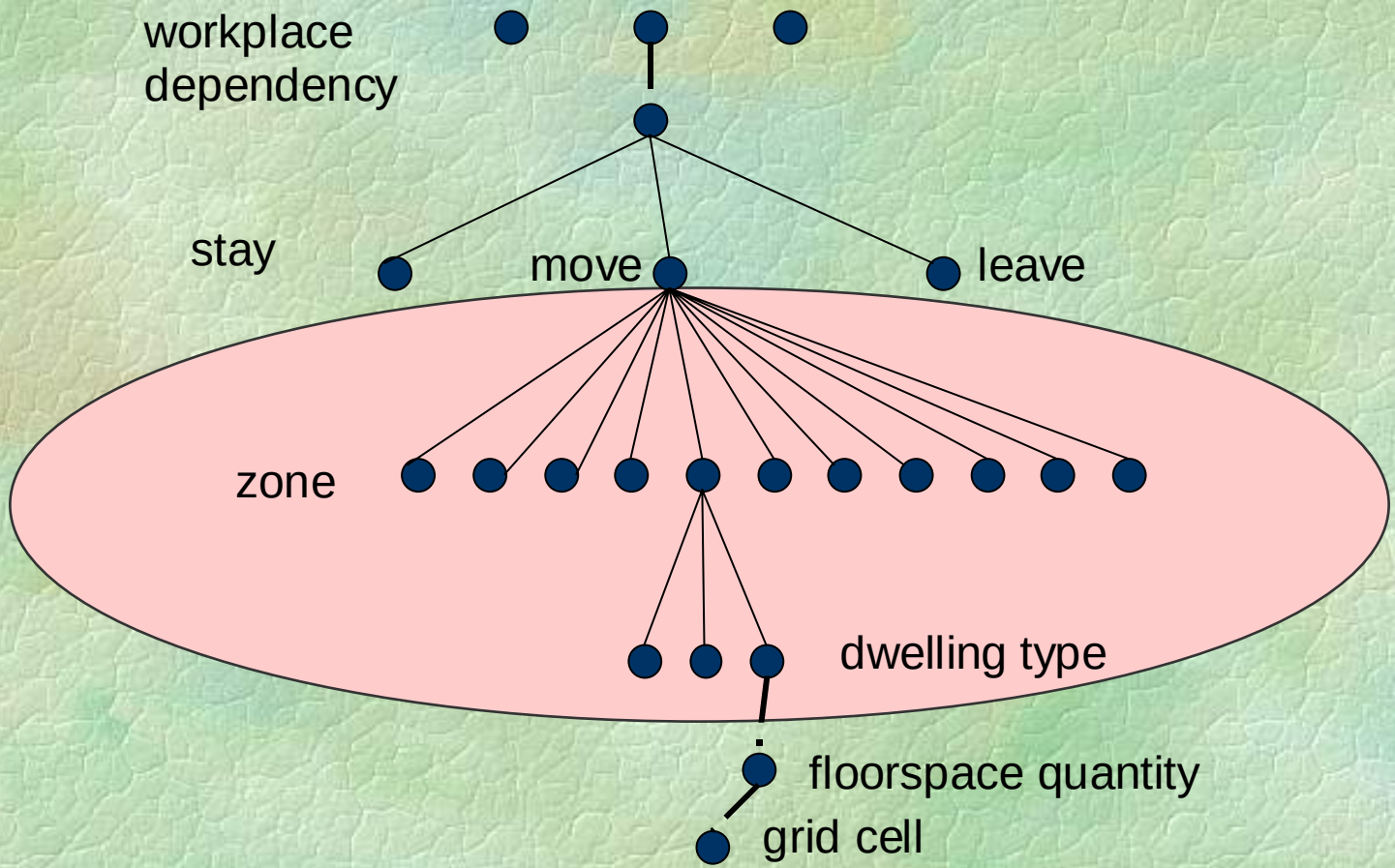
Workplace dependency

- 3 separate possible ways to evaluate home locations
  - conditional on some current workplaces
  - independent of current workplaces
  - conditional on new workplace
- Simple shares model, to be calibrated so that model matches observed aggregate trends
- If 3rd action is chosen, the “new workplace model” is also executed
  - simple logit choice of zone with utility based on number of opportunities and wages established in PI
  - for one household member, based on his/her occupation
- Coefficients based on observed aggregate shares (“calibrated”), not estimated from disaggregate observations

“Move, Stay or Leave”

- Utility function based on workplace/home dependence selection
  - if conditional on workplaces, includes commuting utility to specific jobs
  - if independent on workplaces, includes workplace destination choice logsum to suitable jobs
- Move = log-sum of location choice
  - large computational burden for millions of households
- Leave = fixed utility
- Stay = location choice + constant – size term
- Leave constant, stay constant and dispersion parameter based on observed aggregate shares

Location choice



Utility of zones

- Attributes used
  - residential floorspace price
  - workplaces
    - travel utility (log sum of mode choice) to actual workplace for 3 workers
    - accessibility (log sum of destination choice) to all employment
    - accessibility to labor category employment for 3 workers
    - distance to workplace
    - *all include wage with calibrated parameter*
  - accessibility to shopping, education, recreation...
  - *log sum from space type choice*
- Accessibilities highly correlated
  - use expenditure and trip rate data to weight different accessibility measures
- Sampling issues
  - in a statewide model, considering a subset of the alternatives, it can appear that people always choose the home location that is closest to work.

Dwelling type and square feet

- Based on Census PUMS regression
- Number of rooms as a function of dwelling type, income, number of workers, presence of retired people.
- American Housing Survey used to convert number of rooms to square feet of housing
  - 283.09 sq ft per room

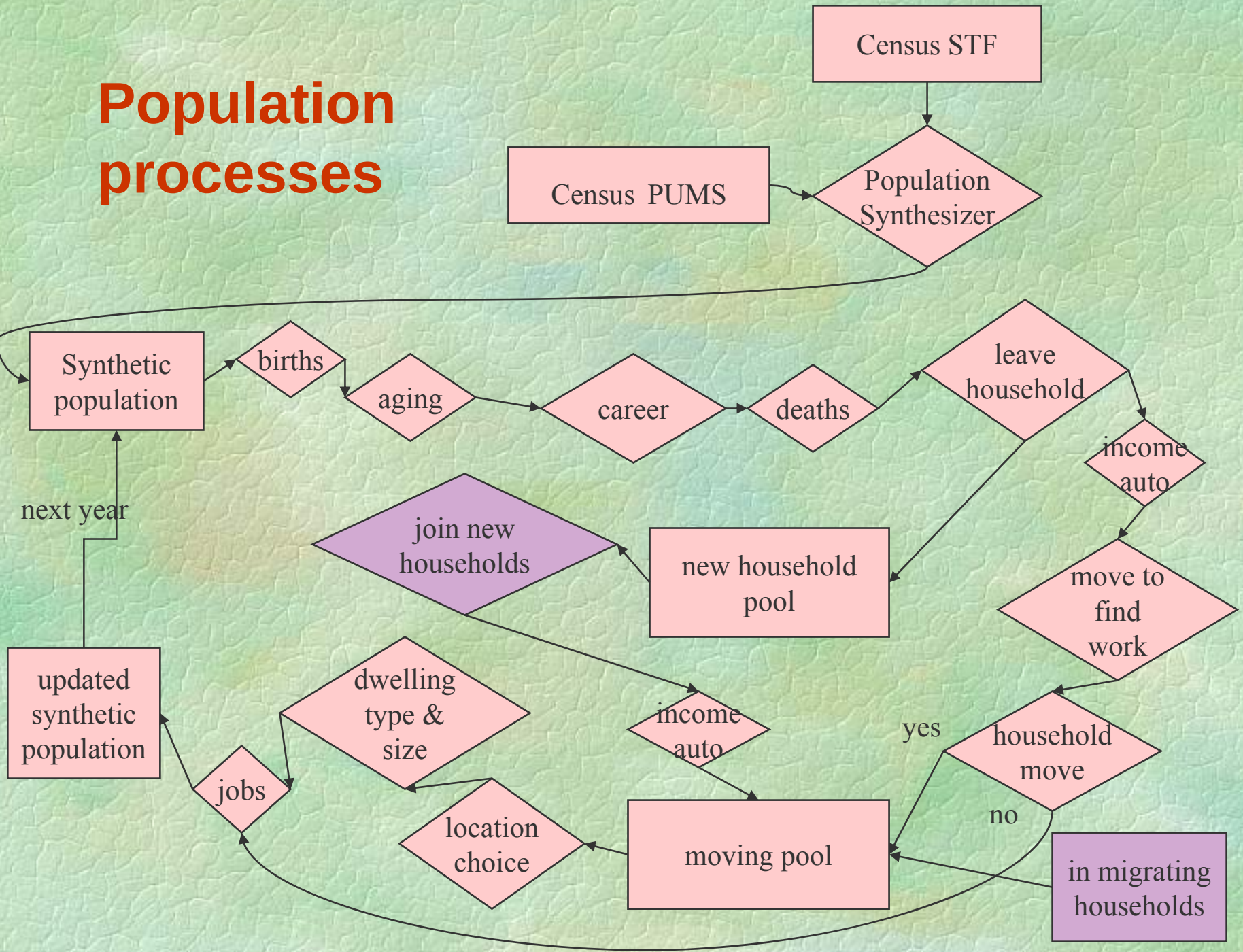
Number of rooms needed

Variable	Estimate	t Values*
Intercept	3.91819	159.58
S. Fam Detached	1.08305	55.95
Multifamily housing	-0.80047	-35.78
2 workers	0.54122	-34.58
3 workers	0.60794	-14.01
4+workers	0.91644	-12.47
each retired person	0.15886	10.17
$1000 income	0.01656	68.24
* t Values shown here are approximate – they are based on a slightly transformed estimation		

Grid cell allocation

- Random selection of grid cell with appropriate development type in the TAZ
  - potential to use grid cell attributes to do a finer spatial allocation

Population processes



Category shares for new households

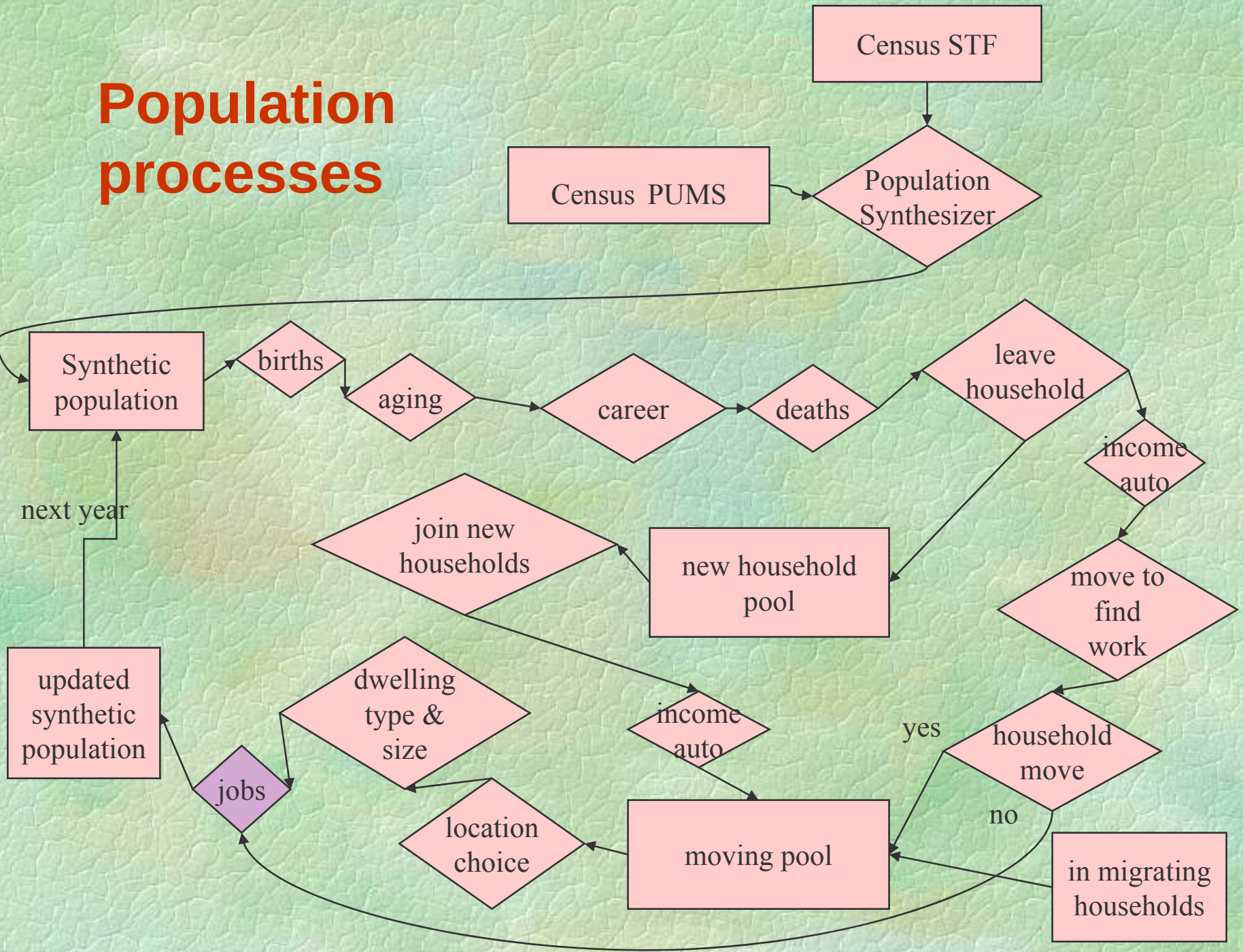
- Applies to in migration, and to those people who left existing households
- Categories with weights, applied directly for in migrating households
- New Household Pool Individuals processed one at a time
  - household category selected based on weights
  - see if individual fits
  - try to fill rest of household from within the pool
- Weights and categories based on overall demographic calibration

Current categories

- Single individual
- Young married couple
- Married couple with 1 young child
- Married couple with 2 young children
- Older married couple with older child
- 2, 3, 4, 5 and 6 person households

- Weights control overall demographics
  - need longer term demographic forecasts!
- Note strategy:
  - leaving probabilites are disaggregate, but fixed based on history in the national and biased PSID data set
  - joining probabilities aggregate, but adjustable

Population processes



Workplace Location Choice

- Connect the right workers with the right jobs
- Employment in size term aggregated into two categories: Primary, Residual
  - Primary employment
    - The employment type most likely to attract the industry/occupation of the worker
  - Residual employment
    - All other employment types

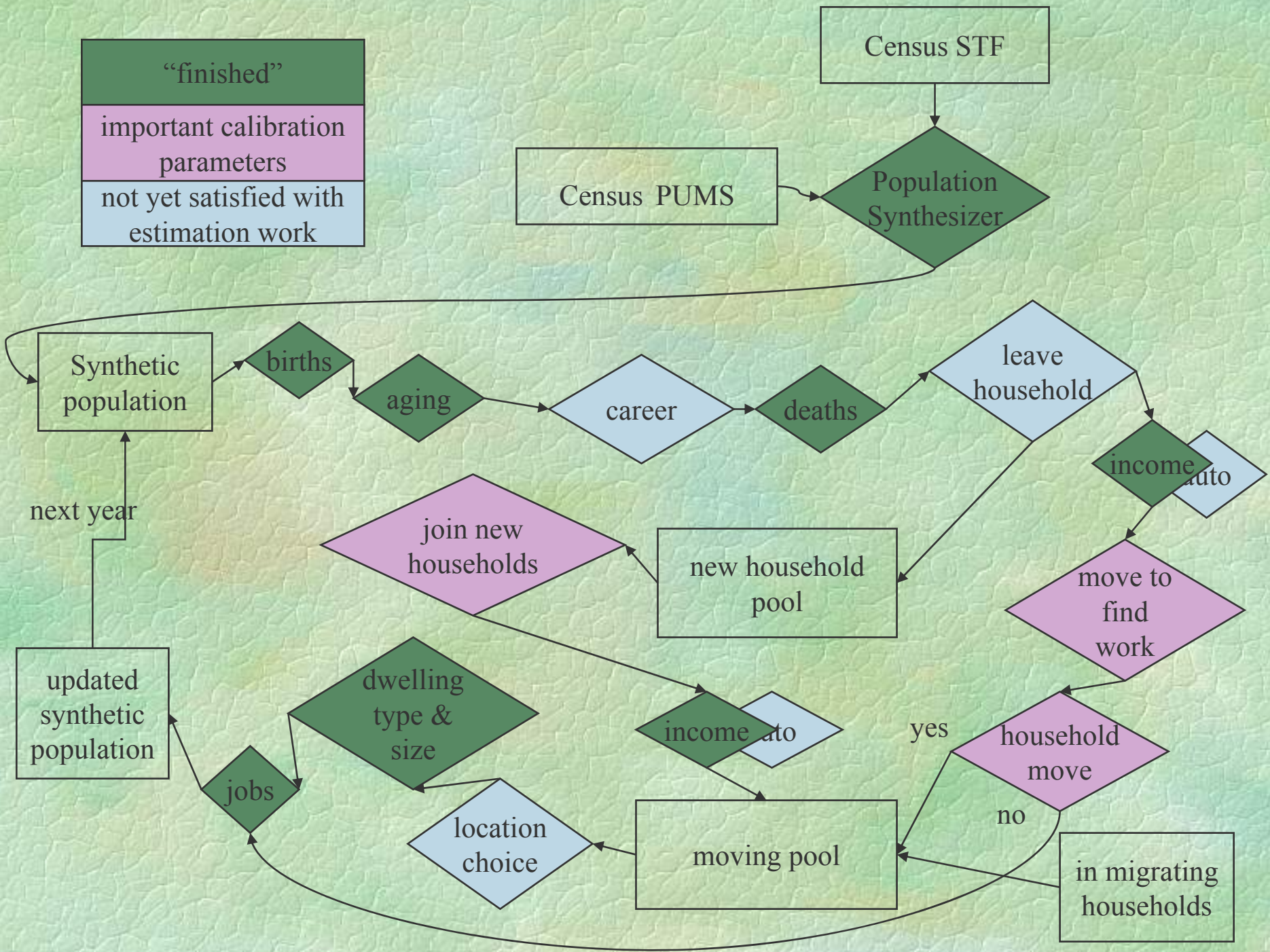
Segment	Occupation & Industry	Primary Employment Type	Residual Employment Type
1	1_ManPro 4_OthP&T	Office	Non-Office
2	5_RetSls	Retail	Non-Retail
3	7_NonOfc	Industrial	Non-Industrial
4	1a_Health 2_PstSec 3_OthTchr	Other Employment	

Workplace Destination Choice Estimation Results

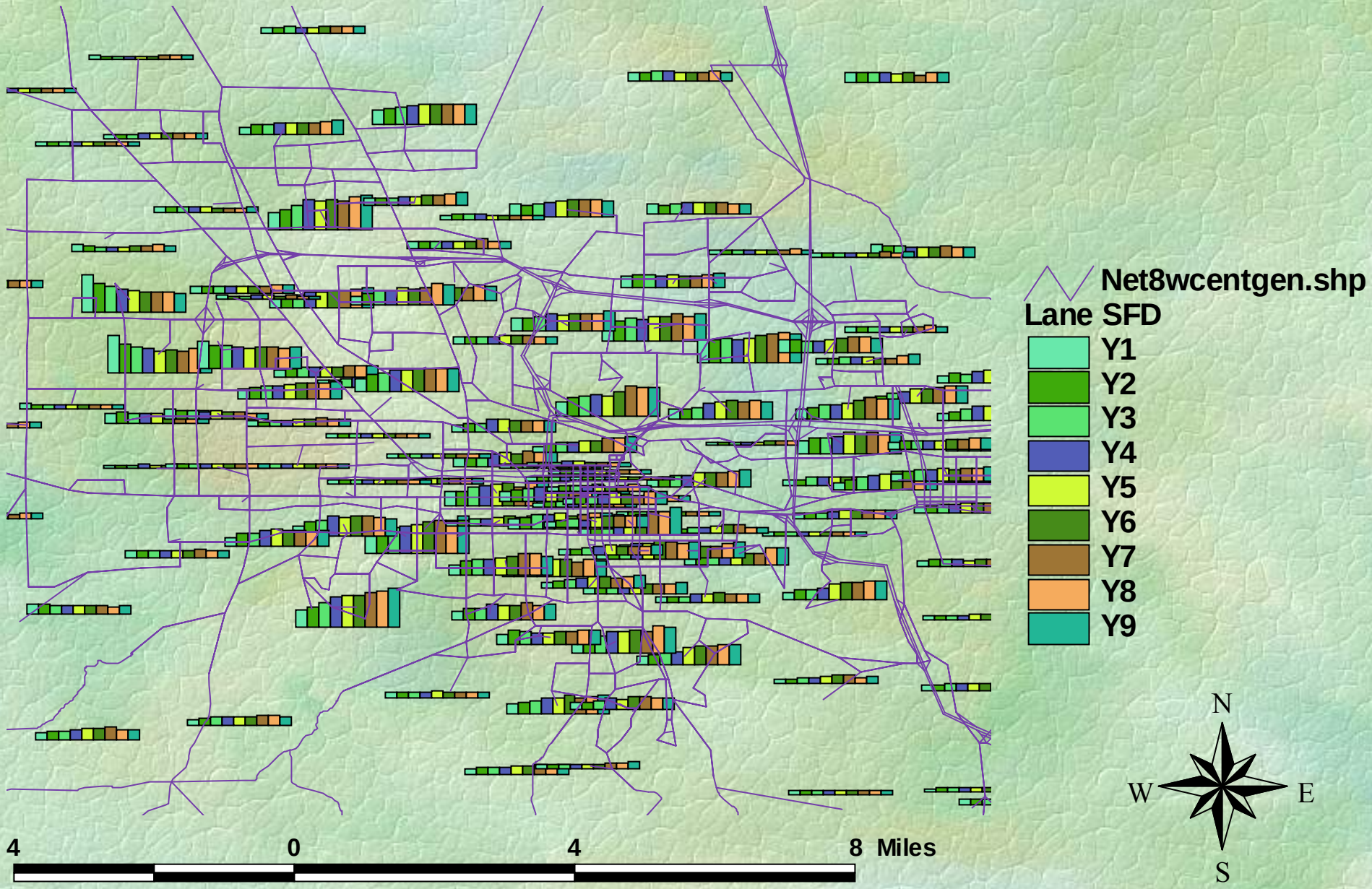
- Sampling of alternatives required (chosen + 40)
  - Based on distance from tour origin TAZ
- Significant mode choice logsums
- Size terms reasonable, but difficult to estimate reasonable parameters for multiple terms
- May require distance-based adjustment in application

Setup

- 10 zone test system for long term dynamics
  - linear city, one type of residential floorspace, poorly synthesized households, commercial floorspace included
  - 4000 households, 100 years, household and grid cell demographics
- Lane County – more realistic data, with 10 year dynamics
  - synthetic population, several types of residential floorspace, every household needs a 1200sq ft dwelling
  - 111000 households, residential floorspace only
- All of Oregon – yet more realistic data, 2 year runs
  - synthetic population, income and floorspace requirement models, actual labor categories.


"finished"
important calibration parameters
not yet satisfied with estimation work
Census STF
Census PUMS
Population Synthesizer
Synthetic population
births
aging
career
deaths
leave household
income
auto
next year
join new households
new household pool
move to find work
updated synthetic population
dwelling type & size
income
ato
yes
household move
jobs
location choice
moving pool
no
in migrating households

Eugene/Springfield population trends



Household transitions example

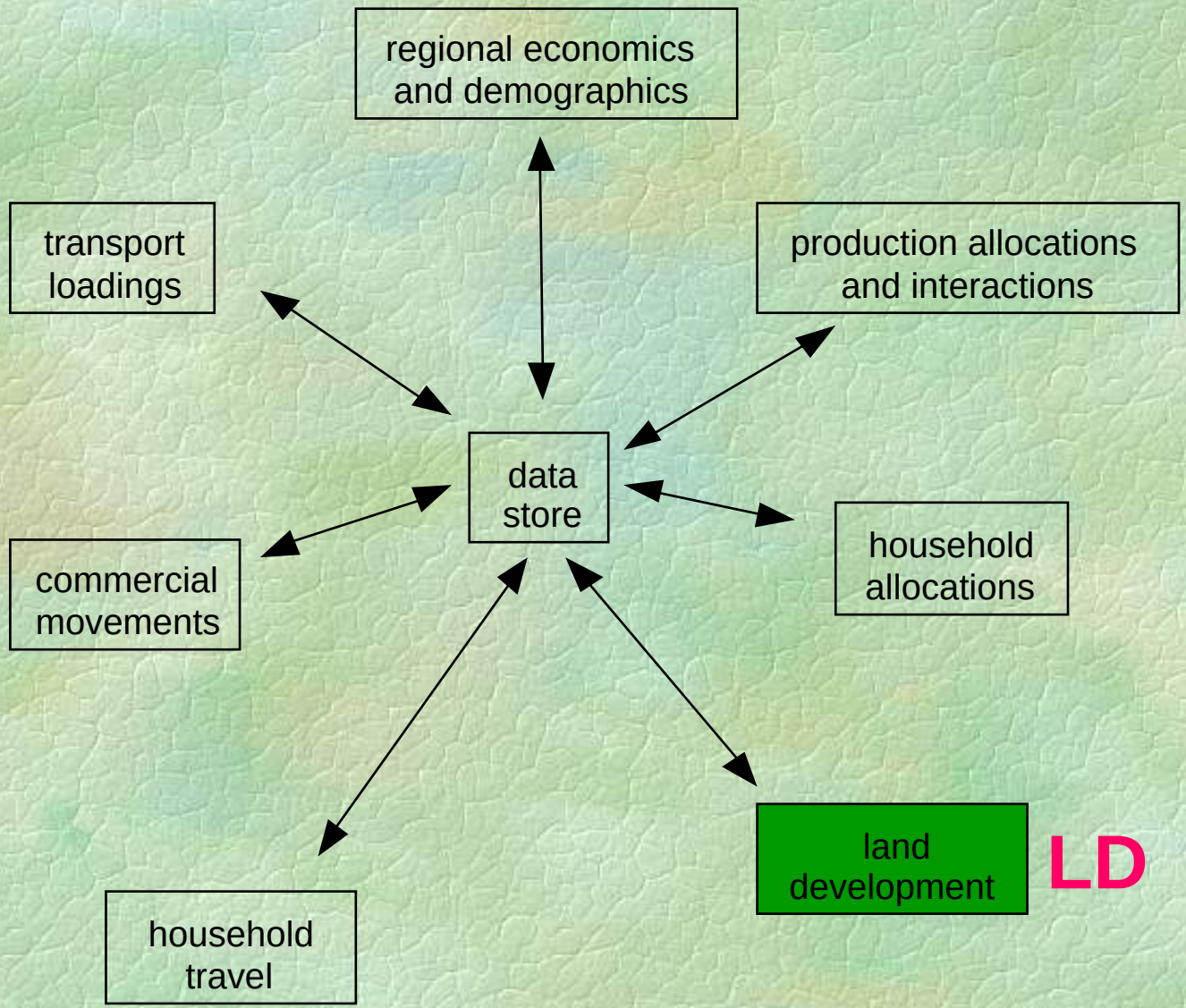
1 auto	26560/yr	TAZ 1550				
	64yo	Female	no job	no occupation	high school	not in school
	30yo	Male	empl ft	6_OthR&C	11th grade	not in school
1 auto	26560/yr	TAZ 181				
	66yo	Female	no job	no occupation	high school	not in school
	32yo	Male	empl ft	6_OthR&C	11th grade	not in school

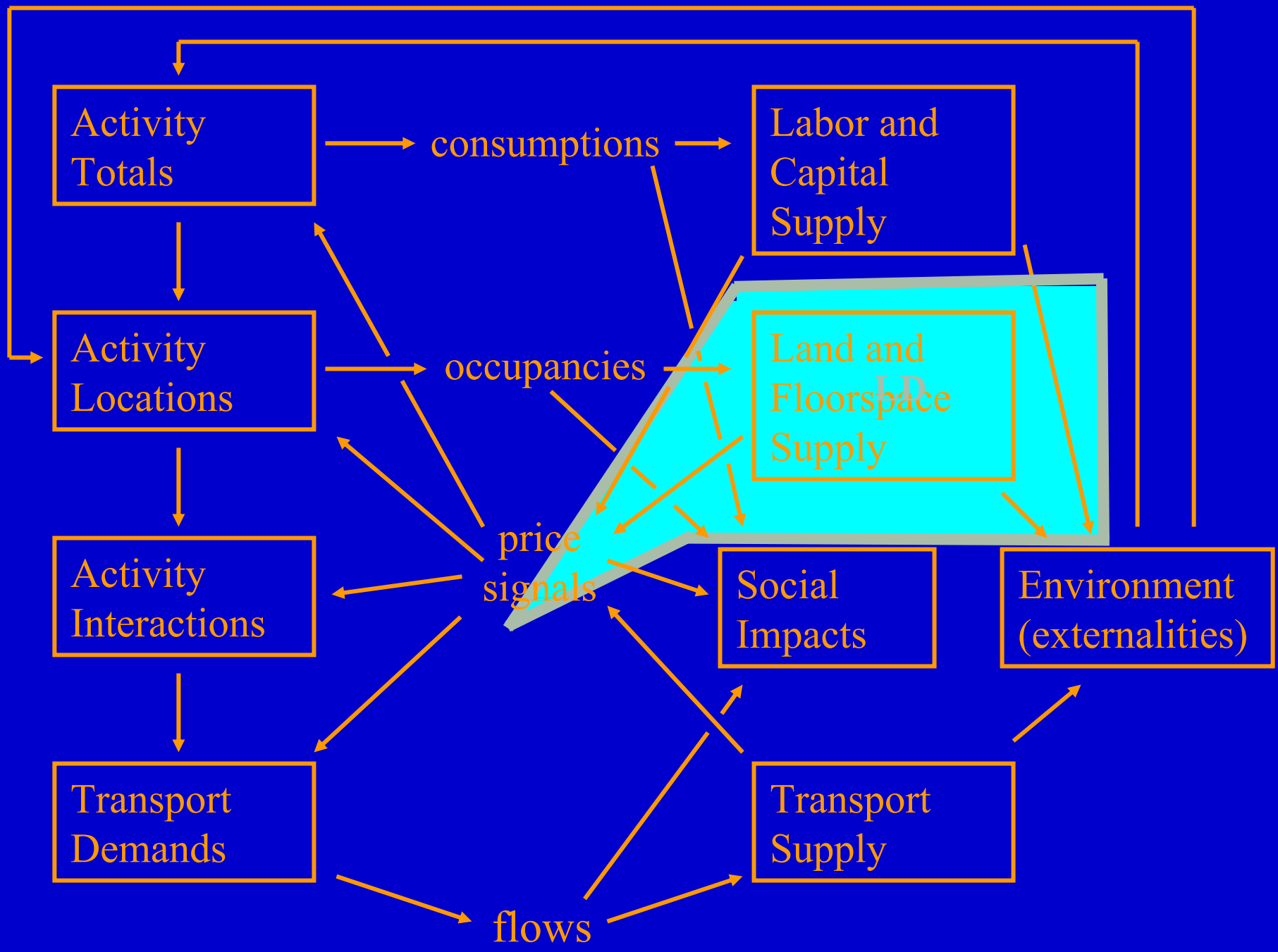
Household transitions example

5 autos	31805/yr	TAZ 3130				
	42yo	Male	empl ft	7_NonOfc	bachlrs	not in school
	43yo	Female	no job	no occupation	some college	not in school
	19yo	Female	empl ft	7_NonOfc	HighSch	not in school
	16yo	Female	empl pt	7_NonOfc	10th grd	in school
	5yo	Female	no job	no occupation	no grds	in school
4 autos	21016/yr	TAZ 143				
	44yo	Male	empl ft	7_NonOfc	bachlrs	not in school
	46yo	Female	no job	no occupation	some college	not in school
	18yo	Female	empl pt	7_NonOfc	HighSch	in school
	7yo	Female	no job	no occupation	2nd grd	in school

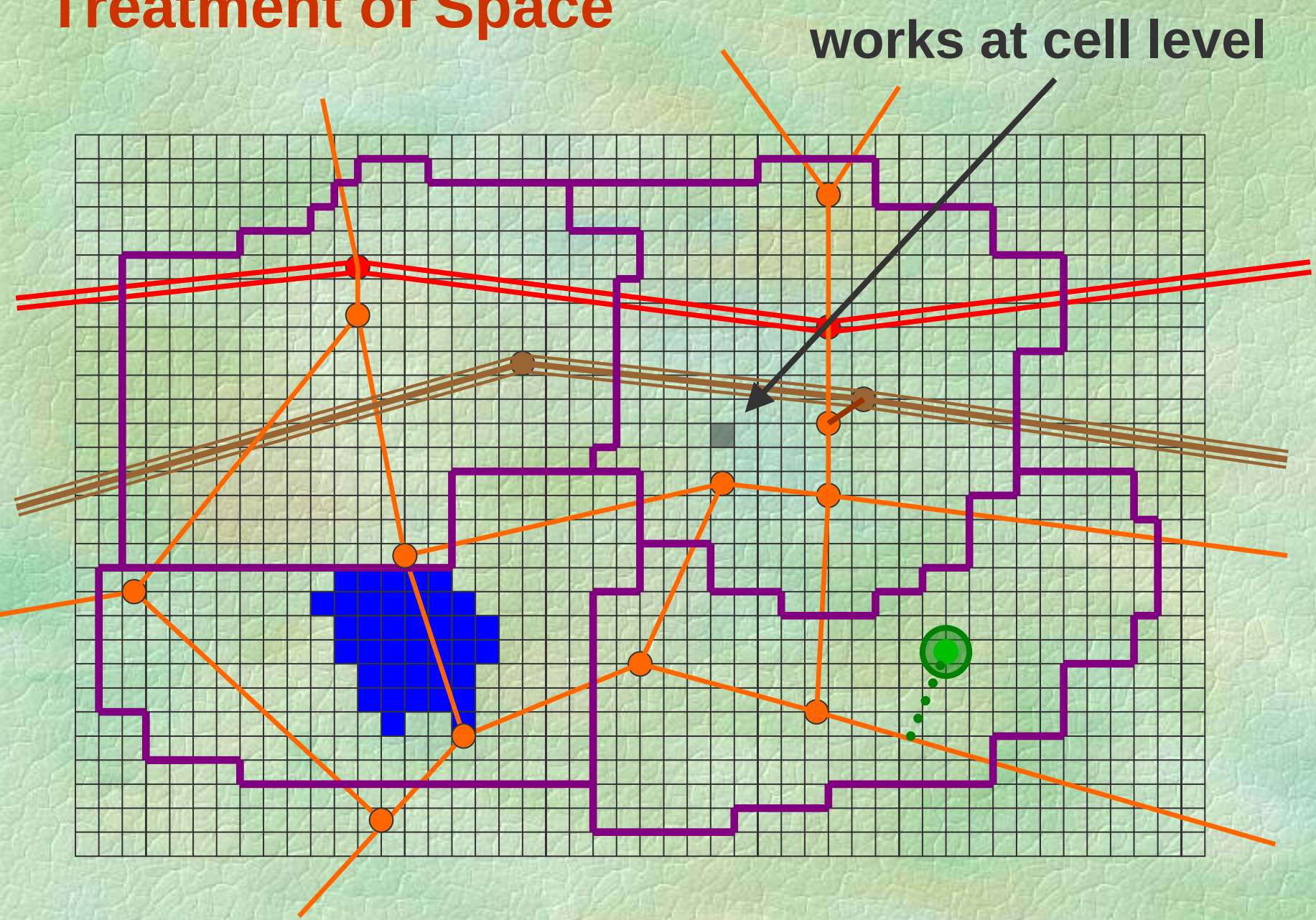
Model Framework; Treatment of Behaviour

Land Development




Activity Totals
consumptions
Labor and Capital Supply
Activity Locations
occupancies
Land and Floorspace Supply
price signals
Activity Interactions
Social Impacts
Environment (externalities)
Transport Demands
Transport Supply
flows

Treatment of Space



Overview

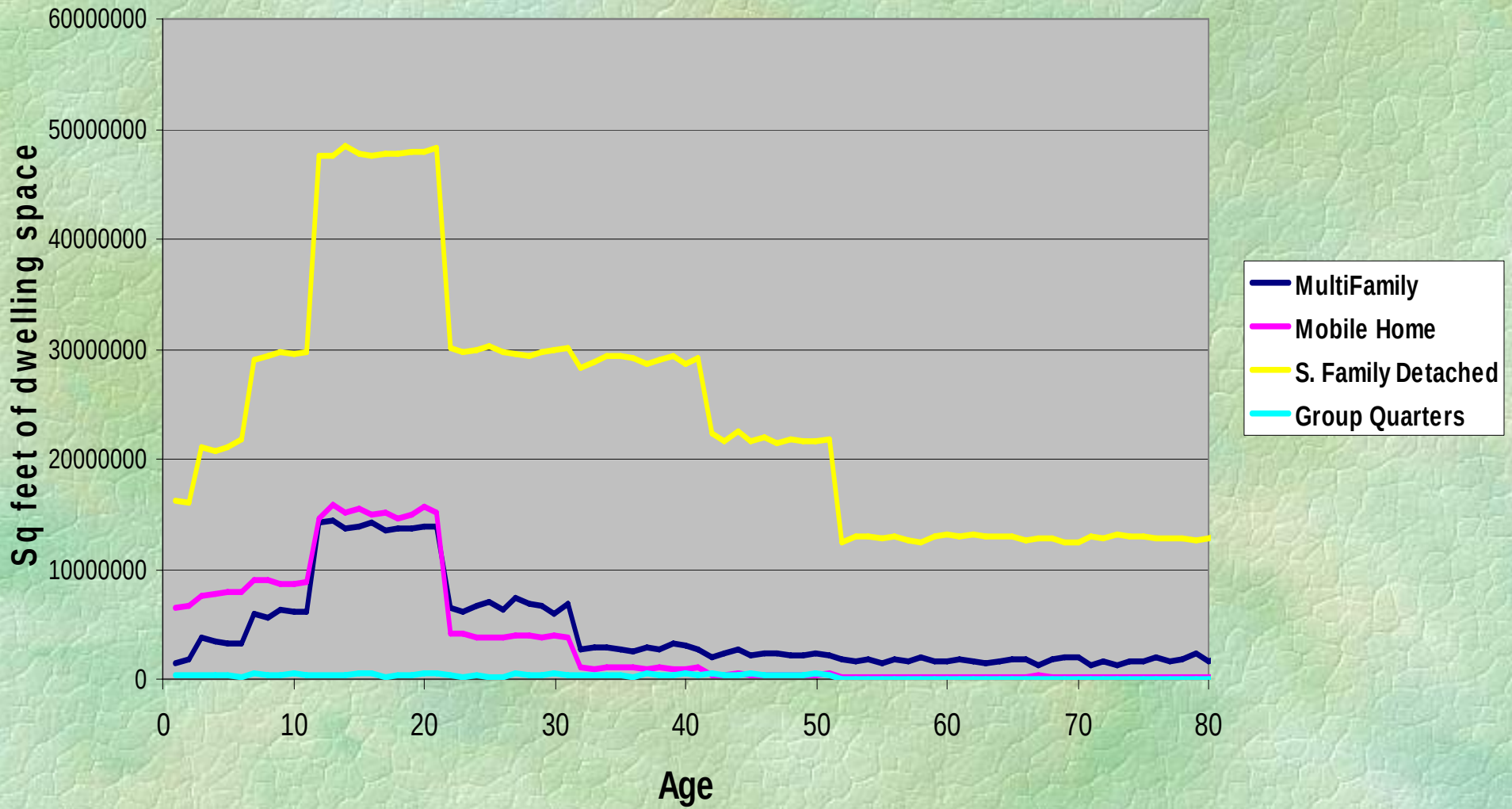
- Disaggregate representation of space based on single values for some attributes:
  - floorspace age
  - density
  - zoning
  - development type
- Transitions microsimulated
  - emphasis on changes in development type and density (quantity)
- Responsive to the zonal prices and vacancy rates established in PI and HA

Base year inputs

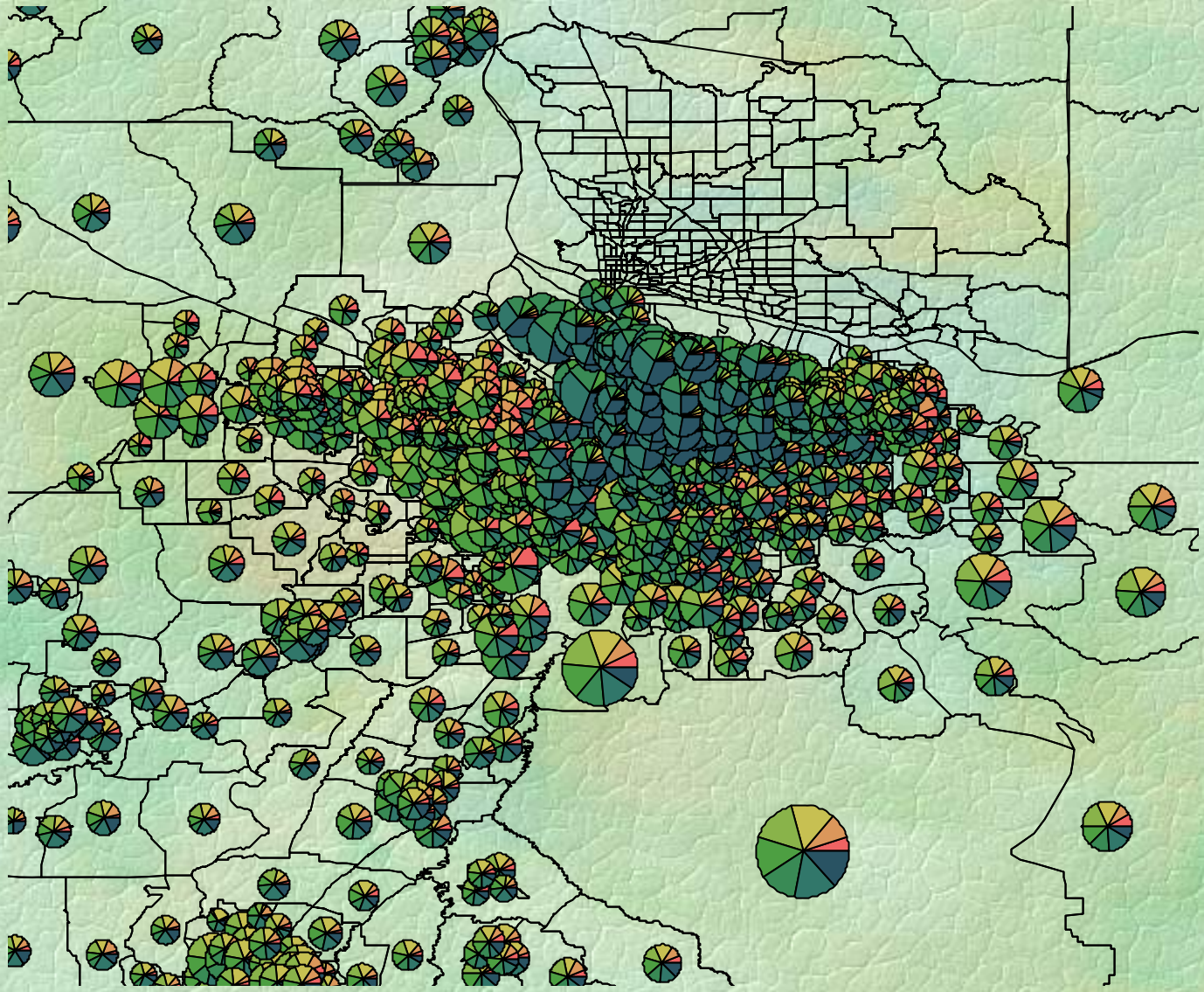
- PI employment inputs imply industrial/commercial space quantities
- HA synthetic households imply residential space quantities
- PNWERC zoned land data and floorspace quantities combined with
  - structure age data from census
  - floorspace ↔ zoning compatibility
  - less developable space
- PNWERC actual grid cells planned
  - provide spatial detail beyond zone system

Age distribution of residential floorspace

471,000 residentially developed grid cells



Portland SFD Housing by Age

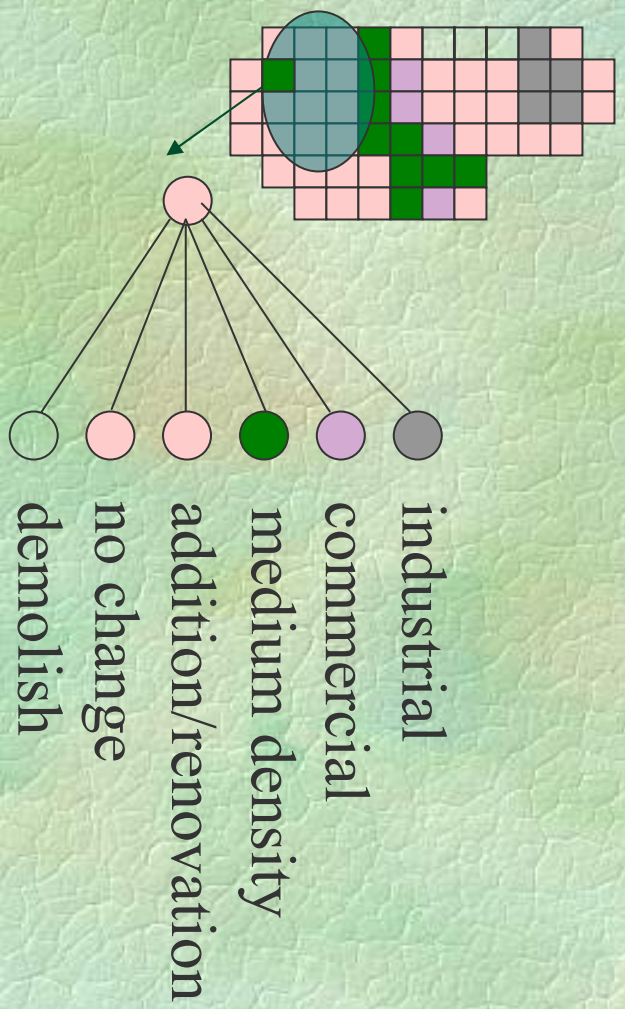


SFD Housing Stock by Age
- 0-5
- 6-10
- 11-15
- 16-20
- 21-30
- 31-40
- 41-60
- 61+

N
W
E
S

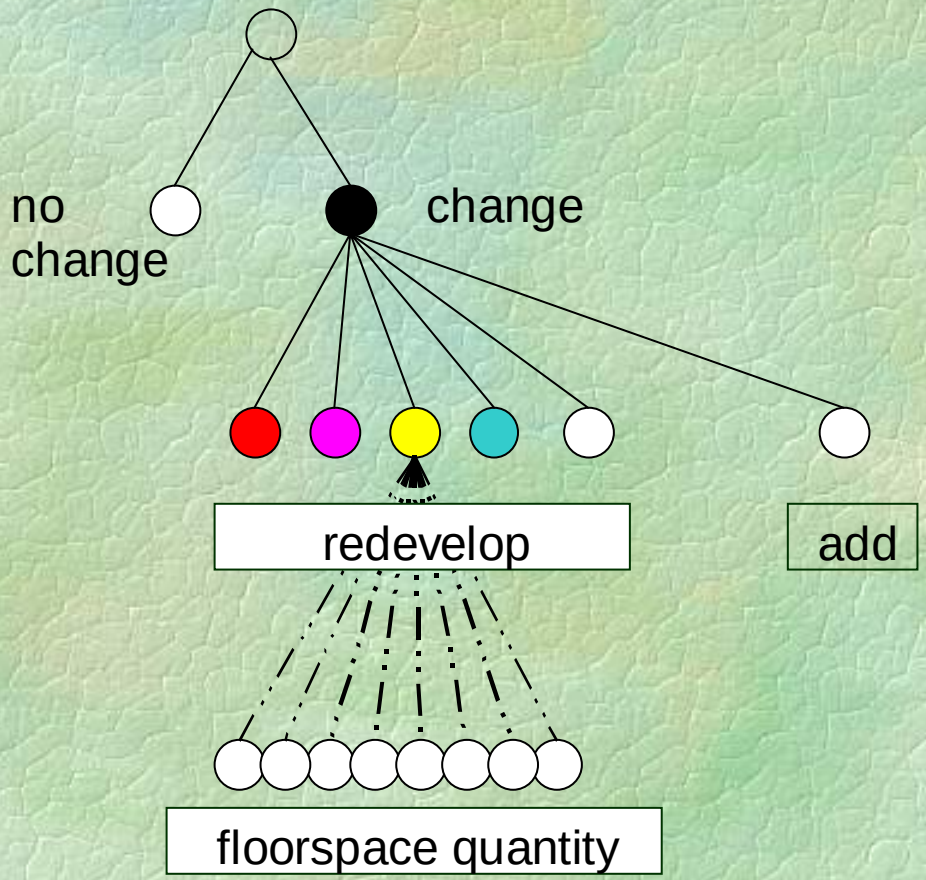
20 0 20 40 Miles

Land microsimulation



- commercial
- residential low
- residential medium
- industrial
- vacant

Structure



Utility function

-demolition costs + potential profit – zonal vacancy risk

- $U_{ij} = K_{demo} + a_2 * densb_j * (rent_{jn} - devcost_i) + K_z * zoneVacancy_j + urbanSimPara$
- $U_{i.} = K_{i.} + a * dens * rent_{in} + K_a * age + K_c * cellVacancy + K_z * zoneVacancy$
- $U_{i+} = a * dens * rent_{in} + a_2 * (maxdens_j - dens) * (rent_{in} - devcost_i) + K_z * zoneVacancy$

constant + revenue, less effect of age and effect of vacancy in cell and vacancy rate in zone

current revenue, plus potential profit – zonal vacancy risk

Draws on Lane County Urbansim development model

- Some coefficients used directly
  - effect of flood plane, slope, etc
  - inertia variables (converted to zonal)
- Others established to reproduce transition rates in Lane County, based on Urbansim application in Lane County
- Others not used because of an effort to separate demand side and supply side

Table of coefficients used

Kc	cellVacancyUtilityCoefficient	-1
Kz	zoneVacancyUtilityCoefficient	-1
a	rentPerAcreCoefficientOnCurrentDevelopment	0.002
Ka	ageCoefficientToKeep	-0.2
Ki.	keepConstant	12
Kdemo	demolitionNotAddConstant	-5
a2	profitPerAcreCoefficientOnNewDevelopment	0.002

Change in development type

Initial development	Next development	Development instances	No development instances
Group Q's	Group Q's		808
Multifamily	Group Q's	76	
Multifamily	Multifamily	21	30880
Multifamily	S. Fam Det.	21	
Mob Home	Mob Home	4	31166
Mob Home	S. Fam Det.	3	
Mob Home	Vacant	15	
S. Fam Det.	S. Fam Det.	478	401377
S. Fam Det.	Vacant	1374	

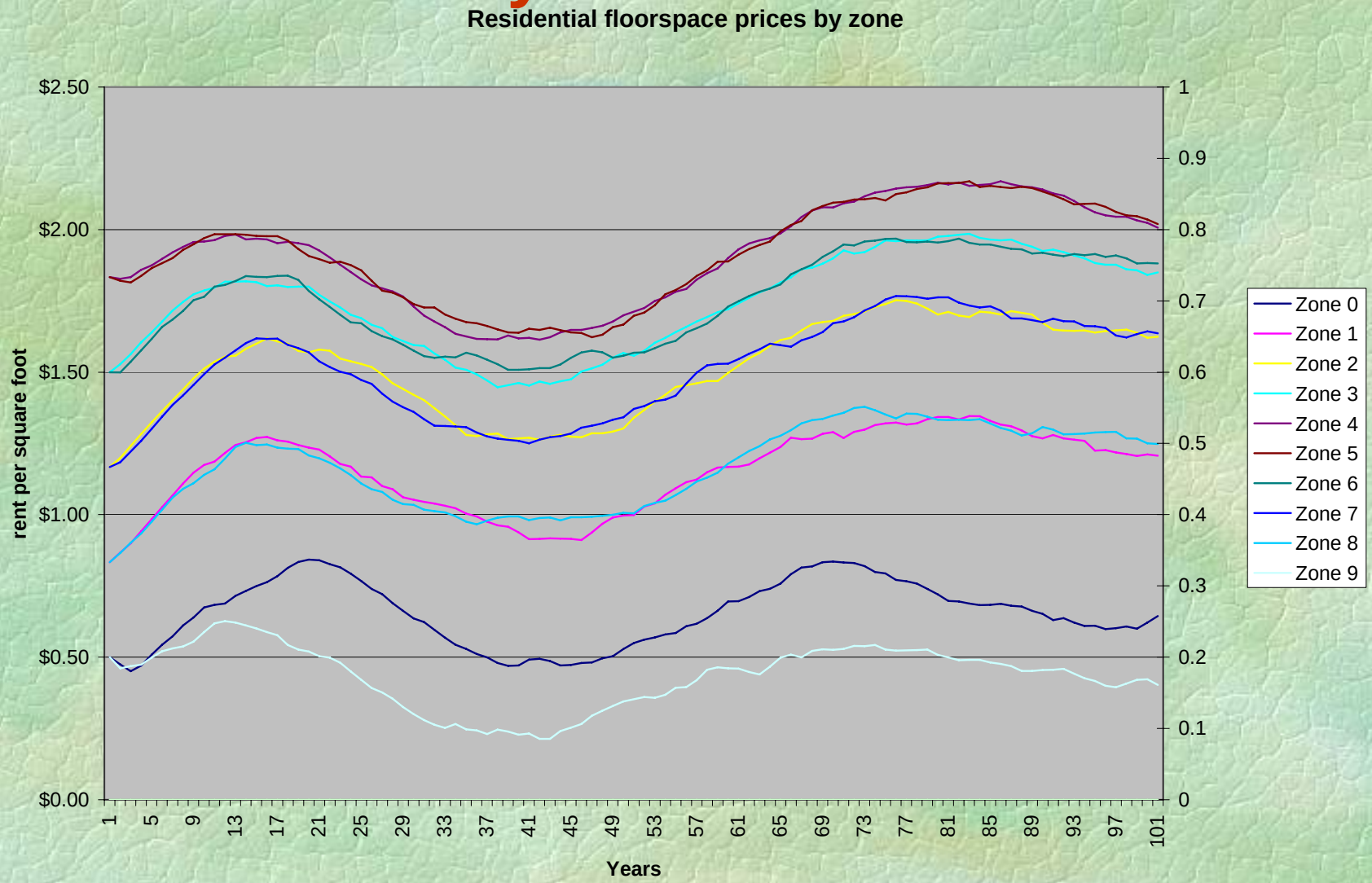
Redevelopment characteristics and notes

- Average age of redeveloped properties is 65 years
- In statewide geography, inventory of future developable land (from PNWERC data and from general plans) is important

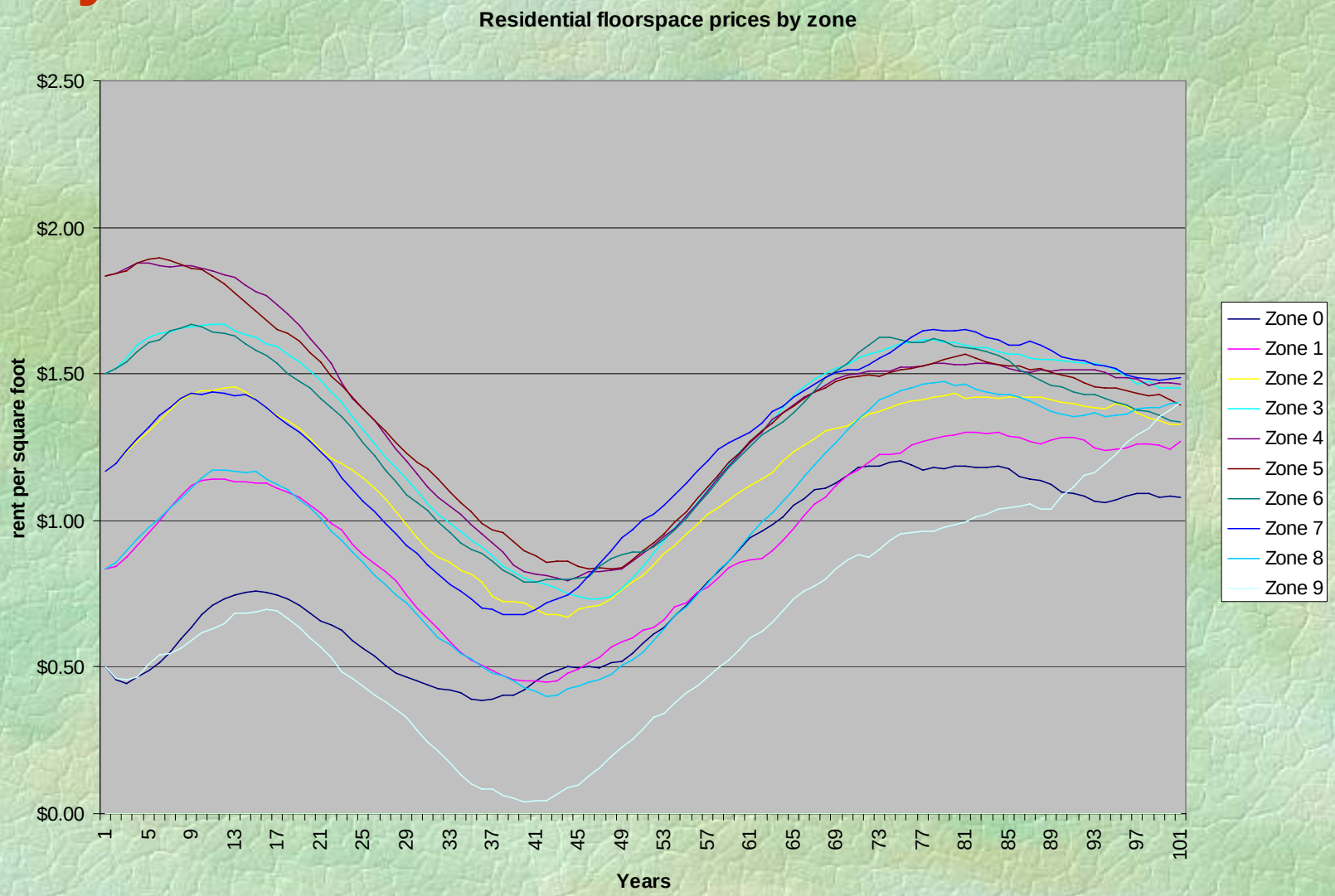
Features

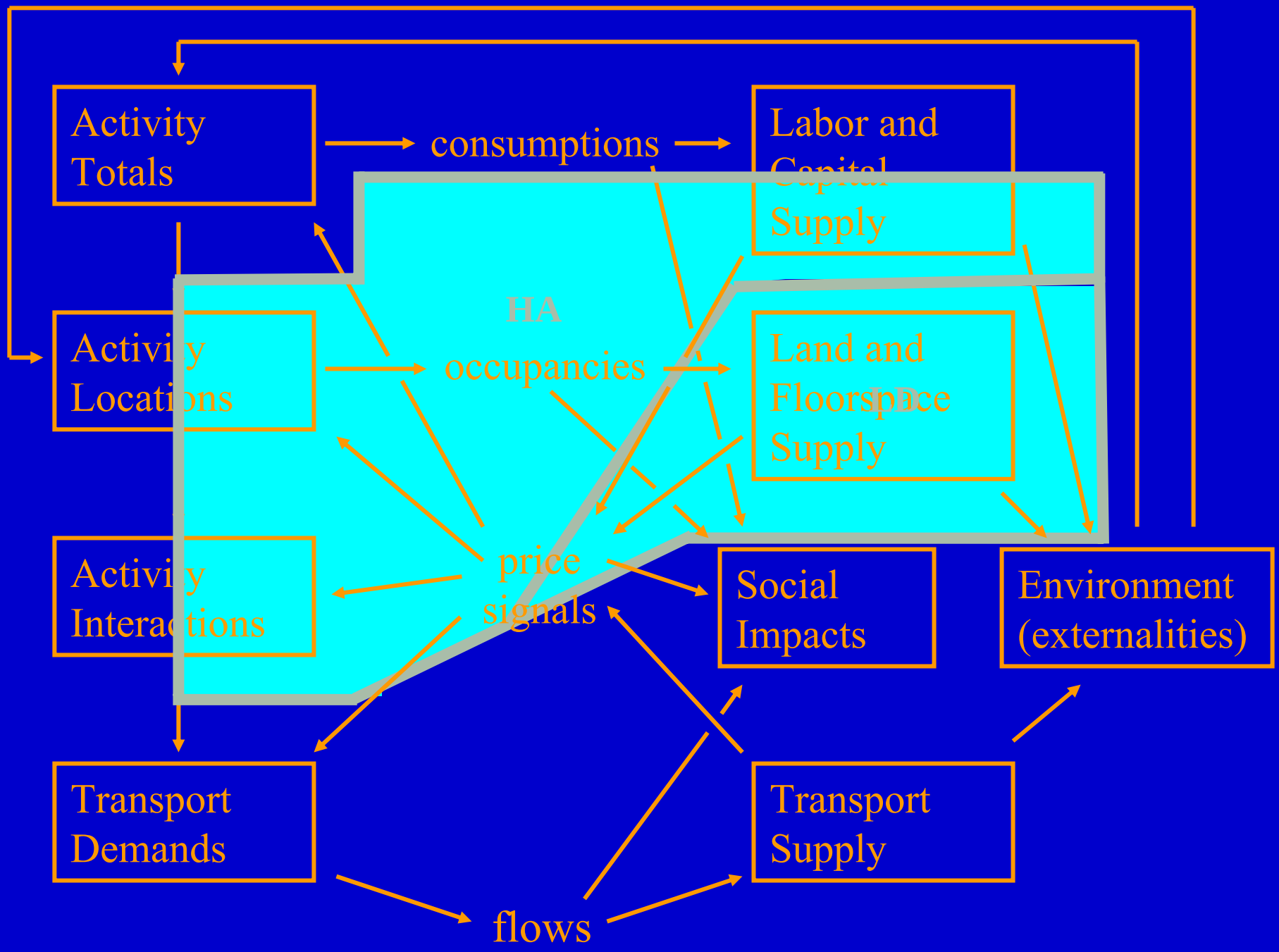
- Behavioral at the grid cell level
- Diversity in land inventory
- Effect of age of building
- Profit maximization per unit of land leads to higher density in higher rent areas

HA Simulation – long term dynamics with 10 zone system



HA and LD Simulation – 10 zone system




Activity Totals
consumptions
Labor and Capital Supply
HA
Activity Locations
occupancies
Land and Floorspace Supply
Activity Interactions
price signals
Social Impacts
Environment (externalities)
Transport Demands
Transport Supply
flows

Three market representations

- Described in Abraham & Hunt 2002
- 1. Quasi-dynamic
- 2. Spatial-dynamic price update
- 3. Full transaction simulation

Labor markets:
Quasi-dynamic markets

- space aggregated into zones
- price at each place established to equilibriate short term demand and supply (in PI)
- “short term” long enough so that there are a range of transactions at each place (1yr)
- long term aspects of demand and supply updated based on prices established in short term equilibrium (participation/home location HA, economic transitions in ED)

Residential floorspace markets: Spatial-dynamic price-update

- Agents (households in HA, developers in LD) assume fixed prices
- Space is aggregated into zones, one price for each housing type at each zone
- Demand modules (HA) remove product, supply modules (LD) provide product
- Every 6 months (or every month), prices updated based on “excess demand” (vacancy)
- Separates demand and supply completely
  - no module needs to represent both demand and supply

Conclusions

- Disaggregate microsimulations!
  - represent the diversity of population and of built form
  - not achievable in an aggregate/categorization model
- Lots of little modules
  - many approximately independent in any year (death, birth, career transitions…)
  - others tightly linked (job dependency, zone, space type)

Conclusions

- Testing scenarios important
  - would have been swamped trying to debug model software and model behavior at the statewide level
- Grid cell data doesn’t have to be complete
  - framework works with “synthetic grid cells” just as well as it works with “synthetic households”
  - can use distributions to sample any grid cell variable from a higher level of geography
  - would be nice not to have to sample ‘x-y’ coordinates though
  - full potential of model requires geocoded grid cells

Conclusions

- Long term bootstrapping and calibration issues are important
  - synthetic population generator not as rich the model
  - could use the model to “build up” the population over time, but then wouldn’t match in base case
  - expect some “start up transition” period
- Dozens of remaining “time series calibration” parameters
  - more data on household and grid cell transitions could reduce these
  - cross sectional data somewhat limited in a path-dependent modeling excercise