SOUTHERN CALIFORNIA ASSOCIATION of GOVERNMENTS

Evaluation of Census 2010 Block Level Data

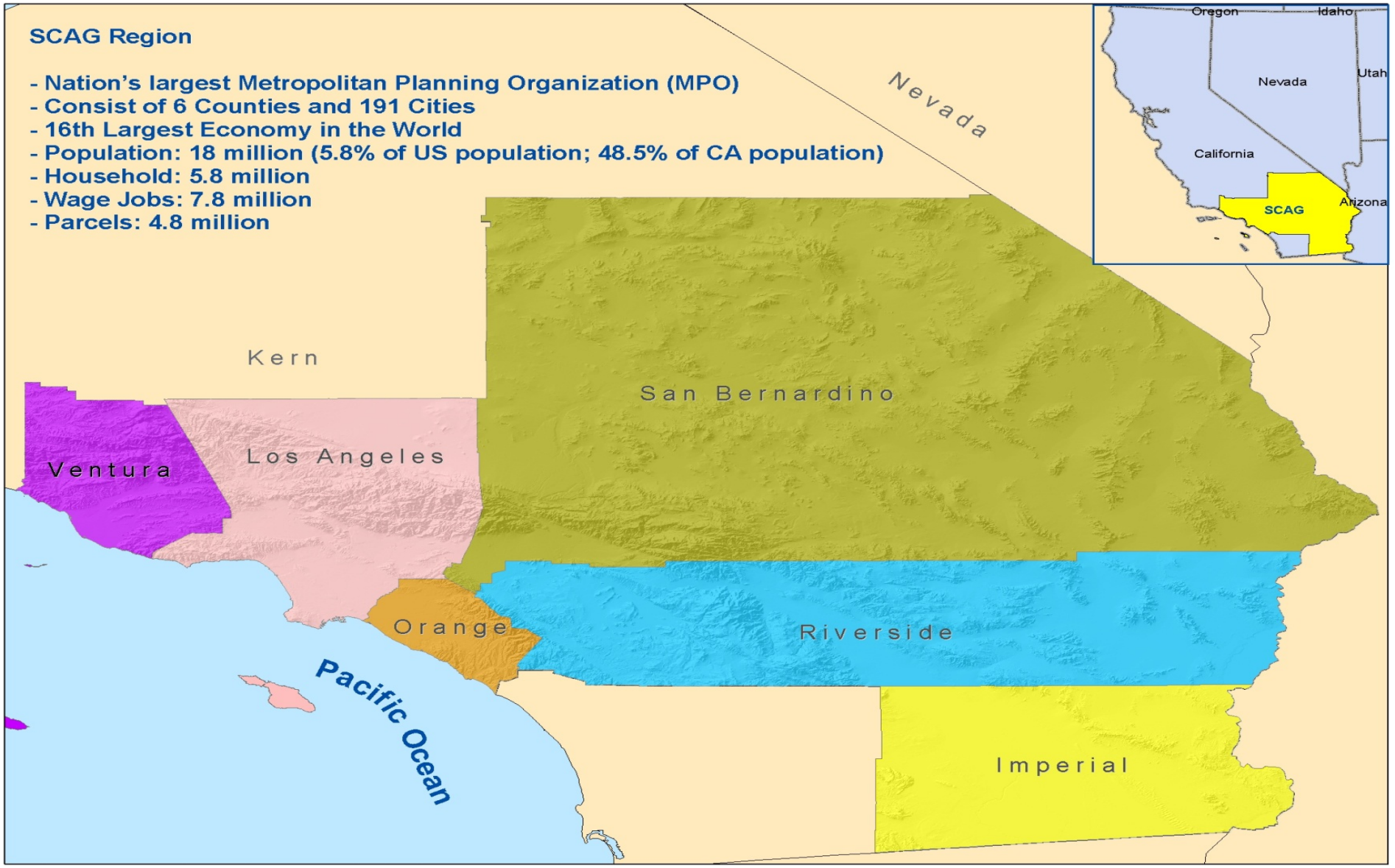
Sung Ho Ryu, Cheol-Ho Lee, Jung H. Seo, Simon Choi, SeongHee Min

2012 ESRI International User Conference
July 26, 2012
San Diego, California

Contents

- About SCAG and Small Area Forecasting
- Importance of Census 2010 Data
- Evaluation of Census 2010 Data with GIS
- Census Count Question Resolution Program

About SCAG


SCAG Region
- Nation's largest Metropolitan Planning Organization (MPO)
- Consist of 6 Counties and 191 Cities
- 16th Largest Economy in the World
- Population: 18 million (5.8% of US population; 48.5% of CA population)
- Household: 5.8 million
- Wage Jobs: 7.8 million
- Parcels: 4.8 million
Nevada
Kern
San Bernardino
Ventura
Los Angeles
Orange
Riverside
Pacific Ocean
Imperial
Oregon
Idaho
Nevada
Utah
California
SCAG
Arizona

Small Area Forecasting

- Small Area: Below City (Census Tracts / BG / Block, TAZ, Parcel, MPU)
- Developing Reasonable, Consistent, Acceptable, and Collaborative Small Area Data Set
- Two Key relationships
  - Geographical relationship (Aggregation and Disaggregation)
  - Temporal relationship (Monotonous vs. Fluctuation)

Importance of Census Data

- Key Data Source (Best Reference)
- The number of seats in Congress
- The allocation of federal fund.
- Public and Private sectors use census data to make important decisions
- Census Block data is the base of small area data development

Census 2010 Block Data

- Redistricting Data Summary File (Public Law 94-171/PL94)
  - Released March, 2011

    (254,085 blocks in the SCAG region)
  - 5 Major Variables:
    - Population by Race
    - Population by Hispanic Origin
    - Population over 18 by Race
    - Population over 18 by Hispanic Origin
    - Occupancy Status

Census 2010 Block Data - cont.

- Table P-42 (group quarters information) from the 2010 Census Summary File 1 (GQ data)
  - Released April, 2011

    (8,495 blocks in SCAG region)
  - 10 variables: Group Quarters Population by Group Quarters Type

Census 2010 Block Data - cont.

- 2010 Census Summary File 1
  - Released June, 2011
  - 331 variables: Race, Hispanic Origin, Age, Sex, Household Type, Household Size, Family Type, Family Size, Group Quarters, Occupancy Status, Householder, ETC

RTP08 vs. RTP12

- 2008 Regional Transportation Plan
  - Based on DOF data
- 2012 Regional Transportation Plan
  - Based on Census data (2011)

DOF Data vs. PL94 Data

County	POP08 DOF	POP08 PL94	HH08 DOF	HH08 PL94	POP Diff.	HH Diff.	POP % Diff.	HH % Diff.
Imperial	177,441	170,003	49,791	48,611	-7,438	-1,180	-4.38%	-2.43%
Los Angeles	10,347,642	9,778,470	3,267,616	3,227,833	-569,172	-39,783	-5.82%	-1.23%
Orange	3,123,058	2,989,426	1,001,255	987,104	-133,632	-14,151	-4.47%	-1.43%
Riverside	2,090,235	2,128,306	675,459	678,767	38,071	3,308	1.79%	0.49%
San Bernardino	2,052,929	2,015,862	608,022	605,918	-37,067	-2,104	-1.84%	-0.35%
Ventura	831,676	813,039	267,665	265,517	-18,637	-2,148	-2.29%	-0.81%
SCAG	18,622,981	17,895,106	5,869,808	5,813,750	**-727,875**	**-56,058**	**-4.07%**	**-0.96%**

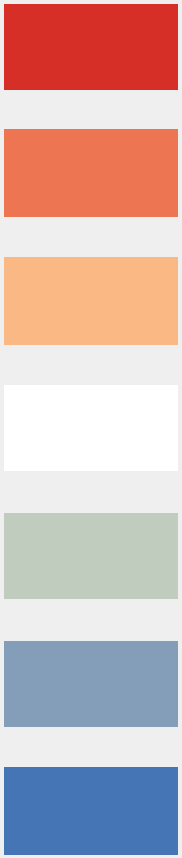
Compare SED at TAZ (DOF vs. PL94)

- SCAG TAZ: consists of 4,109 Tier1 zones and 11,267 Tier2 zones
- Symmetric Mean Absolute Percentage Error (sMAPE)

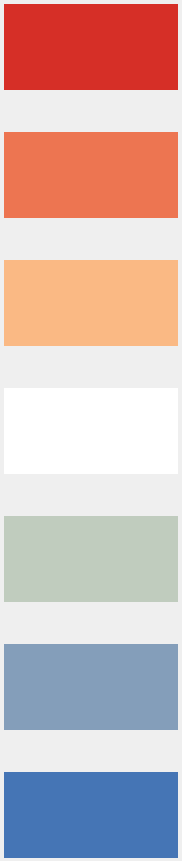
$$\text{sMAPE} = \frac{1}{n}\sum_{t=1}^{n}\frac{|At - Ft|}{At + Ft}$$

	Population 2008	Household 2008
sMAPE	0.1125	0.0873

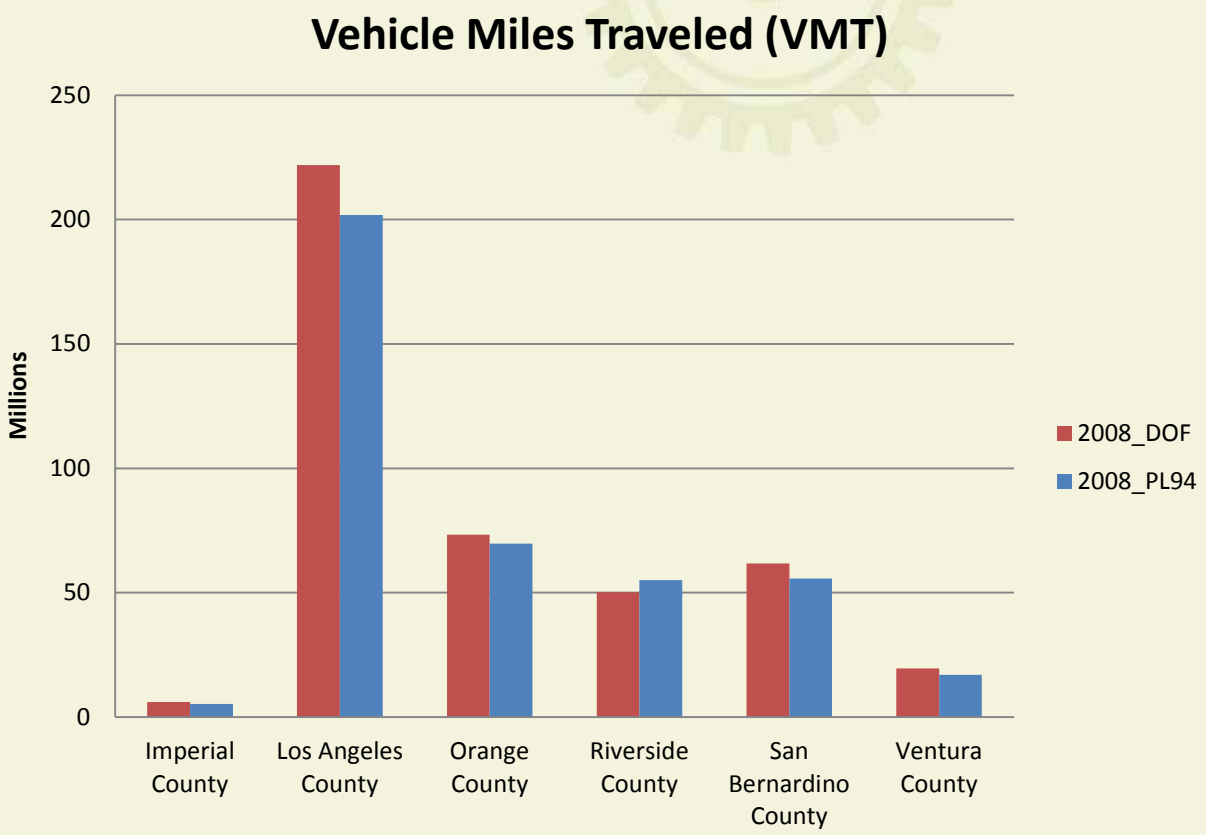
Difference in Households at TAZ



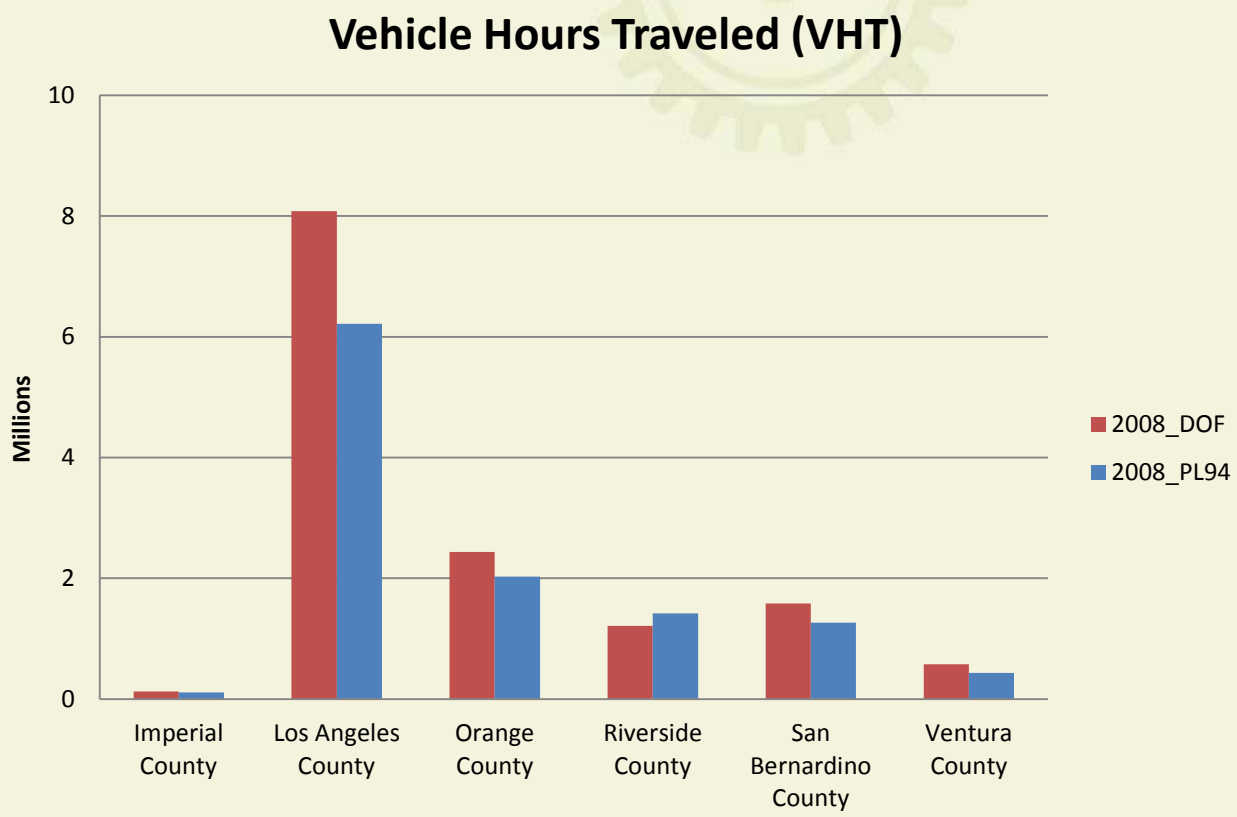
Difference in Populations at TAZ



Difference in VMT (DOF vs. PL94)



Difference in VHT (DOF vs. PL94)



Method for Finding Problematic Blocks

- Select blocks with boundaries that intersect freeways, ramps, or major roads.
- From the previous step, select blocks that have housing units or group quarter population.
- Then identify blocks with areas less than 4,000 square meters. Then identify blocks where the ratio of area and housing units (area/housing units) is less than 600.
- We observed about 800 blocks with possible errors and *manually verified* the presence of problematic data by checking aerial imagery.

Types of Census 2010 Data Error

- Housing units in blocks that are entirely located within freeways, ramps, or major roads.
- Housing units in blocks that are entirely located within water bodies.
- Group quarters population in blocks that are entirely located within freeways, ramps, or major roads.
- Group quarter institutions that has been located in a wrong census block.

Problematic Blocks

County	Number of Blocks by Type			Total
	Road	Water	GQ	
Imperial County	1	-	3	4
Los Angeles County	204	3	83	290
Orange County	76	-	7	83
Riverside County	5	-	2	7
San Bernardino County	4	-	4	8
Ventura County	3	-	1	4
SCAG Region	293	3	100	396

Blocks on the Roads

County	SubRegion	City Name	Block ID		Raw Population	Raw Housing	Raw Household	Fixed Population	Fixed Housing	Fixed Household
037	Las Virgenes	Agoura Hills	Problem	060378003273007	23	10	9	0	0	0
			Proposed	060378003271000	1715	810	757	1738	820	766

Blocks on the Roads

County	SubRegion	City Name	Block ID		Raw Population	Raw Housing	Raw Household	Fixed Population	Fixed Housing	Fixed Household
037	City of Los Angeles	Los Angeles	Problem	060371132371022	275	125	123	0	0	0
			Proposed	060371132371006	5	3	2	280	128	125

Problematic Blocks

County	Number of Blocks by Type			Total
	Road	Water	GQ	
Imperial County	1	-	3	4
Los Angeles County	204	3	83	290
Orange County	76	-	7	83
Riverside County	5	-	2	7
San Bernardino County	4	-	4	8
Ventura County	3	-	1	4
SCAG Region	293	3	100	396

Blocks on Water Bodies

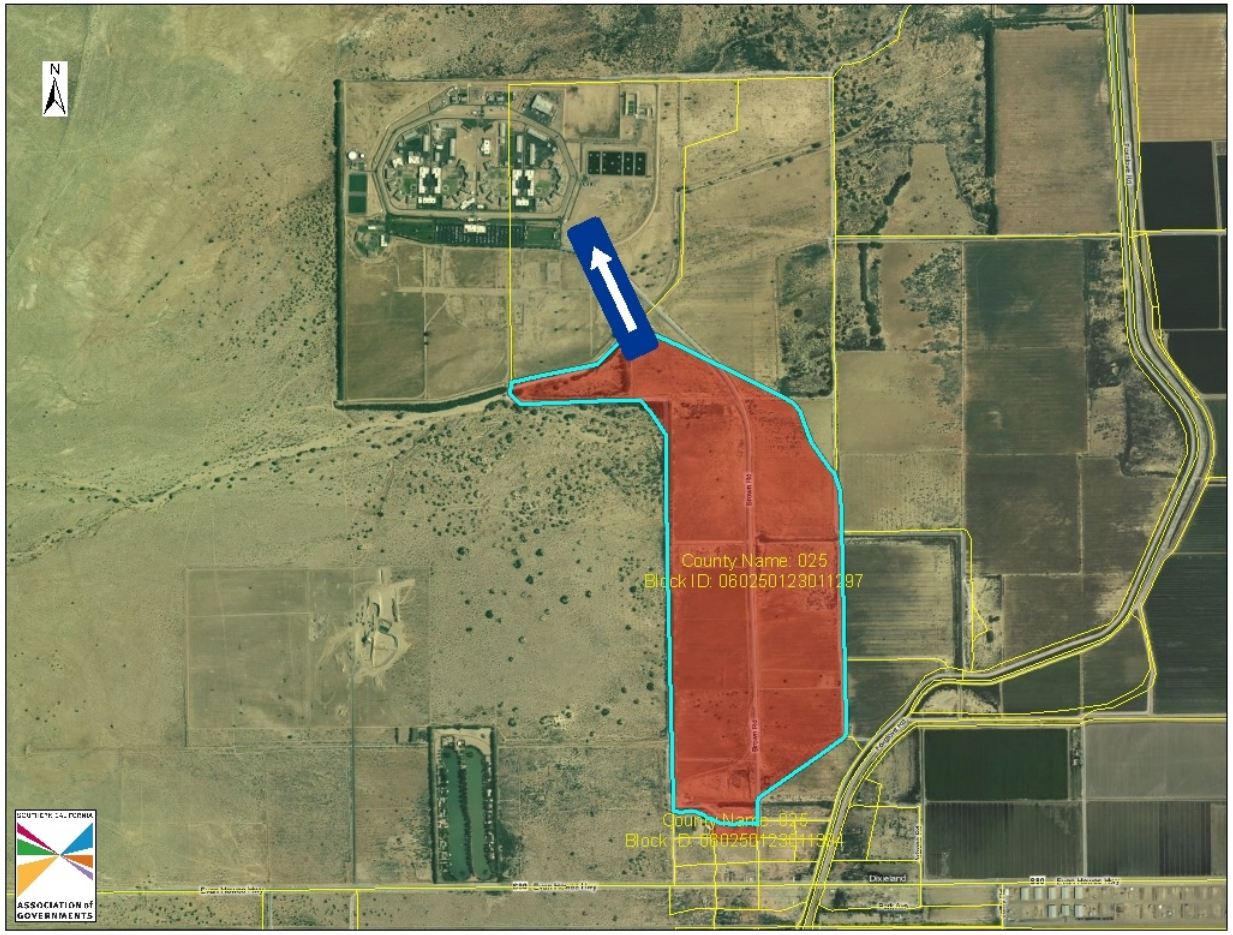
County	SubRegion	City Name	Block ID		Raw Population	Raw Housing	Raw Household	Fixed Population	Fixed Housing	Fixed Household
037	City of Los Angeles	Los Angeles	Problem	060372742021044	11	5	5	0	0	0
			Proposed	060372742021001	3	2	2	14	7	7

Problematic Blocks

County	Number of Blocks by Type			Total
	Road	Water	GQ	
Imperial County	1	-	3	4
Los Angeles County	204	3	83	290
Orange County	76	-	7	83
Riverside County	5	-	2	7
San Bernardino County	4	-	4	8
Ventura County	3	-	1	4
SCAG Region	293	3	100	396

GQ Block

County	Block ID		Raw Population	Raw Housing	Raw Household	Raw GQ	Fixed Population	Fixed Housing	Fixed Household	Fixed GQ
25	Problem	060250123011297	4311	0	0	4311	0	0	0	0
	Proposed	060250123011299	0	0	0	0	4311	0	0	4311



GQ Block

County	Block ID		Raw Population	Raw Housing	Raw Household	Raw GQ	Fixed Population	Fixed Housing	Fixed Household	Fixed GQ
59	Problem	060591100141000	209	4	3	203	0	0	0	0
	Proposed	060591101082016	0	0	0	0	209	4	3	203

CQR Program

- 2010 Census Count Question Resolution
- A mechanism for State, local and Tribal officials of governmental units in the U.S. and Puerto Rico to challenge the 2010 Census population counts of housing units and group quarters.
- The Count Question Resolution (CQR) Program will address questions about housing unit and group quarters counts for three types of challenges for the 2010 Decennial Census: (1) boundary, (2) geocoding, and (3) coverage.

CQR Program – cont.

- The CQR Program will accept CQR challenges beginning June 1, 2011 and ending with challenges that are received and postmarked, faxed, or emailed as of June 1, 2013.
- All functioning governmental units such as States, Counties, and Cities can challenge the 2010 Census counts.
- http://2010.census.gov/2010census/about/cqr.php

Thank You !

SOUTHERN CALIFORNIA
ASSOCIATION of GOVERNMENTS
www.scag.ca.gov