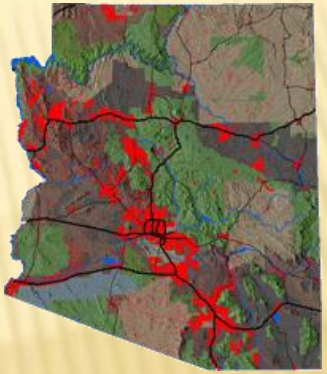


AZ-SMART

Arizona's Socioeconomic Modeling, Analysis, & Reporting Toolbox

AZ-SMART POPTAC WORKSHOP
FEBRUARY 28, 2012

Copyright © 2012

AZ-SMART – OVERVIEW

- AZ-SMART: What and why?
- Model System Overview
- Input and Base Year Data
- Models: Simulation Sequence & Description
- 2012 Official Projections

Copyright © 2012

AZ-SMART……..WHAT IS IT?

- A framework for developing socio-economic projections
- Tools for data development and maintenance
  - Land use, general plan, inventories, etc.
  - Population and employment synthesis
- Tools for data analysis
- Tools for configuring, estimating, calibrating, and running simulation models
- Tools to assist in scenario building

Copyright © 2012

AZ-SMART……WHY DO IT?

- We are required to produce population projections
  - Executive Order 2011-04
- Travel model demands
  - Travel models depend on socio-economic data
  - Socio-economic models are improved by travel model feedbacks
- We are getting asked questions that previous models were less equipped to answer, e.g.
  - Corridor studies
  - Activity-based travel models
- Policy analysis, what happens if:
  - Employment sector X is predicted to grow or decline substantially?
  - We invest or don't invest in transportation project X?
  - Several cities change land use policies to encourage density?
  - Etc….

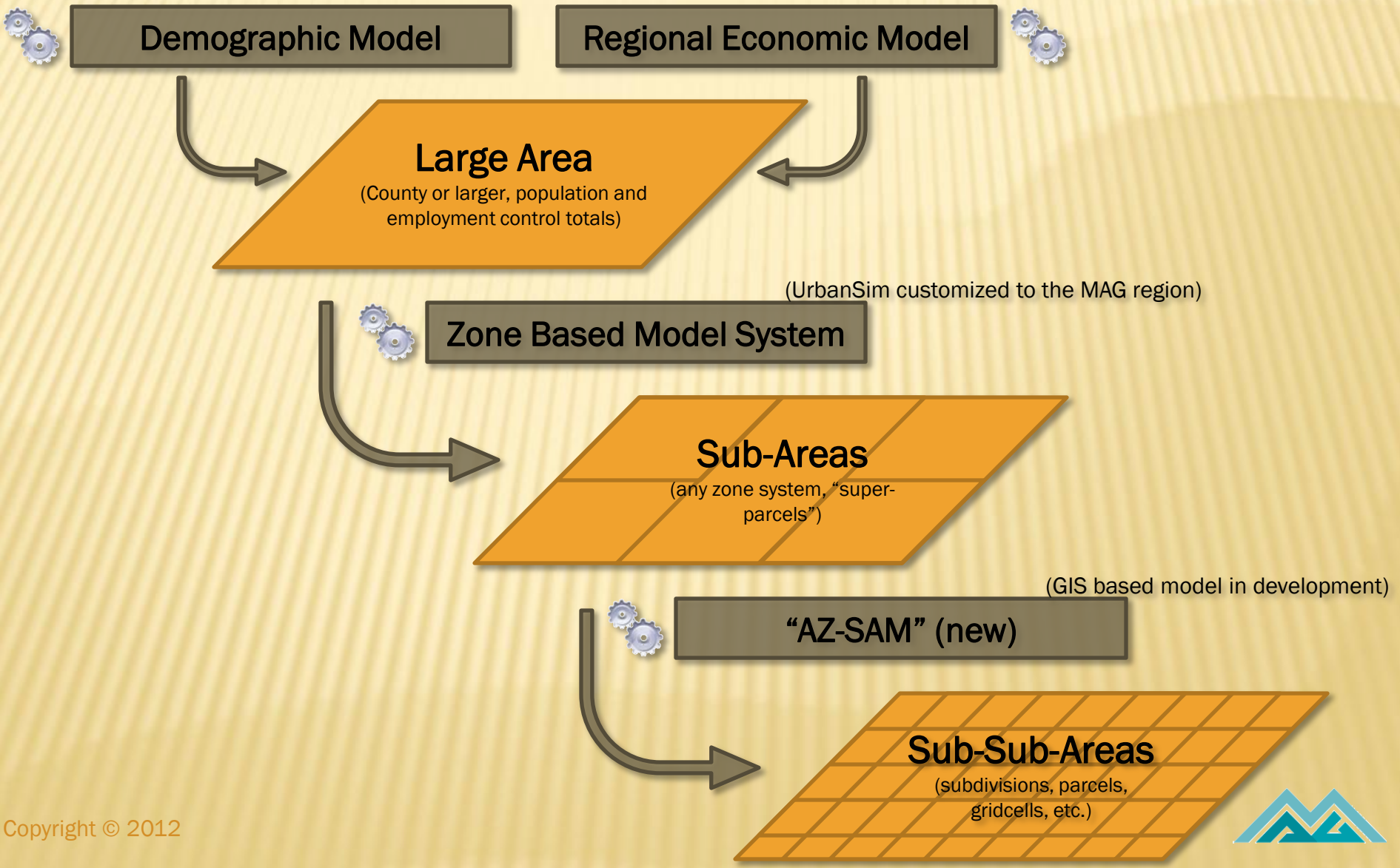
Copyright © 2012

AZ-SMART

Model System Overview

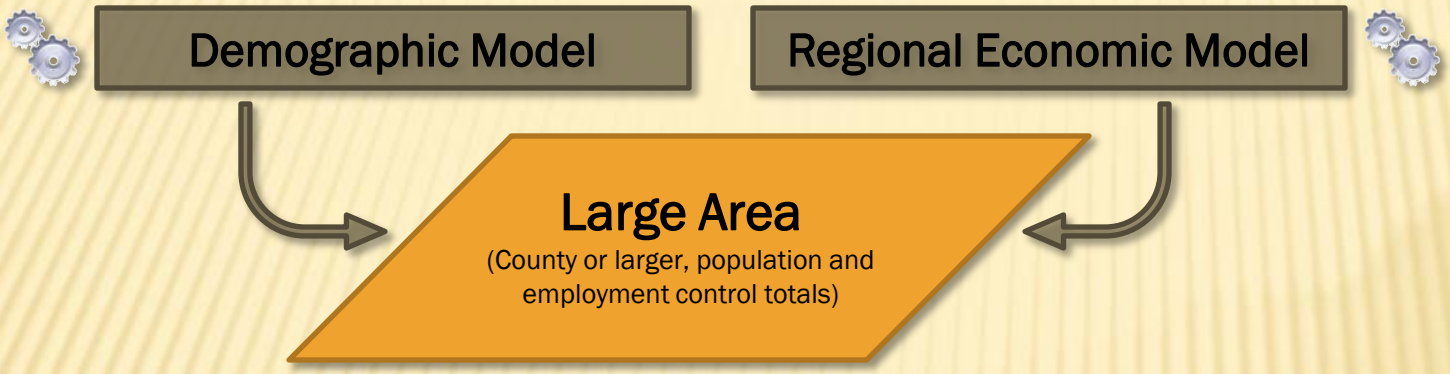
Copyright © 2012

AZ-SMART FULL MODEL SYSTEM



Copyright © 2012

AZ-SMART LARGE AREA MODELS

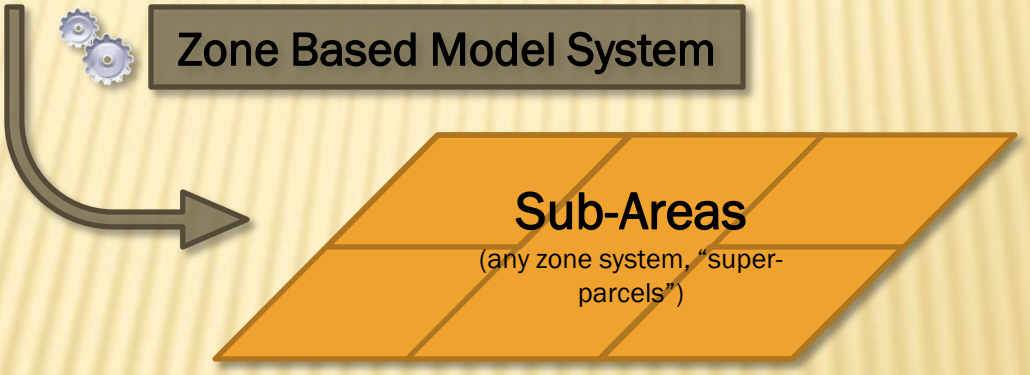


- **Demographic Model:**
  - Official population projections at the county level for the State of AZ
  - Model developed by the AZ Office of Employment and Population Statistics (AZ Department of Administration)
- **Regional Economic Model:**
  - Projections of employment by NAICS code at the county level
  - Planning to purchase projections of employment from a commercial source for 2012 official projections (e.g. Global Insight, Moody's Analytics, etc.)
  - Currently scoping a model to produce these projections in-house

Copyright © 2012

AZ-SMART MID-LEVEL MODEL: URBANSIM

- UrbanSim
  - Fully disaggregate microsimulation model
  - Can model at any level of geography (zones, parcels, gridcells)

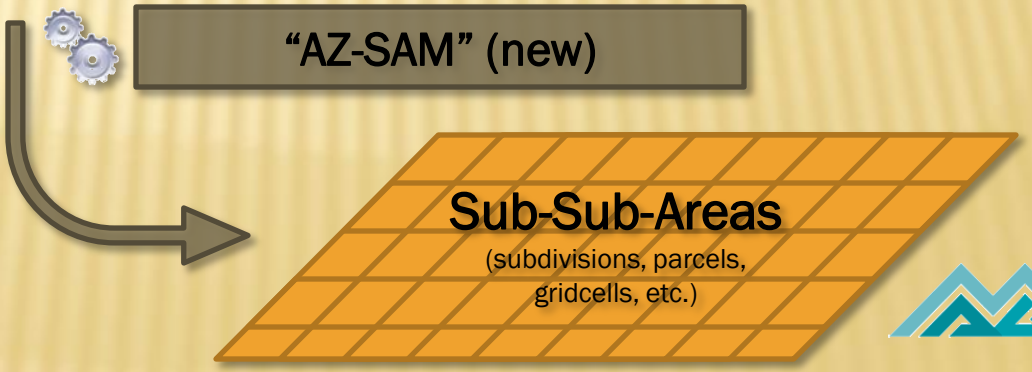


  - Developed at the University of WA by Paul Waddell
  - Funded by the EPA, NSF, FHWA, plus numerous COGs\MPOs
  - Emerging as a national standard for integrated land use\transportation modeling

Copyright © 2012

AZ-SMART SMALL AREA MODEL

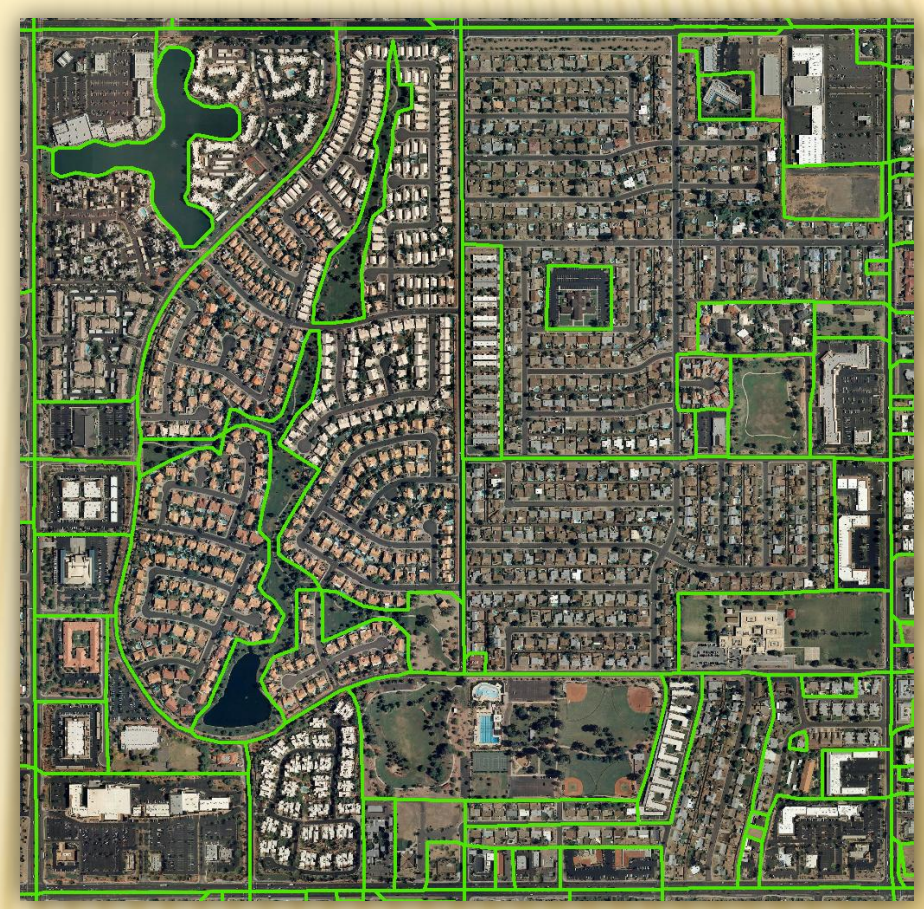
- GIS based model to allocate UrbanSim results down to smaller levels of geography
- Currently being scoped and developed in collaboration with a consultant
- Will allow aggregation of results to any geography
- Will have additional tools for exporting results to a travel model, reporting, visualization, etc.



Copyright © 2012

URBANSIM: ZONE BASED MODEL SYSTEM

- Boundaries are arbitrary, can match Census Tract, Block group, TAZs, or some other geography, such as "super-parcels" (see picture).
- Land and built space is aggregate in nature – land use plans and constraints incorporated as capacities
- Retains individual household, person, and job records



Copyright © 2012

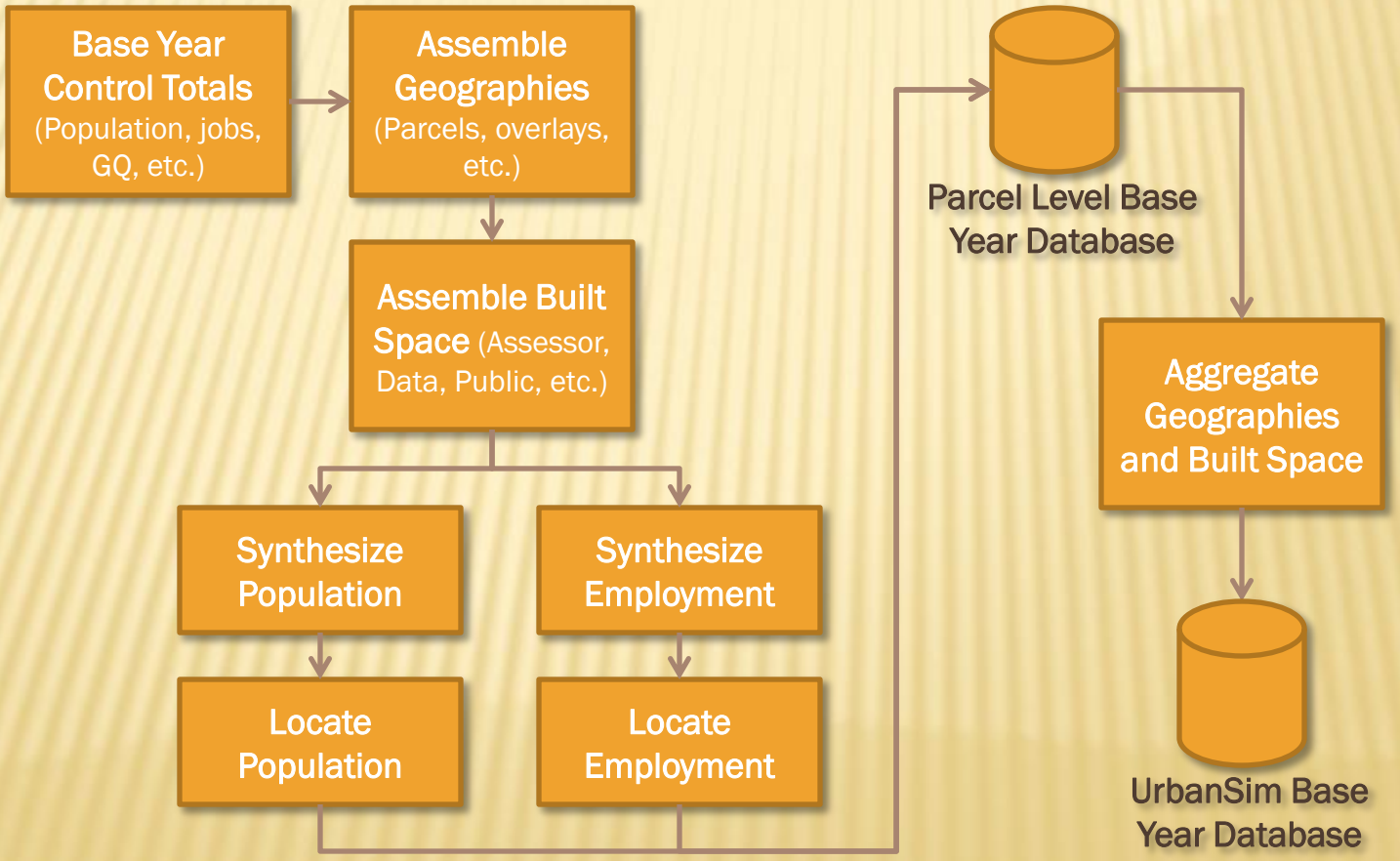
AZ-SMART

Input and Base Year Data

Copyright © 2012

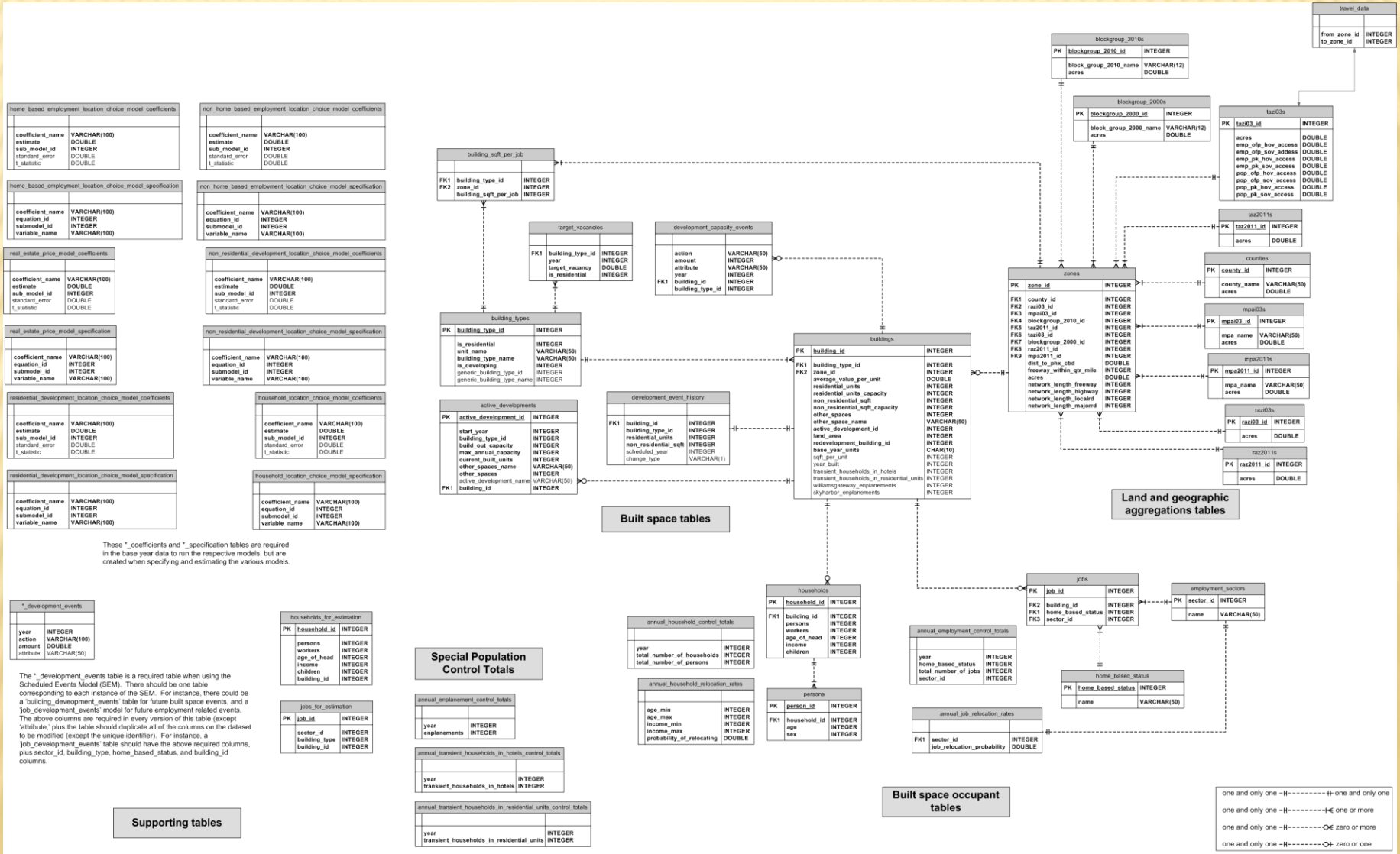
BASE YEAR DATA PREPARATION PROCESS

- Data are prepared, maintained, and synthesized at the parcel level, then aggregated as necessary for simulation modeling



Copyright © 2012

AZ-SMART BASE YEAR DATA MODEL DIAGRAM



AZ-SMART Base Zone Data Model – *DRAFT*

DATA MODEL – KEY TABLES & RELATIONSHIPS (1)

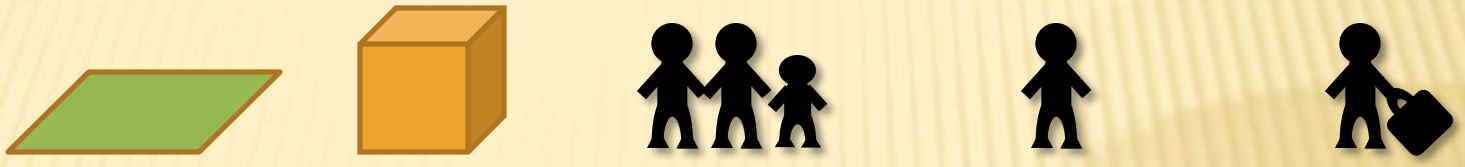
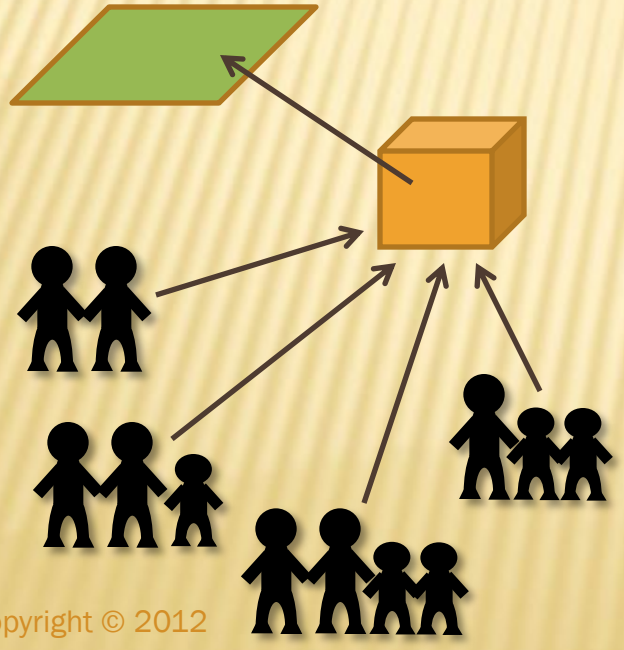


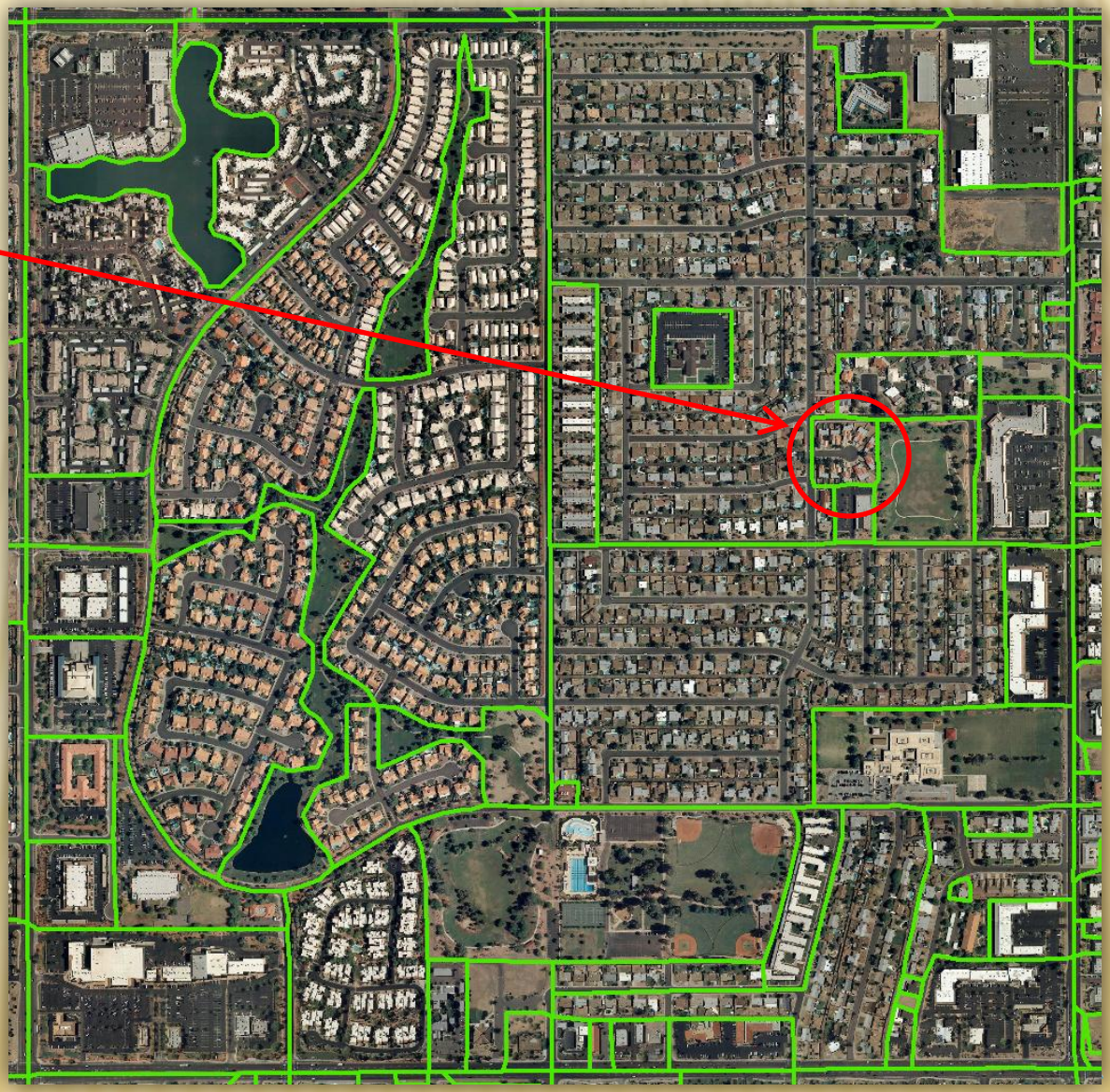
Table	Zones (“super-parcels”)	Buildings (built space)	Households	Persons	Jobs
Primary Key	zone_id	building_id	household_id	person_id	job_id
Foreign Key	taz_id, mpa_id, tract_id, etc.	zone_id	building_id	household_id	building_id
# Records	~71 k	~300 k	~1.44 million	~3.93 million	~1.79 million

Copyright © 2012

(counts are approximate from Maricopa County test database)

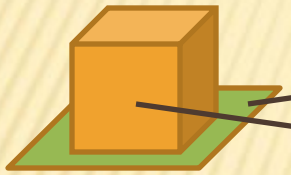
DATA MODEL – KEY TABLES & RELATIONSHIPS (2)





Copyright © 2012

DATA MODEL – KEY TABLES & RELATIONSHIPS (3)



Zones

zone_id	raz_id	mpa_id	county_id	area	freeway_within_qtr_mi	dist_to_phx_cbd	flood_plain
1671	204	26	1	174240	0	185042	0

Buildings

building_id	building_type_id	zone_id	land_area	residential_units	residential_units_capacity	non_residential_sqft	non_residential_sqft_capacity	sqft_per_unit
875	1	1671	174240	10	10	0	0	1964

Households

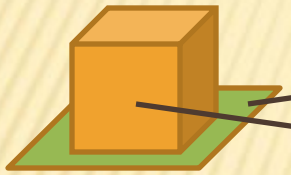
household_id	building_id	children	age_of_head	income	persons	workers
812943	875	1	60	122616	2	2
812960	875	1	30	79910	2	1
813534	875	0	39	44671	2	0
813556	875	0	38	37990	2	0
814097	875	0	71	74408	2	0
814118	875	0	65	76766	2	1
814258	875	0	31	93010	3	3
814287	875	0	67	129690	2	1
814434	875	0	25	0	1	0
814435	875	0	44	0	1	0

Persons

person_id	household_id	person_no	head_of_hh	age	sex	education	student_status	marriage_status
2170137	812943	1	1	60	1	11	1	1
2170138	812943	2	0	17	2	7	2	6
2170171	812960	1	1	30	1	8	1	1
2170172	812960	2	0	2	1	0	0	0
2171893	813534	1	1	39	1	5	1	1
2171894	813534	2	0	35	2	1	1	1
2171937	813556	1	1	38	1	9	1	1
2171938	813556	2	0	39	2	10	1	1
2173361	814097	1	1	71	1	9	1	1
2173362	814097	2	0	65	2	12	1	1
2173403	814118	1	1	65	2	14	1	1
2173404	814118	2	0	56	1	14	1	1
2173725	814258	1	1	31	1	10	1	6
2173726	814258	2	0	29	1	13	1	6
2173727	814258	3	0	28	1	13	1	6
2173857	814287	1	1	67	1	9	1	1
2173858	814287	2	0	59	2	9	1	1
2174154	814434	1	1	25	1	13	3	6
2174155	814435	1	1	44	1	11	1	6

Copyright © 2012

DATA MODEL – KEY TABLES & RELATIONSHIPS (4)



Zones

zone_id	raz_id	mpa_id	county_id	area	freeway_within_qtr_mi	dist_to_phx_cbd	flood_plain
8479	265	27	1	849800	1	86950	0

Buildings

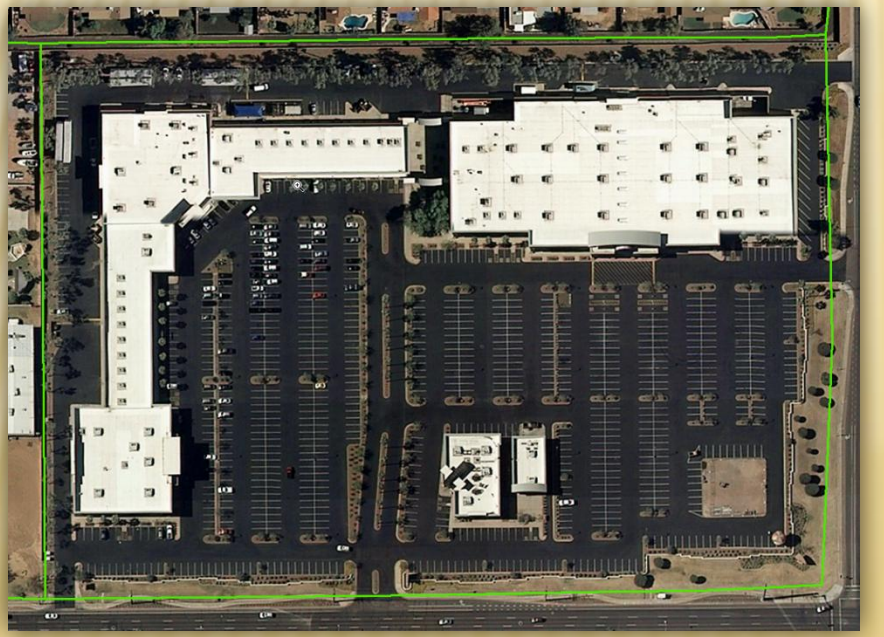
building_id	building_type_id	zone_id	land_area	residential_units	residential_units_capacity	non_residential_sqft	non_residential_sqft_capacity	sqft_per_unit
875	3	8479	849800	0	0	162500	162500	1

Jobs

job_id	building_id	sector_id	home_based_status
814097	875	3	0
814118	875	3	0
814258	875	3	0
814287	875	3	0
814434	875	3	0

building_type_id	building_sqft_per_job
1	150
2	150
3	800
4	400
5	350
6	500
7	500
8	750
9	1000
10	250

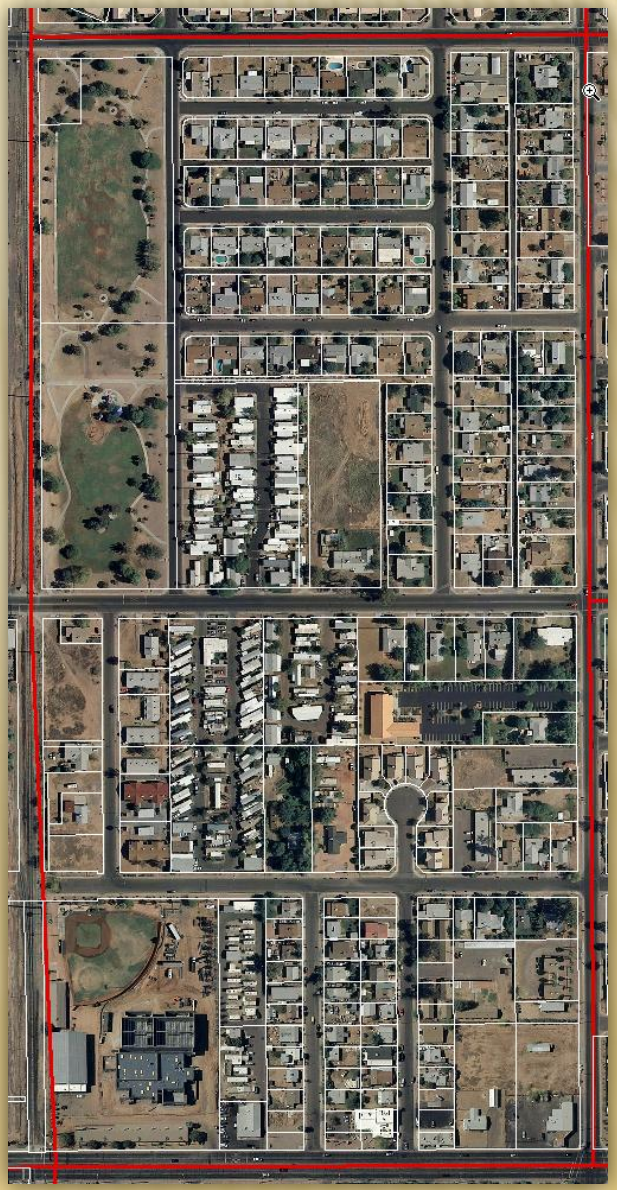
building_sqft_per_job



Copyright © 2012

POPULATION AND EMPLOYMENT SYNTHESIS

1. Given Census Block Group information (e.g. #hh by income categories & size, #pp by age and sex)
2. Generate frequency matrix (via ASU's IPF/IPU methodology) of household and person type constraints by block group
3. Draw from Census PUMS data detailed household and person records that match the constraints
4. Use detailed information from PUMS records (structure size, type, value, etc.) to match or "locate" households to Assessor parcels and built-space information with similar attributes



Copyright © 2012

DRAFT BUILDING TYPES

building_type_id	building_type_name	building_type_description
1	rsf	Single Family Detached Home
2	rmf	Multi Family Attached Home
3	mh	Mobile / Manufactured Home
4	retl	Retail
5	stor	MiniStorage
6	ware	Warehouse
7	ind	Industrial
8	off	Office
9	med	Medical
10	hot	Hotel
11	civic	Civic
12	edu	Education
13	gq	Group Quarters
14	pubf	Public - Federal
15	pubs	Public - State
16	publ	Public - Local
17	ag	Agriculture
18	tpt	Transportation
19	other imp	Other
20	os	Open Space

Copyright © 2012

DRAFT EMPLOYMENT SECTORS

sector_id	name	naics_code
1	Agriculture, Forestry, Fishing, Hunting	11
2	Mining, Quarrying, Oil, Gas	21
3	Utilities	22
4	Construction	23
5	Manufacturing	31-33
6	Wholesale Trade	42
7	Retail Trade	44-45
8	Transportation and Warehousing	48-49
9	Information	51
10	Finance and Insurance	52
11	Real Estate and Rental and Leasing	53
12	Professional, Scientific, Technical Svcs	54
13	Management of Companies and Enterprises	55
14	Admin., Support, Waste Mgmt, Remediation	56
15	Educational Services	61
16	Health Care and Social Assistance	62
17	Arts, Entertainment, and Recreation	71
18	Accommodation	721
19	Food Services and Drinking Places	722
20	Other Services (except Public Admin)	81
21	Public - Federal	92
22	Public - State	92
23	Public - Local	92

Copyright © 2012

AZ-SMART

Models: Simulation Sequence & Description

Copyright © 2012

WHAT IS A MODEL?

- A mathematical or statistical abstraction to help us understand and make predictions about real world systems, often consists of:
  - A set of simplifying assumptions
  - Initial conditions
  - A range of applicability
- Models can be used to:
  - Make predictions about the future
  - Understand what variables influence certain outcomes

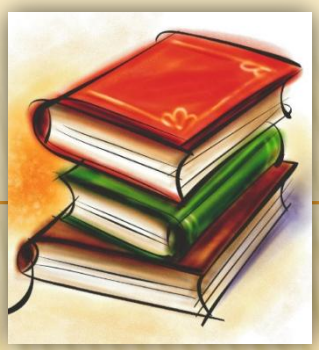
Copyright © 2012

AZ-SMART GENERAL MODEL TYPES

- **"Simple" model:**
  - Can be thought of as a simple calculation on tables (e.g. "SQL-like")
- **Allocation model:**
  - Allocates a given total of 'X' to dataset 'Y' based on weights in 'Y'
- **"Rate Based Change" model:**
  - Randomly selects some percentage of some dataset based on pre-computed change rates
- **Regression model:**
  - Can be used to predict any continuous variable in a dataset during the simulation
- **Probability or Choice model:**
  - Discrete choice: binary (2 choice) or multinomial logit (2+ choices)

Copyright © 2012

AZ-SMART MODELS: THEORETICAL BASES



- Random Utility Theory (McFadden)
  - Agents choose amongst alternatives based on relative utility
- Urban Economics/Bid Rent Theory (Alonso, Mills, Muth)
  - Explains land use as a function of willingness to pay for amenities and location trade-offs
- Hedonic Price Theory (Rosen)
  - Composite goods (housing) can be “disaggregated” into their attributes and priced
- Dynamic Market Equilibrium, Disequilibrium
  - Real-world markets rarely in equilibrium, imperfect information abounds
- Microsimulation (Orcutt)
  - Helps avoid aggregation bias and ecological fallacy
- GIS and Spatial Analysis (Tobler, Anselin)

Copyright © 2012

Source: P. Waddell, 2010

AZ-SMART SIMULATION SEQUENCE

Demographic Evolution Models

↓

Real Estate Development Models

↓

Household Location Models

↓

Employment Location Models

- Simulates demographic change in households and persons
- Predicts price, location, timing, and types of real estate development and re-development
- Simulates migration and location choices of households
- Simulates growth (or decline) and location of regional employment

Copyright © 2012

DEMOGRAPHIC EVOLUTION MODELS (1)

- Models predict individual household and person-level demographic events
- Rates applied as probabilities that specific demographic events will occur
- Primary purpose: 1st attempt to reduce use of "static" agents (households and persons)
- Sometimes called Demographic Transition Models (DTM) in academic literature
- Major MAG enhancement

Aging Model

Fertility Model

Mortality Model

Education Model

Marriage Model

Cohabitation Model

Divorce Model

Breakup Model

Child-Leaving-Home Model

Roommate Model

Household Workers Model

Household Income Model

Copyright © 2012

FERTILITY MODEL EXAMPLE

- Divide persons into bins by the attributes in the rates table
- Randomly select persons using the probability as a percentage
- Create a new person record in the household and initialize and update relevant attributes

age_min	age_max	sex_min	sex_max	fertility_probability
11	14	2	2	0.0007
15	19	2	2	0.0549
20	24	2	2	0.1222
25	29	2	2	0.1195
30	34	2	2	0.0942
35	39	2	2	0.0462
40	44	2	2	0.009
45	55	2	2	0.0006

Persons in Household Before Model

person_id	household_id	person_no	head_of_hh	age	sex	education	student_status	marriage_status
35	15	1	1	39	1	10	1	1
36	15	2	0	31	2	9	1	1
37	15	3	0	8	2	2	3	0

New Person Record

Persons in Household After Model

person_id	household_id	person_no	head_of_hh	age	sex	education	student_status	marriage_status
35	15	1	1	39	1	10	1	1
36	15	2	0	31	2	9	1	1
37	15	3	0	8	2	2	3	0
38	15	4	0	0	2	0	0	0

Copyright © 2012

AZ-SMART SIMULATION SEQUENCE
REAL ESTATE DEVELOPMENT

Demographic Evolution Models

↓

Real Estate Development Models

↓

Household Location Models

↓

Employment Location Models

Copyright © 2012

REAL ESTATE DEVELOPMENT MODELS

- Purpose: simulate the development of residential and non-residential space for new or relocating occupants
- Demand conceptualized as vacancy rates falling below "normal" or "balanced" market thresholds
- Supply conceptualized from the real estate developers perspective of a "project looking for a location"

Regression Model	Real Estate Price Model
*Custom model written in-house at MAG	Active Developments Model
Rate-Based Change Model	Development Project Transition Model
Discrete Choice Model	Residential Development Project LCM
Discrete Choice Model	Non-Residential Development LCM
Simple accounting model	Building Construction Model
*Custom model written in-house at MAG	Redevelopment Model

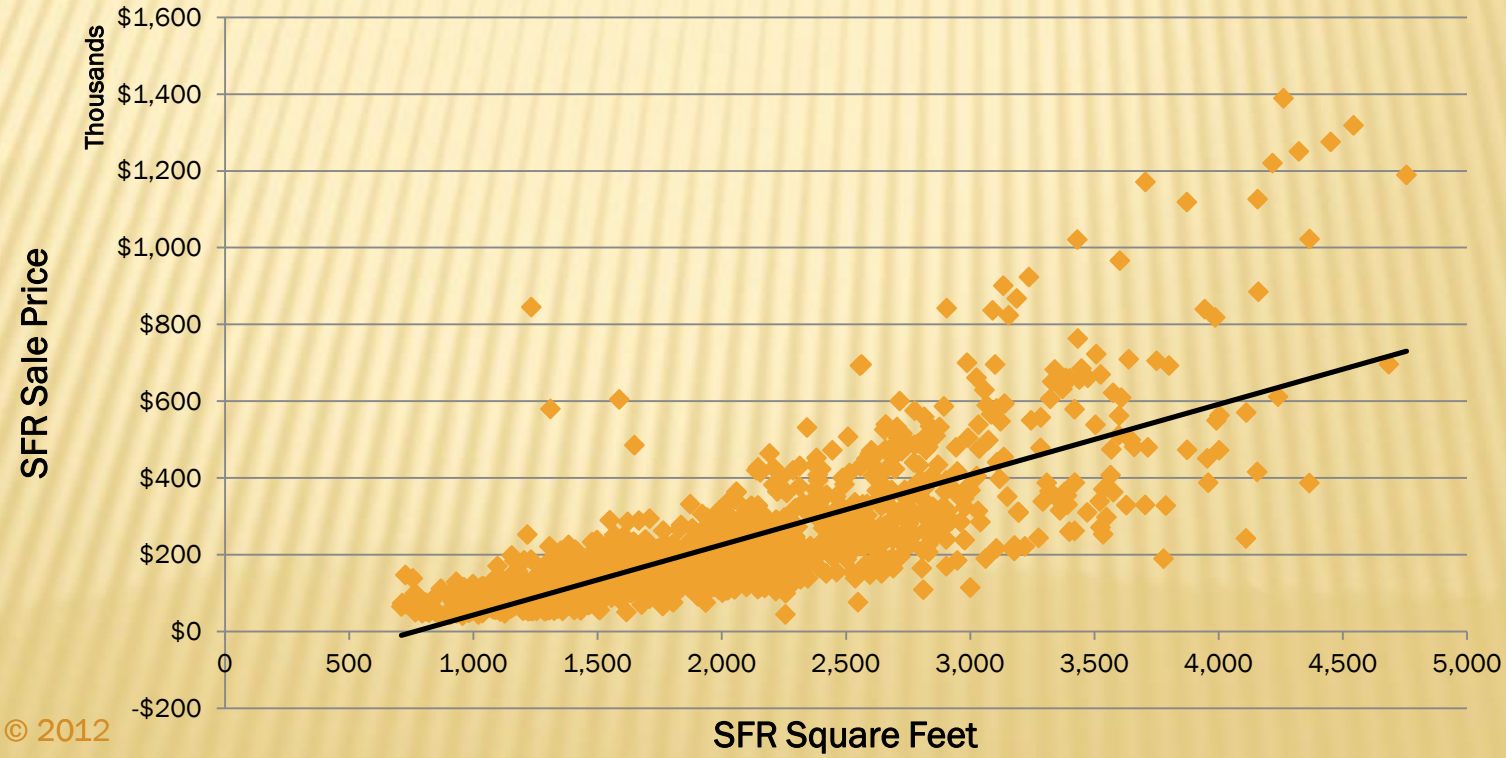
Copyright © 2012

REAL ESTATE PRICE MODEL
PARAMETER ESTIMATION

- Estimation of a linear regression model consists of adjusting the parameters of a model function to best fit a sample set of **observed** values.

- The goal is to find the parameter values for the line which "best fit" the observed data. The REPM is estimated using the standard OLS method to minimize the mean squared error.

price = (β)home sqft



Copyright © 2012

REAL ESTATE DEVELOPMENT
REAL ESTATE PRICE MODEL (REPM): SIMPLE EXAMPLE

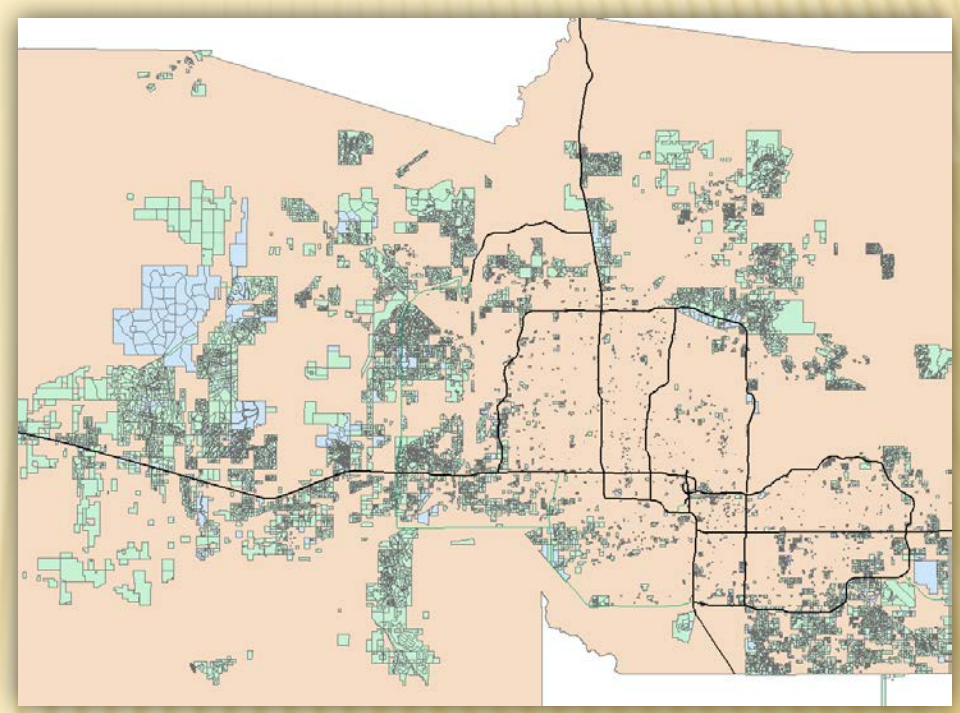
- Primary purpose: prices influence location decisions
- Model stratified into submodels by building type
  - Different variables influence SFR vs. Industrial prices
- Standard OLS estimated regression model with $/sqft as dependent variable
- Predicted prices then used as inputs to influence the behavior of other models (e.g. HLCM, ELCM)

price = (β)home sqft + (β)lot sqft + (β)within 1/4mi of landfill

Price	=	β	*	Home Sqft	+	β	*	Lot Sqft	+	β	*	0.25mi landfill?
Price	=	14.63	*	2365	+	2.34	*	43560	+	-5630	*	1
Price	=	34599.95			+	101930.4			+	-5630		
Price	=	$130,900.40										

REAL ESTATE DEVELOPMENT
ACTIVE DEVELOPMENTS MODEL (ADM)

- Prioritizes developments in the real-estate development process
- Vacancy rates are computed for the simulation year, then compared with the target vacancy rates



- Total number of units demanded by built-space type computed from simulation data vs. target vacancy computation
- Configurable % of total demand satisfied by ADM (based on historic analysis)

Copyright © 2012

REAL ESTATE DEVELOPMENT
GENERAL PLAN AND FUTURE DEVELOPMENTS MODEL

- After the ADM, the remainder of units demanded satisfied with development allowed by the general plan
- Samples from a table of past development until target vacancy rates are satisfied
- Res and Non-Res LCMs locate units
- Built space table updated with information from located developments

Real Estate Price Model

Active Developments Model

Development Project Transition Model

Residential Development Project LCM

Non-Residential Development LCM

Building Construction Model

Redevelopment Model

Copyright © 2012

REAL ESTATE DEVELOPMENT
GENERAL PLAN AND FUTURE DEVELOPMENTS MODEL (2)

Step 1:

Compute Demand for units by unit type

$$TotalUnitsDemanded = \frac{\#\,Households}{TgtVacancyRate}$$

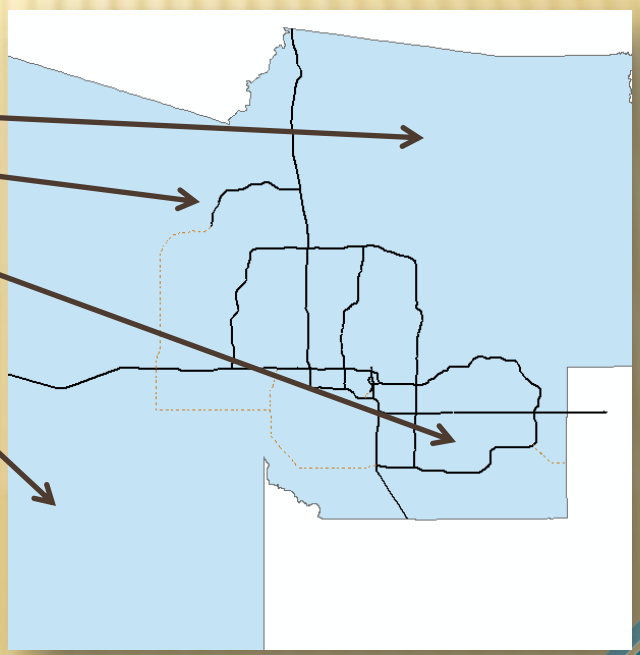
Step 2:

Randomly sample from past developments by type until the total = demand

Step 3:

Use "Development Project LCM" to predict a specific location for each development (subject to land constraints)

Total Units Demanded	Building Type	# DUs	# Non-Res SQFT	Zone ID
107	SFR	2	0	?
	SFR	26	0	?
	SFR	49	0	?
	SFR	16	0	?
	SFR	14	0	?
115	MFR	12	0	?
	MFR	25	0	?
	MFR	19	0	?
	MFR	59	0	?
220,000	RET	0	50,000	?
	RET	0	20,000	?
	RET	0	150,000	?
107,000	OFF	0	25,000	?
	OFF	0	18,000	?
	OFF	0	64,000	?



Copyright

REAL ESTATE DEVELOPMENT
RE-DEVELOPMENT MODEL

- Custom written in-house at MAG
- Based on re-development data from MAG development database
- Designed to satisfy the need to accommodate known redevelopment projects or areas
- Displacement of existing population and employment

Real Estate Price Model

Active Developments Model

Development Project Transition Model

Residential Development Project LCM

Non-Residential Development LCM

Building Construction Model

Redevelopment Model

Copyright © 2012

AZ-SMART SIMULATION SEQUENCE
HOUSEHOLD AND EMPLOYMENT LOCATION CHOICE

Demographic Evolution Models

↓

Real Estate Development Models

↓

Household Location Models

↓

Employment Location Models

Copyright © 2012

HOUSEHOLD AND EMPLOYMENT MODELS

- Adds and removes households and jobs based on control totals
  - Annually rectifies employment, household, and person totals to match the DOA's exogenous prediction
- LCMs locate households and jobs in vacant residential and non-residential space
  - Predicts a specific built-space record for a household or job in need of a location
- Different types of households and employment have different variables that influence location decisions
  - Household LCM stratified into models by type
  - Employment LCM stratified into sectors (roughly corresponding to 2 digit NAICS)

Household Transition Model

Household Location Choice Model

Employment Transition Model

Employment Location Choice Model

Copyright © 2012

HOUSEHOLD AND EMPLOYMENT TRANSITION MODELS

Household Transition Model

Employment Transition Model

- Step 1: Calculate Simulation Population by bins in Control Totals Table and difference
- Step 2: Randomly duplicate or delete population records as necessary

Annual Population Control Totals Table

year	sex	age_min	age_max	Population Control Total	Simulation Population	Difference
2016	1	0	13	53684	52443	+ 1241
2016	1	14	17	52482	53451	- 969
2016	1	18	35	50472	48421	+ 2051
2016	1	36	55	49190	51216	- 2026
2016	1	56	75	48471	50994	- 2523
2016	1	76	-1	49621	48341	+ 1280
2016	2	0	13	48916	47154	+ 1762
2016	2	14	17	51751	50985	+ 766
2016	2	18	35	53478	52187	+ 1291
2016	2	36	55	52971	51924	+ 1047
2016	2	56	75	49168	51369	- 2201
2016	2	76	-1	47154	49271	- 2117

Copyright © 2012

LOCATION CHOICE MODEL EXAMPLE

1. Determine household records w/out a location
2. Identify vacant housing units
3. Randomly sample from vacant units for a reasonably sized choice set
4. Calculate utility and probability (based on household and housing unit characteristics)
5. Use random sampling to determine the choice

House	Utility	Probability
A	300	0.259
B	600	0.517
C	200	0.172
D	50	0.043
E	10	0.009

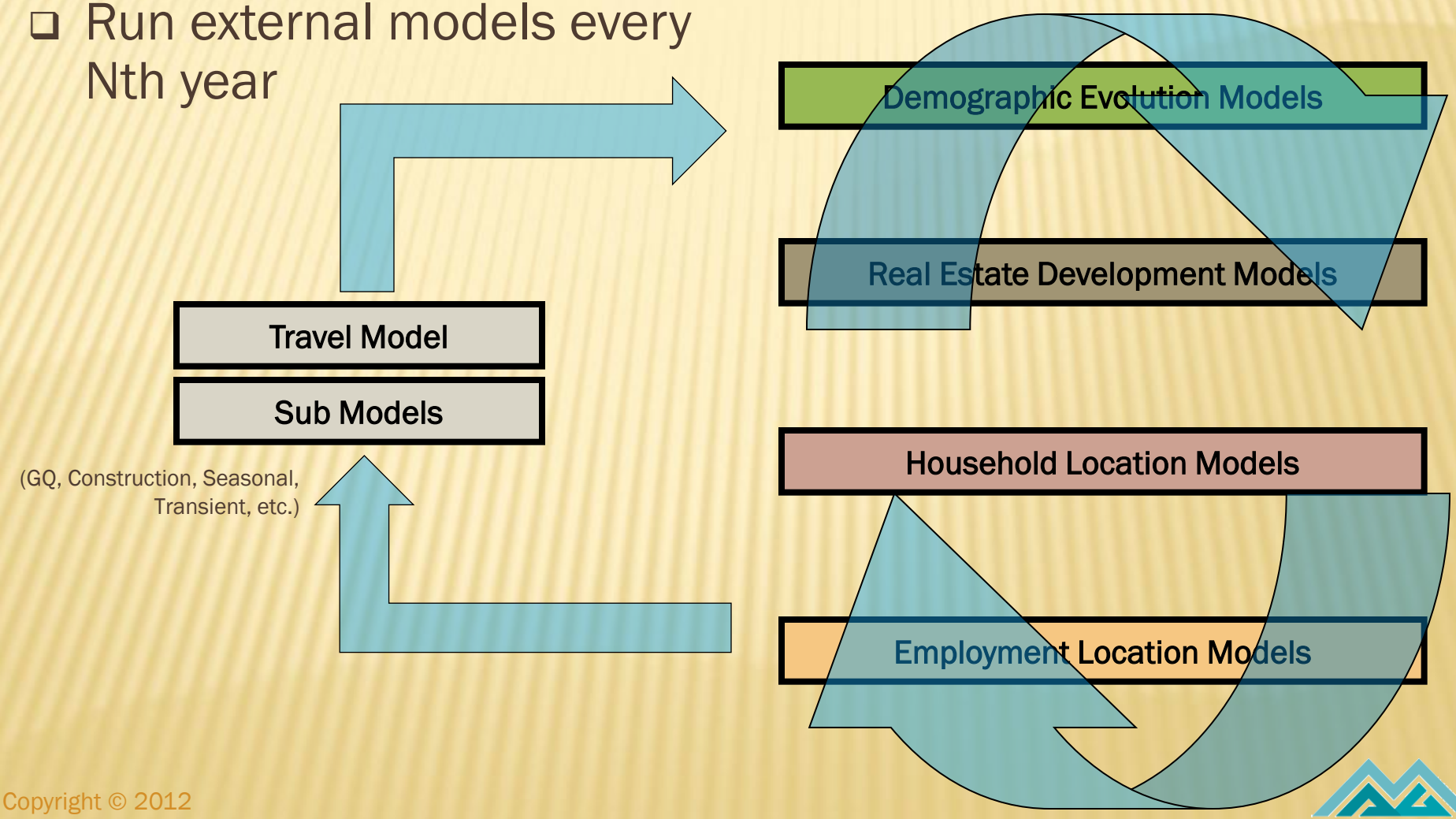
House	Ordered Probability	Cumulative Proability	Random Number
E	0.009	0.009	
D	0.043	0.052	
C	0.172	0.224	
A	0.259	0.483	0.480
B	0.517	1.000	

$$v^n(j) = \sum_i^I \beta_{ji} X_{ji} + \beta_{njt} X_n X_{ji}$$

$$\Pr^n(j) = \frac{e^{V_j}}{\sum_j e^{V_j}}$$

AZ-SMART OVERALL MODEL SEQUENCE: RECAP

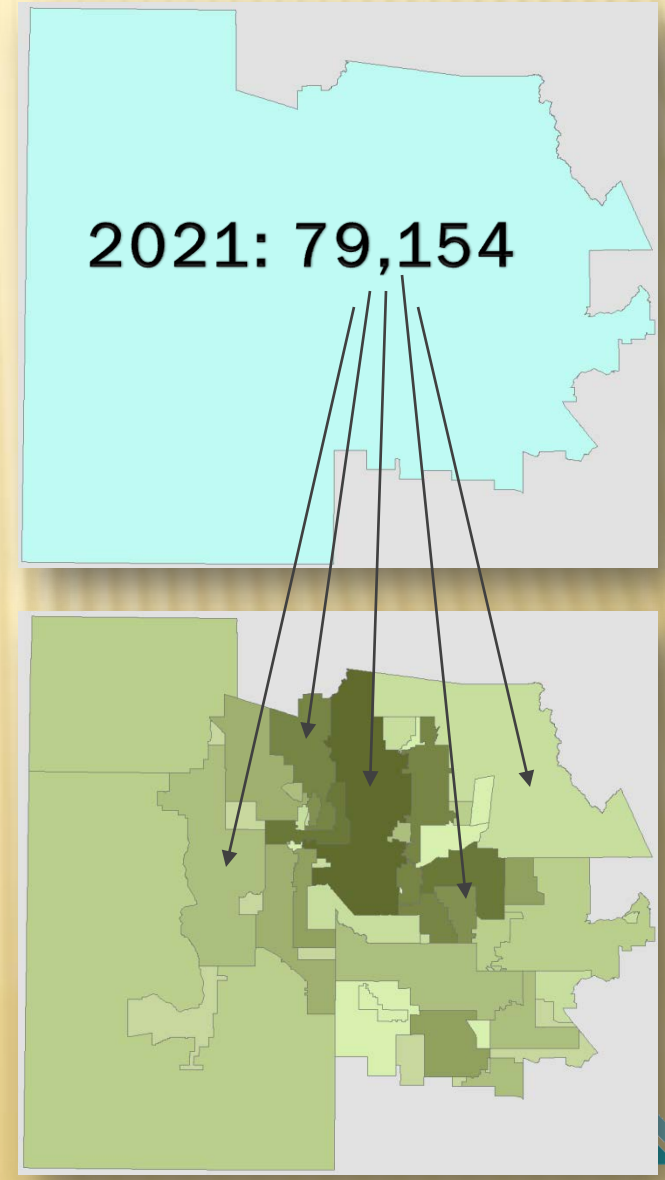
- Models repeat for every year
- Run external models every Nth year



Copyright © 2012

ALLOCATION MODEL EXAMPLE
HOTEL VISITOR SUBMODEL

- Data inputs:
  - forecast of total hotel visitors for every year of the forecast period
  - number of hotel rooms and occupancy rates by city
- Proportionally allocate the total number of hotel visitors to each City based on the number of hotel rooms and occupancy rates
- Constrain the allocation based on the capacity



Copyright © 2012

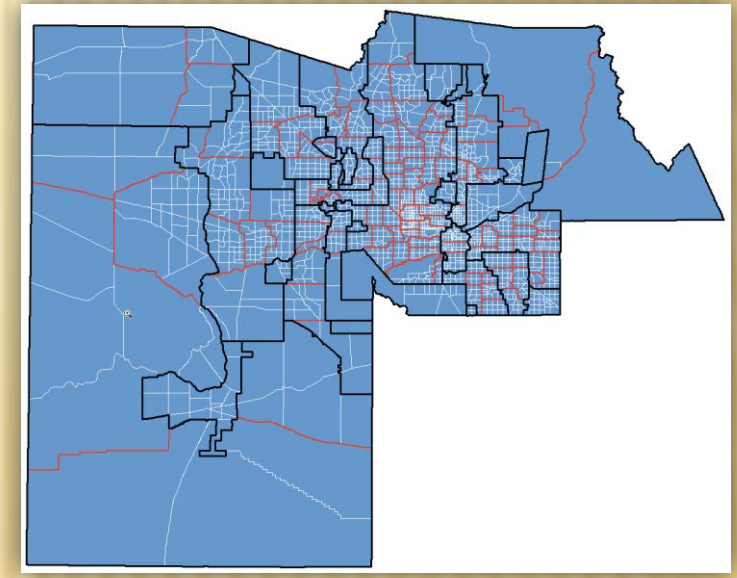
AZ-SMART

2012 Official Projections

Copyright © 2012

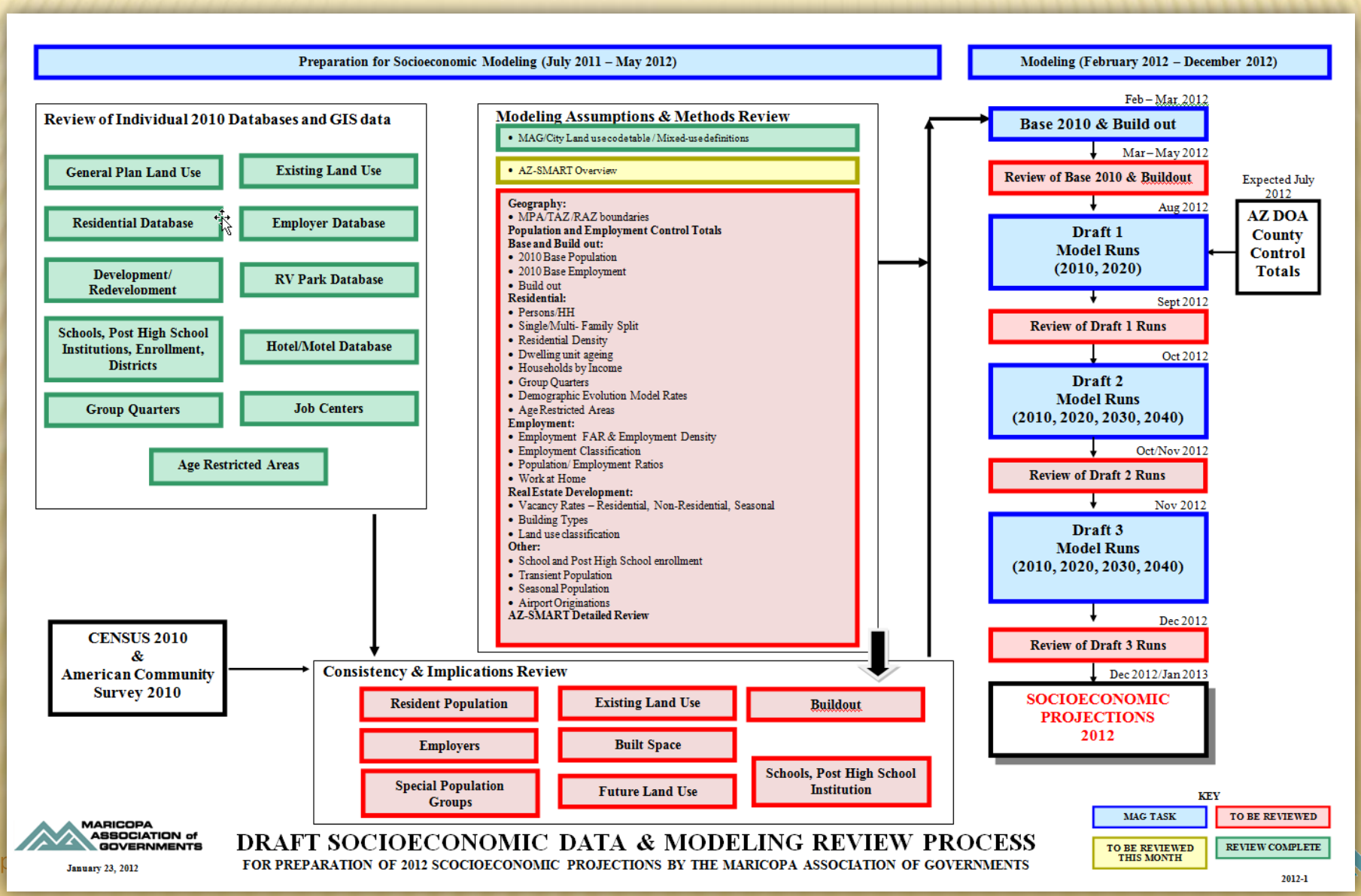
2012 OFFICIAL PROJECTIONS

- 2012 Projections:
  - 2010 Base Year, annual projections to 2040
  - Last set in July 2007
  - Executive Order 2011-04 – Projections prepared every 3 to 4 years
  - Based on DES county control totals
    - Working with AZ State Demographers Office and Council for Technical Solutions
- 2012 Projections at 3 levels of geography:
  - Municipal Planning Areas (MPAs): 29
  - Regional Analysis Zones (RAZs): ~149
  - Socioeconomic Analysis Zones (SAZs): ~2095
- MAG Approval Process
  - POPTAC
  - Management Committee
  - Regional Council
- 40+ variables projected
  - Housing units, households, population in households and group quarters
  - Employment by 5 land use and 2 other categories
  - Employment by 15+ NAICS categories
  - Additional variables like seasonal, transient population, etc.



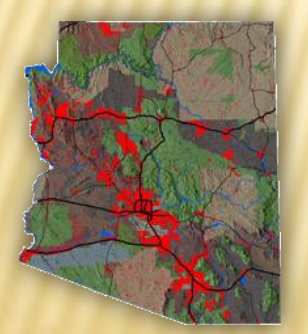
Copyright © 2012

2012 PROJECTIONS REVIEW PROCESS



QUESTIONS?

Team:
Hanyi Li
Jesse Ayers
Mark Roberts
Scott Bridwell



Anubhav Bagley
abagley@azmag.gov

(602) 254-6300

Copyright © 2012