



Application of a Socio-Economic Model System for Activity-based Modeling

Experience from Southern California

Ram M. Pendyala, Karthik C. Konduri and Keith P. Christian
Arizona State University, Tempe
Chandra R. Bhat, Rajesh Paleti, Raghu Sidharthan
The University of Texas, Austin
Konstadinos G. Goulias
The University of California, Santa Barbara
Hsi-hwa Hu and Guoxiong Huang
Southern California Association of Governments, Los Angeles



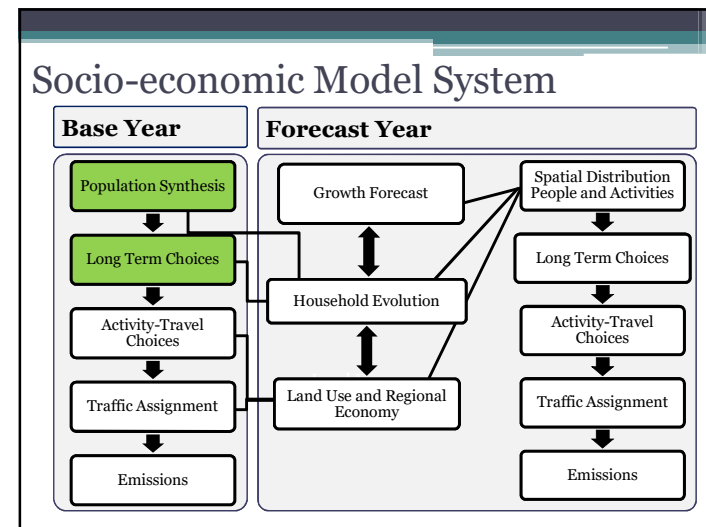
 SOUTHERN CALIFORNIA
ASSOCIATION OF GOVERNMENTS

Inputs to Activity-based Travel Model

- Need socio-economic and demographic attributes of entire population in a region
- Disaggregate data about entire population is generally not available
 - Disaggregate data about a sample of population
 - Control distributions of key variables of interest
- Generate a synthetic population using available data
 - Synthetic population generators are used
 - Key variables of interest are controlled and other attributes of interest are borrowed

Potential Issues with Borrowing

- Repetition and lack of heterogeneity
 - Problem magnified in the context of a large urban region such as SCAG
 - Population of ~17 million in 2003 and ~25 million in 2035
- Fails to capture variance in household- and person-level choice processes
 - Impact of education status → income, vehicle fleet composition
 - Work arrangements → activity-travel patterns
- Census lacks full set of inputs necessary for rich specification and simulation of activity-travel dimensions in a microsimulation model of travel demand

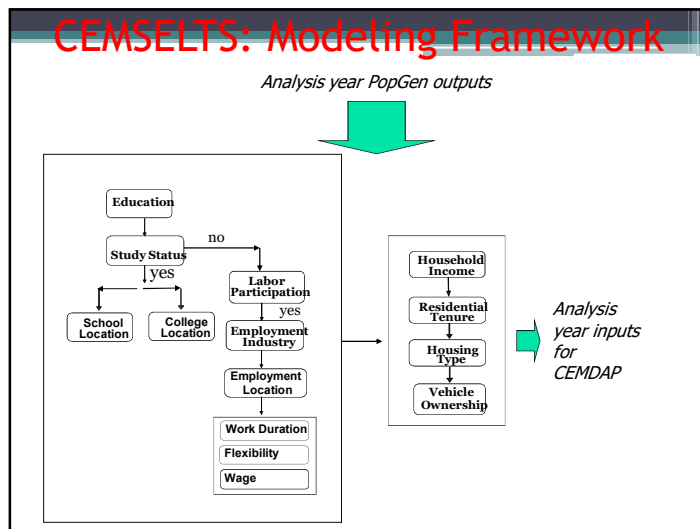


PopGen

- **Population Generator**
 - A population generator that creates synthetic households and persons in a region while ensuring that marginal distributions at household- and person-level are satisfied simultaneously
 - Computationally efficient and tractable
 - Open source software package

CEMSELTS

- **Comprehensive Econometric Microsimulator for Socioeconomics, Land-use, and Transportation System**
 - Simulates medium- and longer-term household- and person- level attributes of interest
 - Comprises a series of econometric models
 - Discrete choice and continuous duration models



8

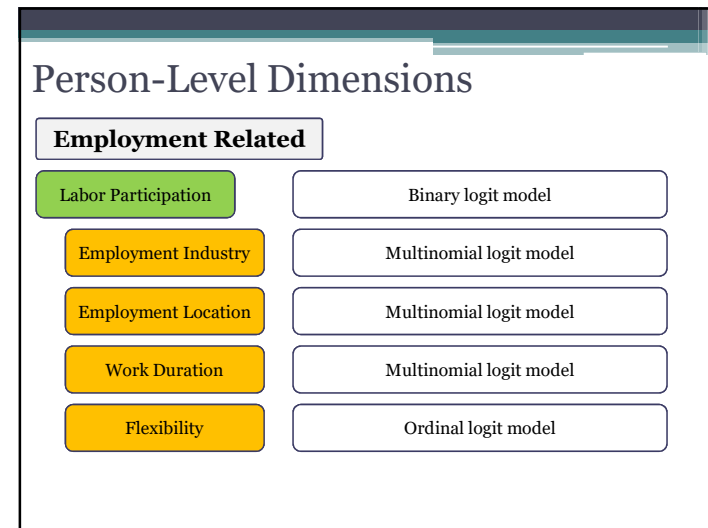
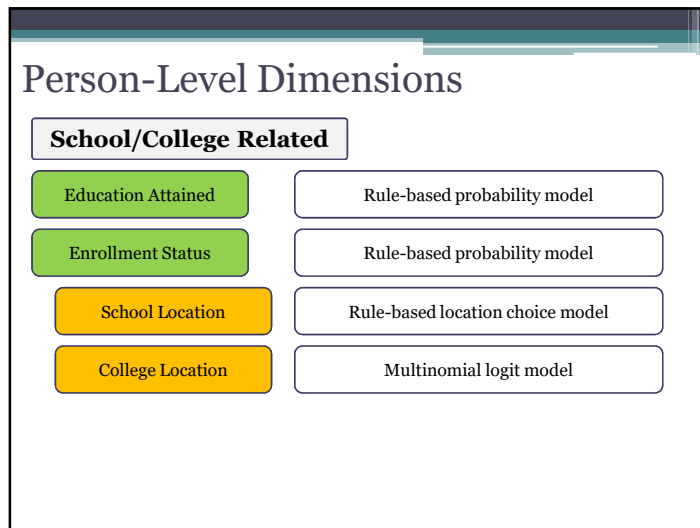
Person Models

Model name	Econometric Structure / Rule Based and Independent Variables	Choice Alternatives/Comments
Schooling		
For children aged < 5 years	Rule based model – all children under the age of 5 are considered as not going to school	Schooling or no schooling
For children between 5 to 12 years	Rule based model – all children between 5-12 years are assumed to attend school, and their grade is based on age	Grades K through 7
If age between 13 and 18 years	Rate-based probability model depending on age, race, and gender	Continue school, drop-out, or complete schooling. If drops out, grade is set to the grade at which drop-out occurs
If age > 18 years	Rate-based probability model for education level based on race.	Associate degree, bachelors, Masters, Ph.D.
School location of children	Deterministic Model: Closest zone which contains a school	List of zones with a school

No part of this publication may be reproduced, distributed, or transmitted in any form or by any means. For permission requests, write to the owner, addressed “Attention: Permissions Coordinator,” at ram.pendyala@asu.edu

Person Models 9		
Model name	Econometric Structure / Rule Based and Independent Variables	Choice Alternatives/Comments
Employment/Drivers License Models		
Labor participation model	Binary Logit model; independent variables include age, gender, race, education, presence of children	Employed, Not employed (Applied for individuals over 16 years of age and not studying)
Employment industry model	Multinomial logit model; ; independent variables include age, race, gender and education level	Construction and Manufacturing, Trade and Transportation, Professional businesses, Government, Retail and Repair, Other (relevant for employed individuals)
College Location	Spatial location choice model; independent variables include whether TAZ is a major /minor college TAZ, distance, race of person, accessibility measures of the TAZ	TAZs of SCAG area with Colleges
Employment location model	Spatial location choice model; independent variables include employment density, transportation level of service, accessibility to population and employment, and zones in central business district	TAZs of SCAG area
Weekly work duration model	Grouped response model; independent variables include gender, education level and industry	< 35 hours, 35-45 hours, and > 45 hours (the results are post processed to estimate a continuous "work hours" variable for each employed individual)
Work flexibility model	Ordered probit model; independent variables include gender, race , education level, employment industry, and hours worked.	No flexibility, Low flexibility, Medium flexibility, High flexibility (flexibility level definition is based on individual response in the survey)
License Model	Binary Logit model; independent variables include age, gender, and race	Has a valid license versus no valid license

Household Models 10		
Model Name	Econometric Structure and Independent Variables	Choice Alternatives
Household Income model	Grouped Response Model, Independent variables include race, presence of elderly individuals, education level of members of households, employment industry of workers in household	0-\$10,000; \$10,000-\$24,999, \$25,000-\$34,999, \$35,000-\$49,999, \$50,000-\$74,999, \$75,000-\$99,999, \$100,000-\$149,999, >\$150,000; Later converted into Continuous income
Residential tenure model	Binary logit model; independent variables include household income, household size, number of employed people, number of children, race, presence of elderly people, single-adult household and presence of unrelated people	Own or rent house
Housing type model	Multinomial logit model; independent variables include household income, race, presence of elderly people, single-adult household, presence of unrelated people and highest education level in the household	Single-family detached, Single-family attached, Apartment, and Mobile home or trailer



No part of this publication may be reproduced, distributed, or transmitted in any form or by any means. For permission requests, write to the owner, addressed "Attention: Permissions Coordinator," at ram.pendyala@asu.edu

Household-Level Dimensions

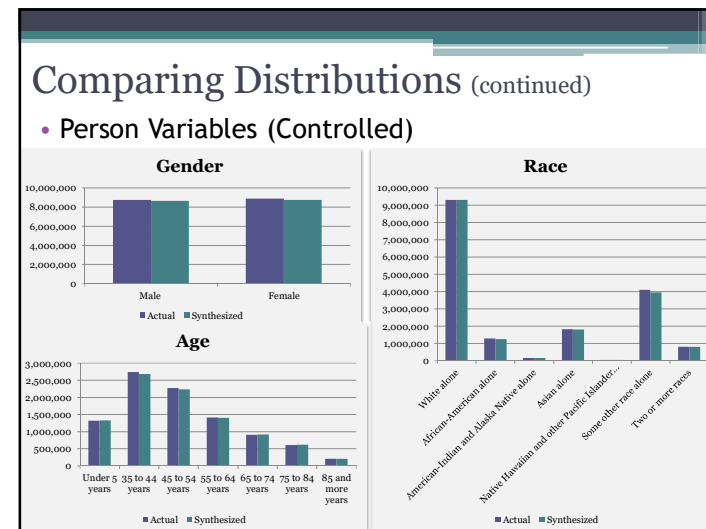
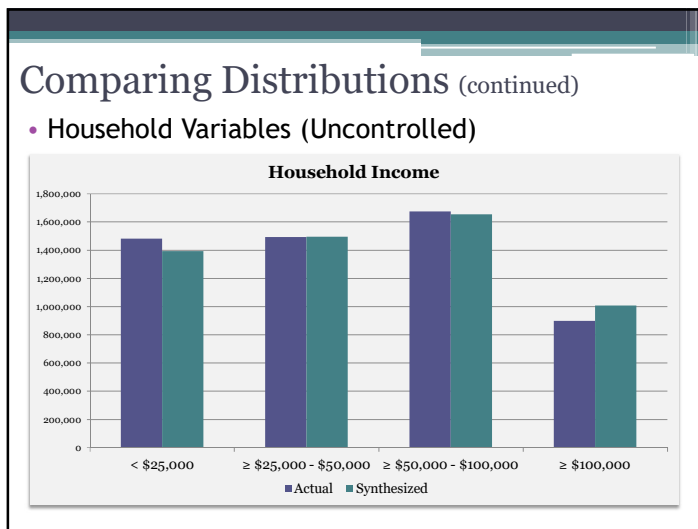
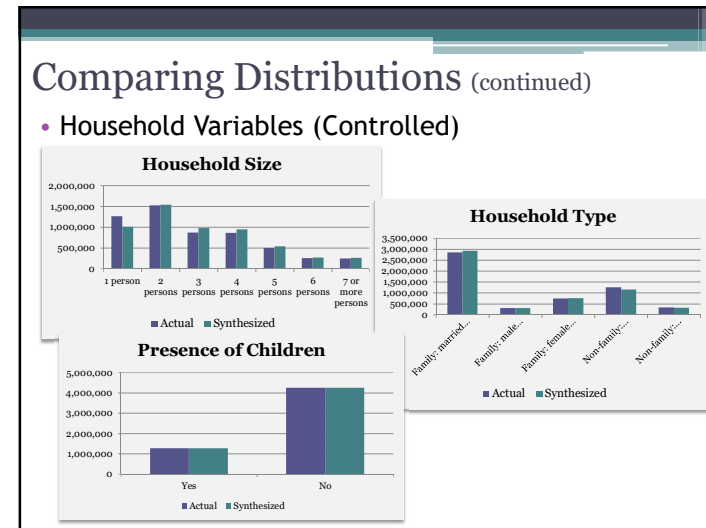
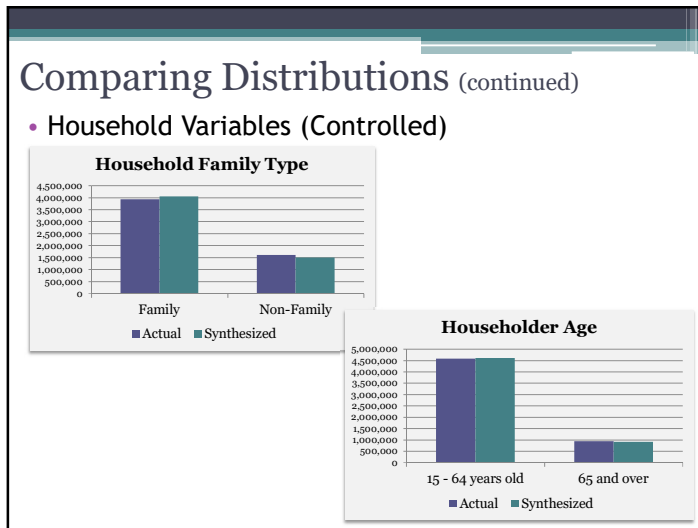
Household Income	Ordered logit model
Residential Tenure	Binary logit model
Housing Type	Multinomial logit model
Vehicle Ownership	
Annual Mileage	Log-linear regression model
Vehicle Fleet including Body Type and Vintage	Multiple Discrete Continuous Extreme Value (MDCEV) model
Vehicle Make	Multinomial logit model
Primary Driver Allocation	Multinomial logit model

- ### Study Area
- Includes 6 counties: Imperial, Los Angeles, Orange, Riverside, San Bernardino, Ventura
 - 4109 Traffic Analysis Zones
 - Population synthesis inputs
 - Census 2000 PUMS
 - Marginal distributions generated by SCAG

- ### PopGen Application
- #### Household attributes
- **presence of children (2 categories)**
1) Yes; 2) No
 - **household type (5 categories)**
1) Family: Married Couple; 2) Family: Male Householder, No Wife; 3) Family: Female Householder, No Husband; 4) Non-family: Householder Alone; 5) Non-family: Householder Not Alone
 - **household size (7 categories)**
1) 1 Person; 2) 2 Persons; 3) 3 Persons; 4) 4 Persons; 5) 5 Persons; 6) 6 Persons; 7) 7 or more Persons
 - **age of householder (2 categories)**
1) 15 – 64 years; 2) 65 years and over
 - **family type (2 categories)**
1) Family; 2) Non-family
 - **income (4 categories; uncontrolled used for validation)**
1) < \$25,000; 2) \$25,000 – \$49,999; 3) \$50,000 - \$99,999; 4) > \$99,999

- ### PopGen Application (continued)
- #### Person attributes
- **age (10 categories)**
1) Under 5 years; 2) 5 to 14 years; 3) 15 to 24 years; 4) 25 to 34 years; 5) 35 to 44 years; 6) 45 to 54 years; 7) 55 to 64 years; 8) 65 to 74 years; 9) 75 to 84 years; 10) 85 and more
 - **gender (2 categories)**
1) Male; 2) Female
 - **race (7 categories)**
1) White alone; 2) Black or African American alone; 3) American Indian and Alaska Native alone; 4) Asian alone; 5) Native Hawaiian and Other Pacific Islander alone; 6) Some other race alone; 7) Two or more races

No part of this publication may be reproduced, distributed, or transmitted in any form or by any means. For permission requests, write to the owner, addressed “Attention: Permissions Coordinator,” at ram.pendyala@asu.edu



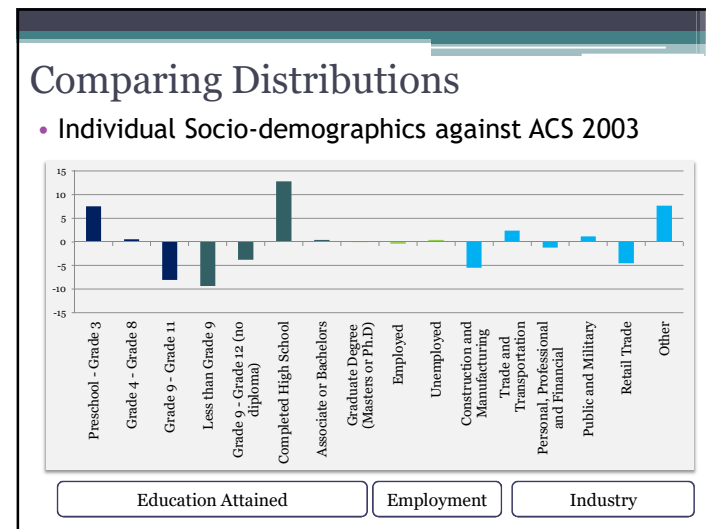
No part of this publication may be reproduced, distributed, or transmitted in any form or by any means. For permission requests, write to the owner, addressed "Attention: Permissions Coordinator," at ram.pendyala@asu.edu

CEMSELTS Application

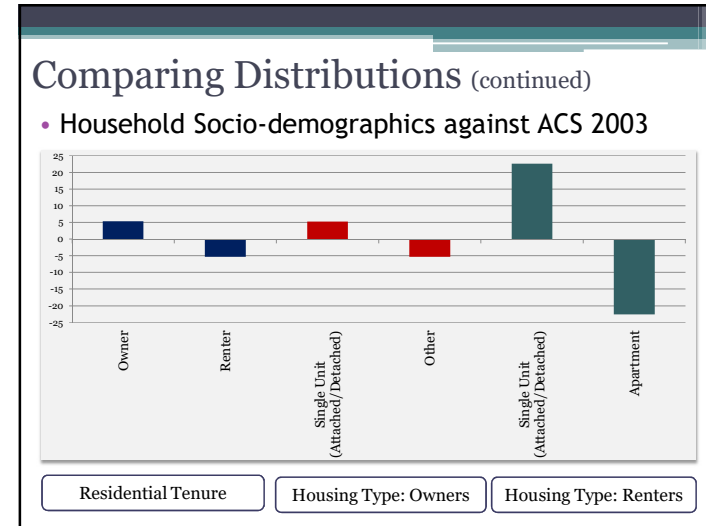
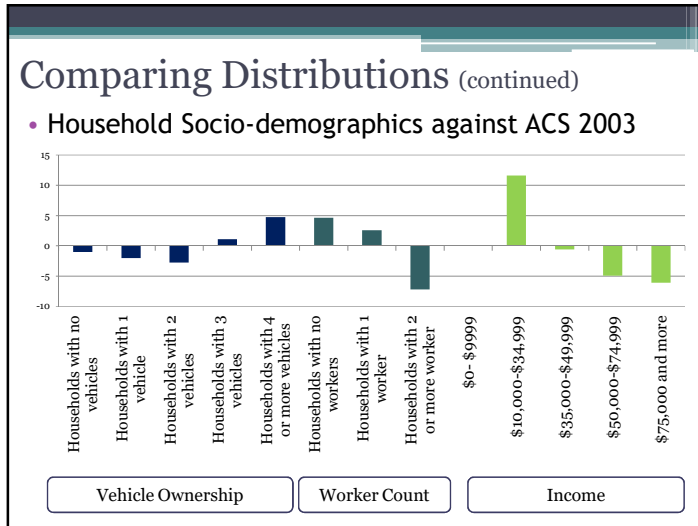
- Socio-economic model system inputs
 - Disaggregate household- and person- records obtained from PopGen
 - Drop-out rates and education attainment
 - Rich accessibility indicators generated (Chen 2011)
 - Household travel survey data for the region
 - California vehicle survey data
 - Data from ACS 2003 used for validation

Household Socio-demographics	Model Component	Percentage Values				
		ACS 2003	Census 2000	Predicted	Difference in Percentage (ACS 2003)	Difference in Percentage (Census)
Number of Vehicles						
Households with no vehicles	CEMSELTS	8.29	10.07	7.27	-1.02	-2.79
Households with 1 vehicle		33.34	34.85	31.32	-2.02	-3.55
Households with 2 vehicles		37.48	37.16	34.71	-2.77	-2.44
Households with 3 vehicles		14.10	12.59	15.17	1.07	2.59
Households with 4 or more vehicles		6.79	5.33	11.52	4.74	6.19
Number of Workers						
Households with no workers	CEMSELTS	12.21	11.31	16.84	4.63	5.54
Households with 1 worker		34.23	32.98	36.80	2.58	3.81
Households with 2 or more worker		53.57	55.71	46.36	-7.21	-9.35
Household Income						
\$0- \$9999	CEMSELTS	8.08	8.98	8.09	0.01	-0.90
\$10,000-\$24,999		17.72	17.86	26.93	9.21	9.07
\$25,000-\$34,999		11.13	11.70	13.52	2.39	1.82
\$35,000-\$49,999		15.05	15.24	14.47	-0.58	-0.77
\$50,000-\$74,999		18.53	18.89	13.58	-4.95	-5.31
\$75,000-\$99,999		11.27	11.16	7.54	-3.73	-3.62
\$100,000-\$149,999		11.08	9.78	7.28	-3.80	-2.49
\$150,000 or more		7.14	6.38	8.58	1.44	2.20
Household Tenure						
Owner	CEMSELTS	55.74	54.78	61.05	5.30	6.25
Renter		44.26	45.22	38.95	-5.30	-6.25

Individual Socio-demographics	Model Component	Percentage Values				
		ACS 2003	Census 2000	Predicted	Difference in Percentage (ACS 2003)	Difference in Percentage (Census 2000)
Enrollment of Children (aged 3 to 17 years)						
Preschool - Grade 3	CEMSELTS	37.07	41.17	44.59	7.52	3.42
Grade 4 - Grade 8		41.64	38.76	42.16	0.52	3.4
Grade 9 - Grade 11		21.29	20.07	13.25	-8.04	-6.82
Educational Attainment (18 years and above)						
Less than Grade 9	CEMSELTS	11.58	13.14	2.23	-9.35	-10.91
Grade 9 - Grade 12 (no diploma)		12.05	14.71	8.28	-3.77	-6.43
Completed High School		45.70	44.00	58.48	12.78	14.48
Associate or Bachelors		22.55	20.77	22.95	0.4	2.18
Graduate Degree (Masters or Ph.D)		8.12	7.37	8.06	-0.06	0.69
Labor Participation						
Employed	CEMSELTS	59.47	56.81	59.07	-0.40	2.26
Unemployed		40.53	43.19	40.93	0.40	-2.26
Employment Industry						
Construction and Manufacturing	CEMSELTS	19.92	20.67	14.46	-5.46	-6.21
Trade and Transportation		4.94	4.86	7.32	2.38	2.46
Personal, Professional and Financial		50.63	49.34	49.42	-1.21	0.08
Public and Military	CEMSELTS	3.94	4.04	5.07	1.13	1.03
Retail Trade		15.29	15.60	10.77	-4.52	-4.83
Other		5.28	5.49	12.96	7.68	7.47

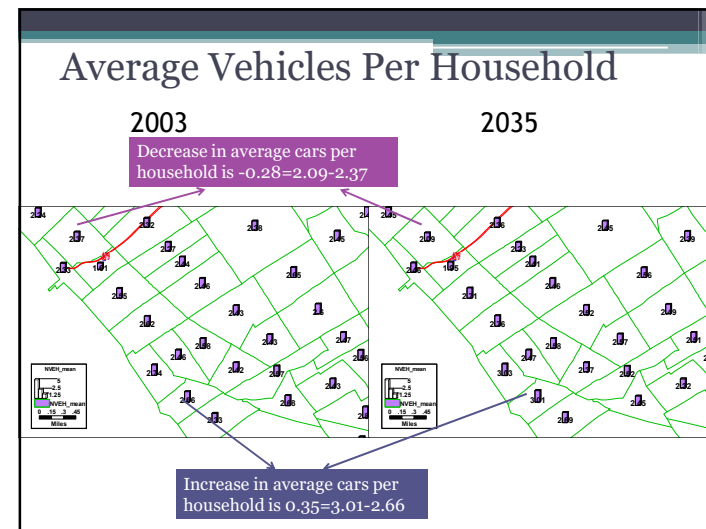


No part of this publication may be reproduced, distributed, or transmitted in any form or by any means. For permission requests, write to the owner, addressed "Attention: Permissions Coordinator," at ram.pendyala@asu.edu



Comparing Work Flow Distribution

Origin County	Destination County														Total			
	Imperial		Los Angeles		Orange		Riverside		San Bernardino		Ventura		Total		Total			
	Census 2000 (%)	CEMSELTS 2003 (%)	Census 2000 (%)	CEMSELTS 2003 (%)	Census 2000 (%)	CEMSELTS 2003 (%)	Census 2000 (%)	CEMSELTS 2003 (%)	Census 2000 (%)	CEMSELTS 2003 (%)	Census 2000 (%)	CEMSELTS 2003 (%)	Census 2000 (%)	CEMSELTS 2003 (%)	Census 2000 (%)	CEMSELTS 2003 (%)		
Imperial	0.60	0.76	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.61	0.78	0.61	0.78	
Los Angeles	0.01	0.00	53.32	52.21	2.39	3.23	0.14	0.31	0.61	1.19	0.48	0.53	56.94	57.46	56.94	57.46	56.94	57.46
Orange	0.00	0.00	2.76	2.80	16.26	14.17	0.17	0.35	0.14	0.28	0.01	0.00	19.35	17.60	19.35	17.60	19.35	17.60
Riverside	0.01	0.00	0.55	0.23	0.77	0.21	6.22	7.59	0.90	1.39	0.00	0.00	8.45	9.43	8.45	9.43	8.45	9.43
San Bernardino	0.00	0.00	1.66	1.03	0.43	0.22	0.78	1.33	6.81	7.52	0.01	0.00	9.69	10.10	9.69	10.10	9.69	10.10
Ventura	0.00	0.00	1.02	0.99	0.01	0.00	0.00	0.00	0.00	0.00	3.93	3.64	4.97	4.63	3.93	3.64	4.97	4.63
Total	0.62	0.76	59.31	57.26	19.86	17.83	7.32	9.59	8.47	10.38	4.43	4.18	100.0	100.0	100.0	100.0	100.0	100.0



No part of this publication may be reproduced, distributed, or transmitted in any form or by any means. For permission requests, write to the owner, addressed "Attention: Permissions Coordinator," at ram.pendyala@asu.edu

29

Phase 1 Version

- Auto Ownership model → Simple MNL with 4 categories
- Implemented as part of CEMSELTS
- Limitations:
 - No vehicle fleet composition information
 - No modeling of make of each vehicle
 - No allocation of vehicles to household members
 - No allocation of vehicles to different trips

30

Key Changes for Phase 2 & 3

- Vehicle Type Choice Model → MDCEV
- Vehicle Make Model
- Primary Driver Model
- Allocation of a household vehicle to each independent tour
- Explicit vehicle type choice model for joint trips of the household
- Integration of all these models within SimAGENT

31

Vehicle Type Choice Model

- Vehicle type choice determines vehicle fleet mix; critical to energy and emissions analysis
- Bhat's Multiple Discrete-Continuous Extreme Value (MDCEV) – and its mixed variants – is capable of modeling multiple vehicle holdings, body types, fuel types, age, and use (miles) simultaneously

32

Vehicle Type Choice Model

- Log-Regression model to predict annual household mileage
- Vehicle Fleet Composition → MDCEV
- 54 Alternatives: Combination of 9 body types and 6 vintage categories
 - Body type: Sub-compact car, Compact car, Medium car, Large Car, Sports car, Medium SUV, Large SUV, Van, Pickup
 - Vintage: New or 1 year, 2-3 years, 4-5 years, 6-9 years, 10 to 12 years, >12 years
 - Plus One non-motorized mileage alternative

Vehicle Make Model

- For each body type and vintage combination predicted by MDCEV → Vehicle Make determined using MNL model
- Vehicle Make information obtained using Wards Automotive year books from 2008, 2006, 2004, 2002, 1999, 1996 and Green Vehicle Guide from EPA
- Variables in the model:
 - Dimensions of the vehicle, Horse Power, Engine Capacity, Type of wheel drive, Curb weight, Green House Gas Rating, Annual Fuel Cost, Purchase Price, Vehicle Make (Honda, Toyota, BMW etc)
 - Interactions with Household Characteristics like Size and income

Vehicle Make Alternatives

A total of 759 Makes across all vehicle type and vintage combinations

Vehicle Type/Vintage	Number of Makes
Van & New or 1 year old	9 makes
Van & 2 or 3 years old	9 makes
Van & 4 to 5 years old	8 makes
Van & 6 to 9 years old	5 makes
Van & 10-12 years old	12 makes
Van & >12 years old	9 makes
PU & New or 1 year old	5 makes
PU & 2 or 3 years old	7 makes
PU & 4 to 5 years old	6 makes
PU & 6 to 9 years old	6 makes
PU & 10-12 years old	7 makes
PU & >12 years old	6 makes

Primary Driver Model

- Multinomial Logit model → allocates each vehicle to a person (primary driver)
- Number of alternatives = Number of licensed individuals in household
- Independent variables → Interaction of several person specific variables like gender, education, employment with vehicle characteristics

Allocation of Vehicles to Tours

- Any independent vehicle tour made by a household member is assumed to be made using the individual's primary vehicle
- The vehicle used for joint tours in a household is determined using an MNL model
- Alternatives: Primary vehicles of all members participating in the joint activity
- Variables: Vehicle characteristics (Engine CC, Horse Power, vehicle body type) and their interaction with joint activity characteristics (such as number of participating people in the joint activity, distance to joint activity location, etc.)

37

Vehicle Type Choice Model Results

Body Type	Survey Data	CEMDAP
Sub-compact Car	3.5	2.7
Compact Car	18.2	23.9
Medium Car	22.3	23.9
Large Car	5.7	3.3
Sports Car	5.6	4.1
Medium SUV	9.5	9.9
Large SUV	11.0	8.9
Van	7.0	5.9
Pickup	17.2	17.3

38

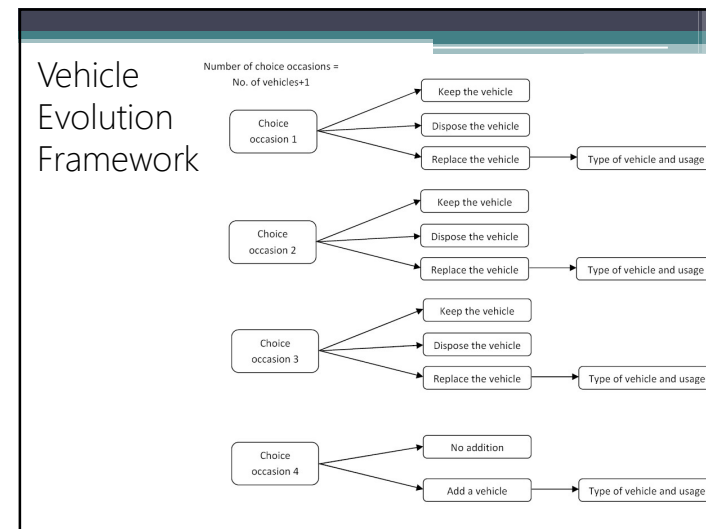
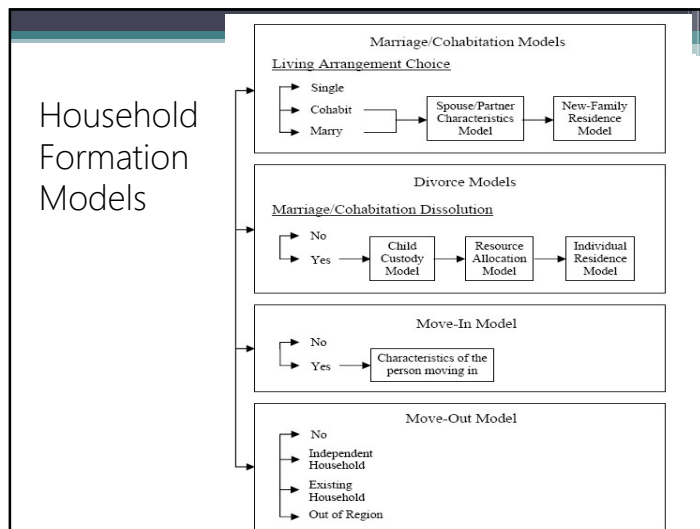
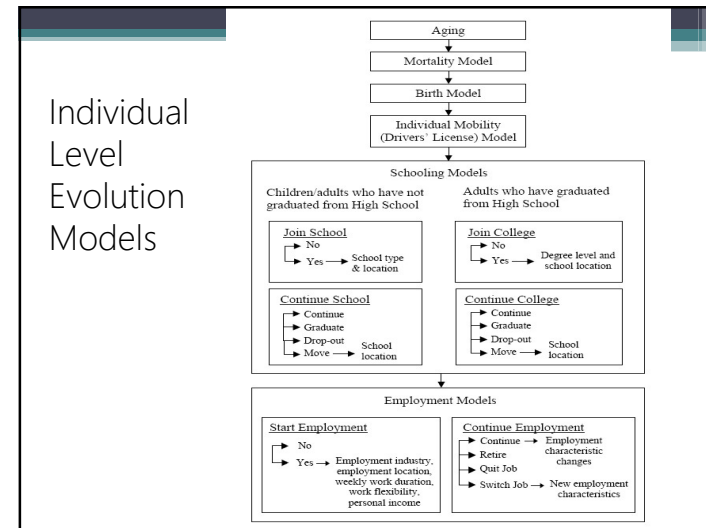
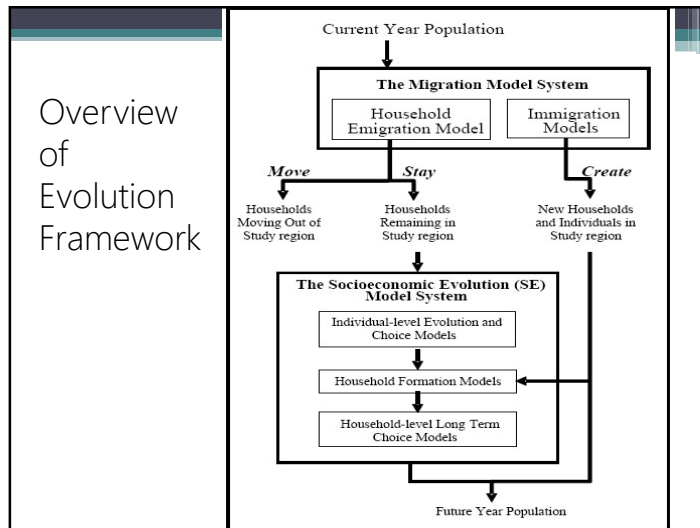
Vehicle Type Choice Model Results

Vintage Category	Survey Data	CEMDAP
New or one year	13.4	11.4
2 to 3 years old	15.9	13.5
4 to 5 years old	15.1	12.5
6 to 9 years old	26.4	27.7
10 to 12 years old	11.2	12.2
> 12 years	18.0	22.7

- ### Conclusions
- Demonstrates the applicability of a socio-demographic model system to generate inputs for use in activity-based model systems
 - Reasonableness of generated attributes
 - Implications for policy and forecasting
 - Scalability of framework
 - Applied to a large model region - SCAG (~17 million people)
 - Application of state-of-the-art econometric formulations - MDCEV for simulating choice dimensions
 - The accuracy of travel forecasts affected by population synthesis process

- ### Household Evolution
- Progresses the resident population year after year using smooth transitions instead of abrupt adjustments based on externally provided demographic data
 - First application in the US
 - Enables to link demographic transition to behavioral change
 - Can demonstrate the market penetration of new technologies and adoption of new behavioral patterns in a realistic and verifiable way

No part of this publication may be reproduced, distributed, or transmitted in any form or by any means. For permission requests, write to the owner, addressed "Attention: Permissions Coordinator," at ram.pendyala@asu.edu



No part of this publication may be reproduced, distributed, or transmitted in any form or by any means. For permission requests, write to the owner, addressed "Attention: Permissions Coordinator," at ram.pendyala@asu.edu