# DATA SOURCES FOR MODEL DEVELOPMENT AND VALIDATION

# HOR







Jason Carbee, AICP jason.carbee@hdrinc.com 402.399.1370

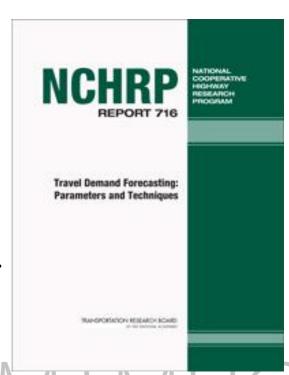
#### Presentation Agenda

- Back to Basics!
- Open Discussion
- Topics:
  - Brief Overview of NCHRP 716
  - Data Sources / Validation Techniques
    - Socio-Economic Data
    - Trip Generation
    - Trip Distribution
  - Combing Data Sources Where Appropriate

### NCHRP 716 – WHAT'S DIFFERENT?

#### What's New?

- Acknowledges Increased Complexity:
  - Vehicle Availability Models
  - ABM, Tour-Based, etc.
- Discussion of Best Practices
- NHTS 2009 for Default Parameters
- Database of 69 MPOs' parameters
- NCHRP 735 supplements for Long-Distance and Rural Travel



#### NCHRP 716 Data Source Discussion

Table 3.5. Example primary and secondary model validation tests.

Model Component	Primary Tests	Secondary Tests	Potential Validation Data Sources
Network topology, including balance between roadway network detail and zone detail		Intrazonal travel distances (model design issue)  Zone structure compatibility with transit analysis needs (model design issue)  Final quality control checks based on review by end users  Transit paths by mode on selected interchanges	GIS center line files     Transit on-board or household survey data
Socioeconomic Data/Models	Households by income or auto ownership Jobs by employment sector by geographic location Locations of special generators Qualitative logic test on growth Population by geographic area Types and locations of group quarters Frequency distribution of households and jobs (or household and job densities) by TAZ	Dwelling units by geographic location or jurisdiction     Households and population by land use type and land use density categories     Historical zonal data trends and projections to identify "large" changes (e.g., in autos/ household from 1995 to 2005)	Census SF-3 data QCEW Private sources, such as Dun & Bradstreet
Trip Generation	Reasonableness check of trip rates versus other areas Logic check of trip rate relationships	Checks on proportions or rates of nonmotorized trips     Reasonableness check of tour rates     Cordon lines by homogeneous land use type	Chapter 4 of this report Traffic counts (or intercept survey data) for cordon lines Historic household survey data for region NHTS (2001 or 2009)
Trip Distribution	Trip length frequency distributions (time and distance) by market segments Worker flows by district District-to-district flows/desire lines Intrazonal trips External station volumes by vehicle class	Area biases (psychological barrier—e.g., river)     Use of k-factors (Design Issue)     Comparison to roadside intercept origin-destination surveys     Small market movements     Special groups/markets     Balancing methods	ACS/CTPP data Chapter 4 of this report Traffic counts (or intercept survey data) for screenlines Historic household survey data for region NHTS (2001 or 2009)

#### NCHRP 716 Data Source Discussion

Table 3.5. Example primary and secondary model validation tests.

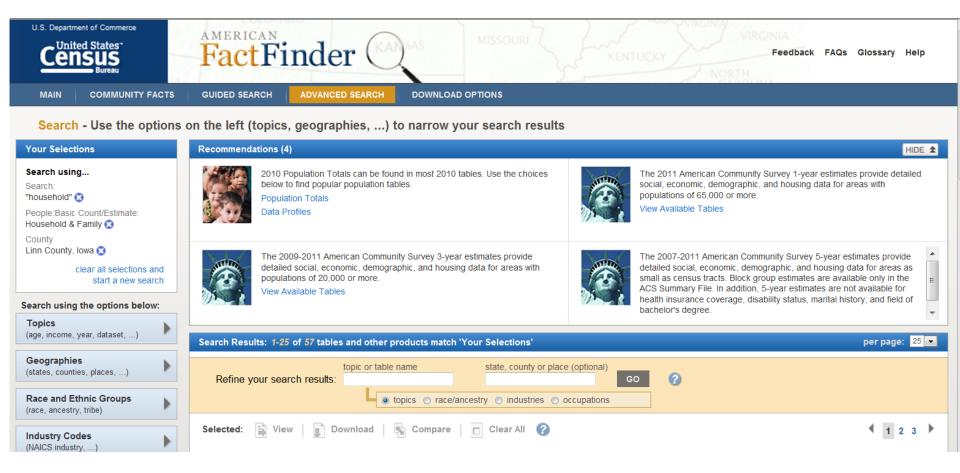
Model Component	Primary Tests	Secondary Tests	Potential Validation Data Sources
Networks/Zones	Correct distances on links Network topology, including balance between roadway network detail and zone detail Appropriateness of zone size given spatial distribution of population and employment Network attributes (managed lanes, area types, speeds, capacities) Network connectivity Transit run times	Intrazonal travel distances (model design issue) Zone structure compatibility with transit analysis needs (model design issue) Final quality control checks based on review by end users Transit paths by mode on selected interchanges	GIS center line files     Transit on-board or household survey data
Socioeconomic Data/Models	Households by income or auto ownership     Jobs by employment sector by geographic location     Locations of special generators     Qualitative logic test on growth     Population by geographic area     Types and locations of group quarters     Frequency distribution of households and jobs (or household and job densities) by TAZ	Dwelling units by geographic location or jurisdiction     Households and population by land use type and land use density categories     Historical zonal data trends and projections to identify "large" changes (e.g., in autos/ household from 1995 to 2005)	Census SF-3 data  QCEW Private sources, such as Dun & Bradstreet
Trip Generation	Reasonableness check of trip rates versus other areas     Logic check of trip rate relationships	Checks on proportions or rates of nonmotorized trips     Reasonableness check of tour rates     Cordon lines by homogeneous land use type	Chapter 4 of this report     Traffic counts (or intercept survey data) for cordon lines     Historic household survey data for region     NHTS (2001 or 2009)
Trip Distribution	Trip length frequency distributions (time and distance) by market segments Worker flows by district District-to-district flows/desire lines Intrazonal trips External station volumes by vehicle class	Area biases (psychological barrier-e.g., river)     Use of k-factors (Design Issue)     Comparison to roadside intercept origin-destination surveys     Small market movements     Special groups/markets     Balancing methods	ACS/CTPP data Chapter 4 of this report Traffic counts (or intercept survey data) for screenlines Historic household survey data for region NHTS (2001 or 2009)

# DATA SOURCES / VALIDATION APPROACHES

### Data Source Examples

- SE Data Development / Checking
  - Control Total Data Sources
  - TAZ-level Data Development / Sources
- Trip Generation
  - Data Sources
  - Validation Methods
- Trip Distribution
  - Data Sources

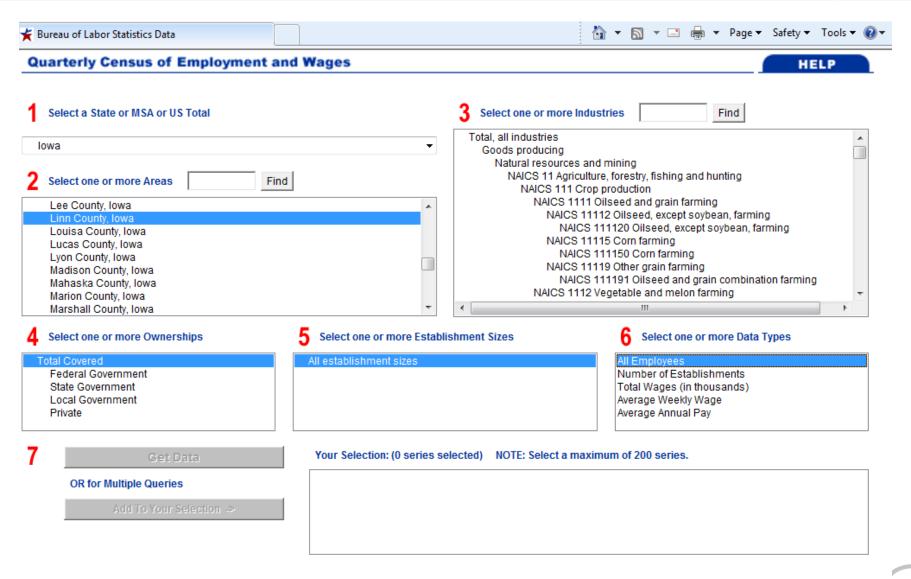
# Census / ACS Data



#### **QCEW Data**

- Quarterly Census of Employment and Wages (QCEW).
  - Was ES-202.
  - Employment and Wage Information at 6-digit NAICS level.
  - Workers Covered by State Unemployment Insurance (UI) laws and Federal Unemployment Compensation for Federal Employees (UCFE) program.
  - Aggregated to County Level.

# QCEW Online Tool – Linn County



# QCEW Example – Linn County

#### **Quarterly Census of Employment and Wages**

Series Id: ENU1911310010

State: Iowa

Area: Linn County, Iowa
Industry: Total, all industries

Owner: Total Covered

Size: All establishment sizes

Type: All Employees

#### Download: .xls

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
2002	115698	115163	115877	117556	118263	118659	115514	115892	116581	117387	118300	117438	116861
2003	113544	113059	114007	115085	115569	115988	114035	114441	115186	115995	116120	115921	114913
2004	113621	113506	114407	115736	116164	117262	115371	115540	116231	117623	117864	118441	115981
2005	115145	115852	117417	118547	118553	119269	116810	116726	118601	118859	119696	119639	117926
2006	116787	117376	118838	120058	120846	122650	119488	119907	121086	121997	122703	122806	120379
2007	120068	120641	121142	122755	124224	126155	122621	122926	123850	124765	125911	125735	123399
2008	123038	123364	124168	125435	126808	127362	125267	125298	126485	127530	128001	127319	125840
2009	123518	123965	123786	124593	124668	125406	122603	121979	123190	123968	124213	123777	123806
2010	120917	120851	121684	123606	124704	125876	123905	123858	124943	126041	126799	126586	124148
2011	121870	122418	123182	125068	125985	127525	124213	124136	125840	125962	126897	127024	125010
2012	124418(P)	124130(P)	125299(P)	126392(P)	127543(P)	129013(P)	124943(P)	125550(P)	126645(P)				

# QCEW Example – Linn County

#### **Quarterly Census of Employment and Wages**

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
2002	115698	115163	115877	117556	118263	118659	115514	115892	116581	117387	118300	117438	116861
2003	113544	113059	114007	115085	115569	115988	114035	114441	115186	115995	116120	115921	114913
2004	113621	113506	114407	115736	116164	117262	115371	115540	116231	117623	117864	118441	115981
2005	115145	115852	117417	118547	118553	119269	116810	116726	118601	118859	119696	119639	117926
2006	116787	117376	118838	120058	120846	122650	119488	119907	121086	121997	122703	122806	120379
2007	120068	120641	121142	122755	124224	126155	122621	122926	123850	124765	125911	125735	123399
2008	123038	123364	124168	125435	126808	127362	125267	125298	126485	127530	128001	127319	125840
2009	123518	123965	123786	124593	124668	125406	122603	121979	123190	123968	124213	123777	123806
2010	120917	120851	121684	123606	124704	125876	123905	123858	124943	126041	126799	126586	124148
2011	121870	122418	123182	125068	125985	127525	124213	124136	125840	125962	126897	127024	125010
2012	124418(P)	124130(P)	125299(P)	126392(P)	127543(P)	129013(P)	124943(P)	125550(P)	126645(P)				

Table 2-4: 2005 Employment by Type (Corridor MPO)

Employment Type	Employees (2005, Corridor MPO)	Percent
Retail	20,294	18.1%
Service	54,559	48.6%
Basic	37,494	33.4%
Total	112,347	100.0%

QCEW Available at Industry Level

### Woods and Poole Projections

- One of Multiple Private Economic Projections Sources
- Econometric Model for Population, Household, Employment (by Sector)
- Data Available at County / MSA Level
- Historical and Projections, 1969 to 2040

# Woods and Poole Example

	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028
TOTAL POPULATION (in thousands)	24.445															
TOTAL POPULATION AGE 0 to 19 YEARS (in thousands)	7.733															
TOTAL POPULATION AGE 20 to 64 YEARS (in thousands)	13.178															
TOTAL POPULATION AGE 65 YEARS and OVER (in thousands)	3.534															
WHITE NON-HISPANIC POPULATION (in thousands)	15.117															
BLACK NON-HISPANIC POPULATION (in thousands)	0.813															
NATIVE AMERICAN NON-HISPANIC POPULATION (in thousands)	0.093															
ASIAN AMERICAN and PACIFIC ISLANDER NON-HISPANIC POPULATION (in thousands)	0.221															
HISPANIC or LATINO POPULATION of ANY RACE (in thousands)	8.201															
TOTAL POPULATION AGE 16 YEARS and OVER (in thousands)	18.088															
TOTAL POPULATION AGE 18 YEARS and OVER (in thousands)	17.329															
TOTAL POPULATION AGE 0 to 6 YEARS (in thousands)	2.773															
TOTAL POPULATION AGE 7 to 11 YEARS (in thousands)	2.004															
TOTAL POPULATION AGE 12 to 14 YEARS (in thousands)	1.216															
TOTAL POPULATION AGE 15 to 17 YEARS (in thousands)	1.123															
TOTAL POPULATION AGE 15 to 44 YEARS (in thousands)	8.965															
TOTAL EMPLOYMENT (in thousands of jobs)	15.566															
FARM EMPLOYMENT (in thousands of jobs)	0.866															
NON-FARM EMPLOYMENT (in thousands of jobs)	14.7															
PRIVATE NON-FARM EMPLOYMENT (in thousands of jobs)	12.506															
FORESTRY, FISHING, RELATED ACTIVITIES and OTHER EMPLOYMENT (in thousands of jobs)	0.237															
MINING EMPLOYMENT (in thousands of jobs)	0.07															
UTILITIES EMPLOYMENT (in thousands of jobs)	0.017															
CONSTRUCTION EMPLOYMENT (in thousands of jobs)	0.844															
MANUFACTURING EMPLOYMENT (in thousands of jobs)	3.934															
WHOLESALE TRADE EMPLOYMENT (in thousands of jobs)	0.597															
RETAIL TRADE EMPLOYMENT (in thousands of jobs)	1.473															
TRANSPORTATION and WAREHOUSING EMPLOYMENT (in thousands of jobs)	0.521															
INFORMATION EMPLOYMENT (in thousands of jobs)	0.114															
FINANCE and INSURANCE EMPLOYMENT (in thousands of jobs)	0.519															
REAL ESTATE and RENTAL and LEASE EMPLOYMENT (in thousands of jobs)	0.294															
PROFESSIONAL and TECHNICAL SERVICES EMPLOYMENT (in thousands of jobs)	0.352															
MANAGEMENT of COMPANIES and ENTERPRISES EMPLOYMENT (in thousands of jobs)	0.167															
ADMINISTRATIVE and WASTE SERVICES EMPLOYMENT (in thousands of jobs)	0.527															
EDUCATIONAL SERVICES EMPLOYMENT (in thousands of jobs)	0.126															
HEALTH CARE and SOCIAL ASSISTANCE EMPLOYMENT (in thousands of jobs)	0.979															
ARTS, ENTERTAINMENT, and RECREATION EMPLOYMENT (in thousands of iobs)	0.13	0.133	0.135	0.137	0.139	0.142	0.144	0.147	0.149	0.151						
ACCOMMODATION and FOOD SERVICES EMPLOYMENT (in thousand Woods and Po	ole Em	ploym	ent	Retai	1 6	Service	Ot	her	Tota	76						
OTHER SERVICES. EXCEPT PUBLIC ADMINISTRATION EMPLOYMENT (										)1						
	20	)10 Cοι	unty	1,49	12	4,080	9	,476	15,04	48						
	Se	ector Sh	hare	9.9	%	27.1%	63	3.0%								
	20	)35 Coι	inty	1,85	51	5,917	11	,266	19,03	34		D.				

9.7%

**Sector Share** 

31.1%

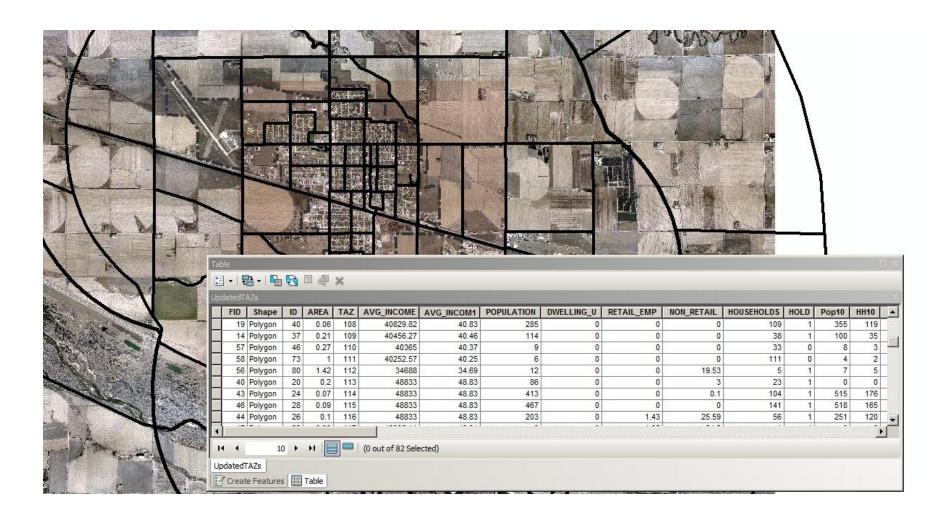
59.2%

# Extracting Sub-County Study Area

- How to Fit County Level Data to Sub-County Model Area?
  - LEHD Can Interpret Current Break Down
  - Make assumptions on future proportions

LEHD	Retail	Service	Other	Total
2010 County	1,137	3,756	6,991	11,884
2010 Study Area	709	1,715	4,313	6,737
Study Area Share of County	62%	46%	62%	57%

# Developing TAZ-Level Data



#### SE Data: TAZ – Level Households

#### Production Data

- ACS Level Population / Households:
   Block Level
- ACS Level Cross-Tab Data: Block-Group Level
  - Auto Ownership
  - Income

# **Example NCHRP 716 Trip Production Rates**

Table C.5. Home-based work trip rates.

Number of Workers by Number of Autos

		Workers							
Autos	0	1	2	3+	Average				
0	0.0	1.0	2.4	5.1	0.5				
1	0.0	1.0	2.6	5.1	0.8				
2	0.0	1.3	2.6	5.1	1.6				
3+	0.0	1.3	2.6	5.1	2.3				
Average	0.0	1.2	2.6	5.1	1.4				

Number of Persons by Number of Autos

_		Persons								
Autos	1	2	3	4	5+	Average				
0	0.2	0.7	1.0	1.0	1.0	0.5				
1	0.6	0.8	1.2	1.7	1.5	0.8				
2	0.7	1.3	2.0	2.0	2.3	1.6				
3+	0.9	1.4	2.6	2.9	3.3	2.3				
Average	0.5	1.2	2.0	2.3	2.4	1.4				

Number of Persons by Income Level

			Per	sons		
Household Income	1	2	3	4	5+	Average
i	0.2	0.6	0.8	1.3	1.8	0.6
ii	0.3	0.8	1.5	1.6	2.0	0.8
iii	0.7	1.0	1.8	2.3	2.6	1.3
iv	0.8	1.5	2.4	2.4	2.6	1.9
v	0.9	1.6	2.4	2.4	2.6	2.0
Average	0.5	1.2	2.0	2.3	2.4	1.4

Note: All averages are weighted.

Source: 2009 NHTS.

# SE Data: TAZ – Level Employment

- Attraction Data:
  - Employment by Sector

# **Example NCHRP 716 Trip Attraction Rates**

Table 4.4. Trip attraction rates from selected MPOs (person trips per unit).

	Number of				Emple	yment	
	MPO Models Summarized	Households <sup>a</sup>	School Enrollment <sup>b</sup>	Basic <sup>c</sup>	Retaild	Service <sup>e</sup>	Total
			All Person Tri	ps			
			Home-Based We	ork			
Model 1	16	•			•		1.2
	•		Home-Based Non	work			
Model 1	2	1.2	1.4	0.2	8.1	1.5	
Model 2	8	2.4	1.1		7.7	0.7	
Model 3	2	0.7		0.7	8.4	3.5	
	•	•	Nonhome Base	ed .			
Model 1	5	0.6		0.5	4.7	1.4	
Model 2	8	1.4			6.9	0.9	
		N	lotorized Person	Trips			
			Home-Based W	ork			
Model 1	8						1.2
		. 1	Home-Based Non	work			
Model 1	1	0.4	1.1	0.6	4.4	2.5	
Model 3	4	1.0		0.3	5.9	2.3	
			Nonhome Base	ed .			
Model 1	6	0.6		0.7	2.6	1.0	

The number of households in a zone.

Source: MPO Documentation Database.

b The number of elementary, high school, or college/university students in a zone.

<sup>&</sup>lt;sup>c</sup> Employment primarily in two-digit North American Industry Classification System (NAICS) codes 1–42 and 48–51 [Standard Industrial Classification (SIC) codes 1–51].

d Employment primarily in two-digit NAICS codes 44-45 (SIC codes 52-59).

Employment primarily in two-digit NAICS codes 52–92 (SIC codes 60–97).

#### Employment Data Lessons Learned

- How to Define a "Job"
  - Data Sources Count Different Jobs
    - QCEW / ES-202 Count Jobs Paying into Unemployment Insurance Pool
      - No Military Jobs
      - No Self-Employed
      - No Proprietors
      - No Railroad Jobs
    - Private Databases Count More
  - FTEs / Primary vs Secondary Jobs
- Headquarters Plant Issue
  - Multi-plant operations
  - Government
  - Schools

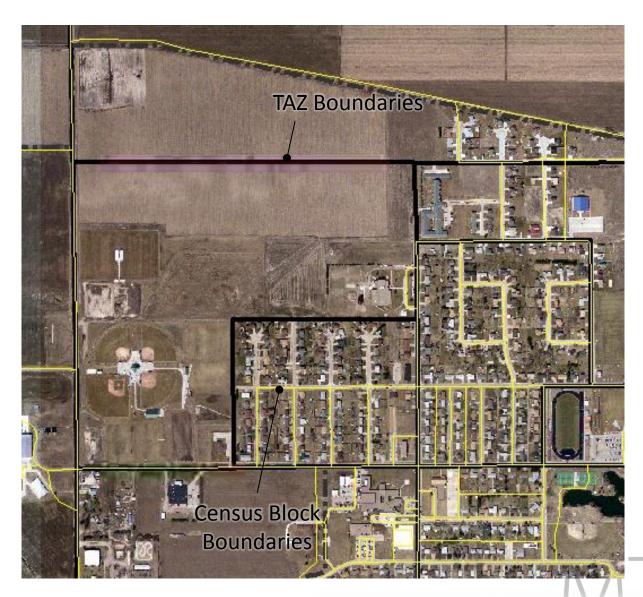
#### SE Data Allocation: Block — TAZ "Perfect Fit"



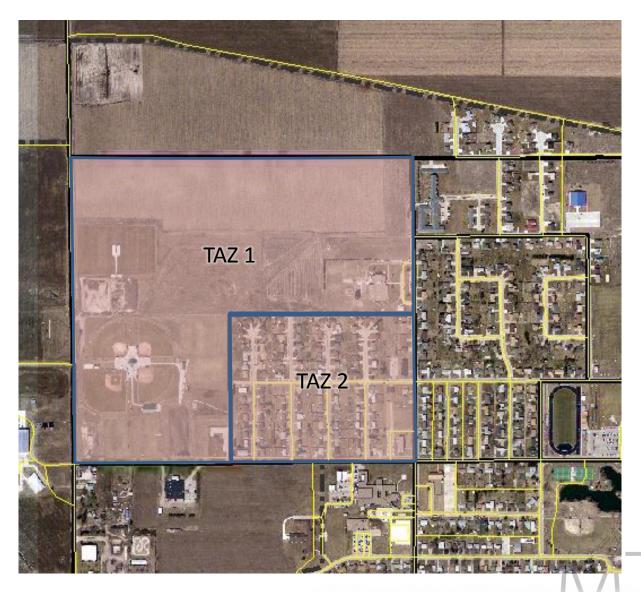
**TAZ Boundaries** 

Census Block Boundaries

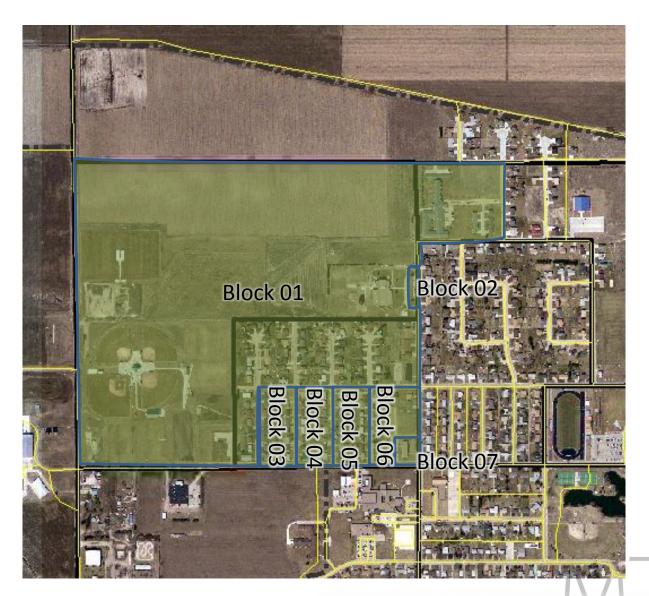
#### SE Data Allocation: Block – TAZ "Post-Processing"



#### TAZ "Post-Processing": TAZ Boundaries Example



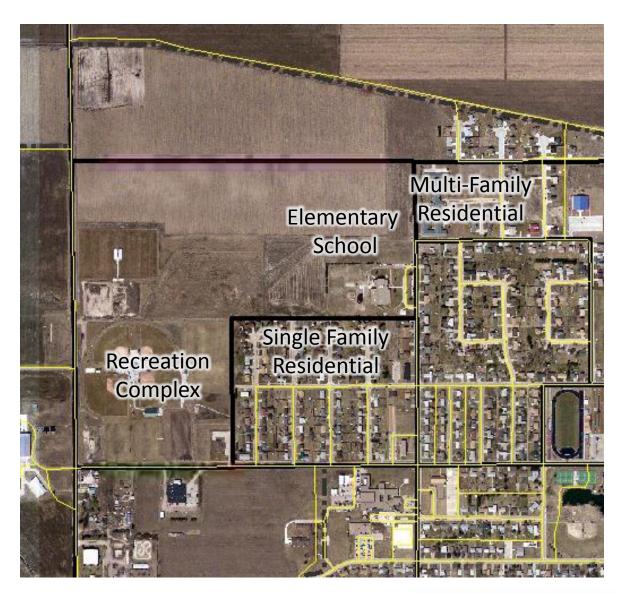
#### TAZ "Post-Processing": Block Boundaries Example



#### TAZ "Post-Processing": The Issue



#### TAZ "Post-Processing": Dig In and Get Your Hands Dirty



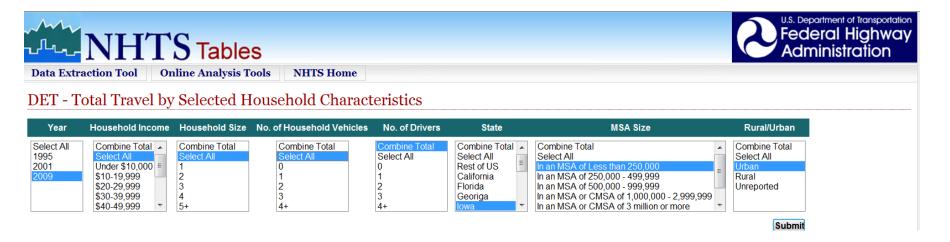
- Find Schools,
   Employers
- Count SF Housing
- Field Verify /
   Estimate MF Housing
- Know Aerial Date
- Review Online Mapping

## **Trip Generation**

- Reasonable Regional Trip Rate Average
- Evaluate Real Trip Rates from Homogenous Land Uses

#### Trip Generation

- NHTS 2009 Data for Trip Gen Rates
  - Relevant Geography
  - Relevant Cross-Classification Variables
  - By Trip Purpose



#### Trip Generation – Example NHTS Validation Check

- NHTS 2009 Trip Productions
  - lowa MSAs
  - Urban
  - Under 250,000 population
- NHTS Rate: 9.42 Person Trips / HH (AADT)
- Current Model Rate: 13.52 Person Trips / HH (AWDT)
- Refine NHTS Data / Evaluate Model Data
  - Refine NHTS data to Weekday Data
  - Is your model Peak Workday or Average Day of Week?

#### Trip Generation Validation - Macro

Table 5.7. Comparison of household trip rates.

	Daily Person Trips per Household							
Urban Area Population	NCHRP Report 187 <sup>a,c</sup> (Published 1978)	NCHRP Report 365ac (Published 1998)	2009 NHTS Data <sup>b</sup>					
50,000 to 100,000	14.1	9.2	9.1					
100,000 to 200,000	14.5	9.2	9.1					
200,000 to 500,000	11.8	9.0	9.1					
500,000 to 1,000,000	7.6	8.6	9.6					
1,000,000 to 3,000,000	7.6	8.5	9.6					
More than 3,000,000	7.6	8.5	9.6					

<sup>&</sup>lt;sup>a</sup> Trip rates are total person trips in motorized vehicles.

Source: Sosslau et al. (1978), Martin and McGuckin (1998), 2009 NHTS.

Table 5.8. Comparison of shares of trips by trip purpose.

Urbanized Area Population	Percentage of Daily Person Trips by Trip Purpose								
	NCHRP Report 187 <sup>a</sup> (Published 1978)			NCHRP Report 365 <sup>a</sup> (Published 1998)			2009 NHTS Data <sup>b</sup>		
	HBW	HBNW	NHB	HBW	HBNW	NHB	HBW	HBNW	NHB
50,000 to 100,000	16	61	23°	20°	57°	23°	15	54	31
100,000 to 200,000	20	57	23°	20°	57°	23°	15	54	31
200,000 to 500,000	20	55	25°	21°	56°	23°	15	54	31
500,000 to 1,000,000	25	54	21°	22	56°	22°	14	56	30
1,000,000 to 3,000,000	25	54	21°	22 <sup>c</sup>	56°	22°	14	56	30
More than 3,000,000	25	54	21°	22 <sup>c</sup>	56°	22°	14	56	30

Shares by purpose are based on person trips in motorized vehicles.

HBW = home-based work; HBNW = home-based nonwork; NHB = nonhome based.

Source: Sosslau et al. (1978), Martin and McGuckin (1998), 2009 NHTS.

**Trip Rates** 

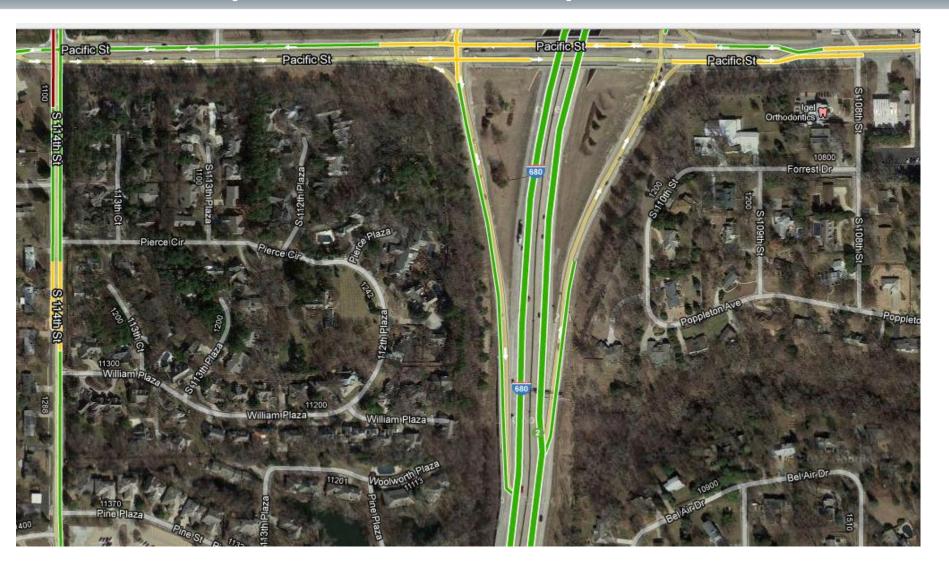
b Trip rates are total person trips by all modes.

<sup>6</sup> Because of differences between urban area categories in the three reports, the rates shown were chosen from the closest matching category. **Example NCHRP 716** 

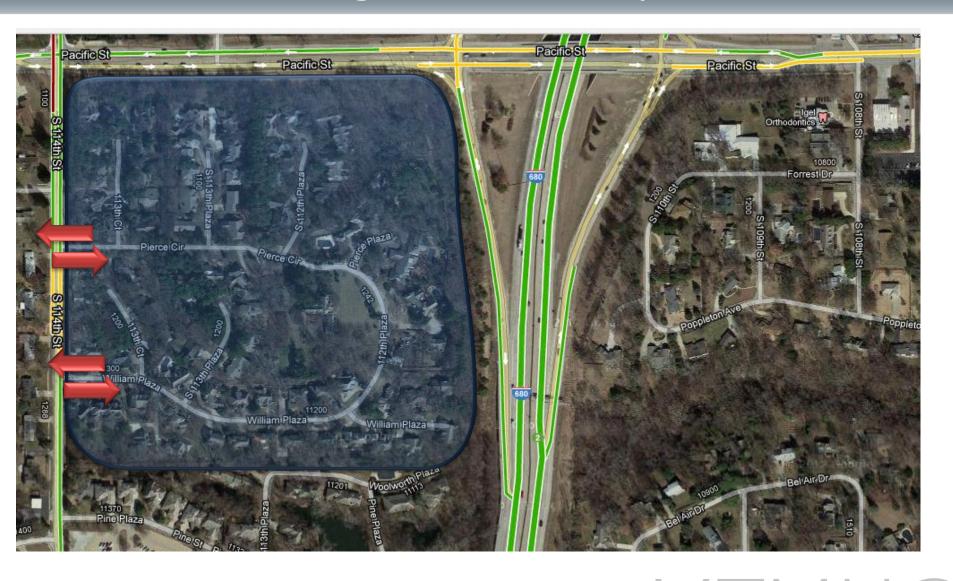
b Shares by purpose are based on person trips by all modes.

<sup>&</sup>lt;sup>c</sup> Because of differences between urban area categories in the three reports, the rates shown were chosen from the closest

# Example Cordon Trip Gen Check



#### Define Homogenous Development Area



# How Much Development?



#### How Much Traffic at Cordon?



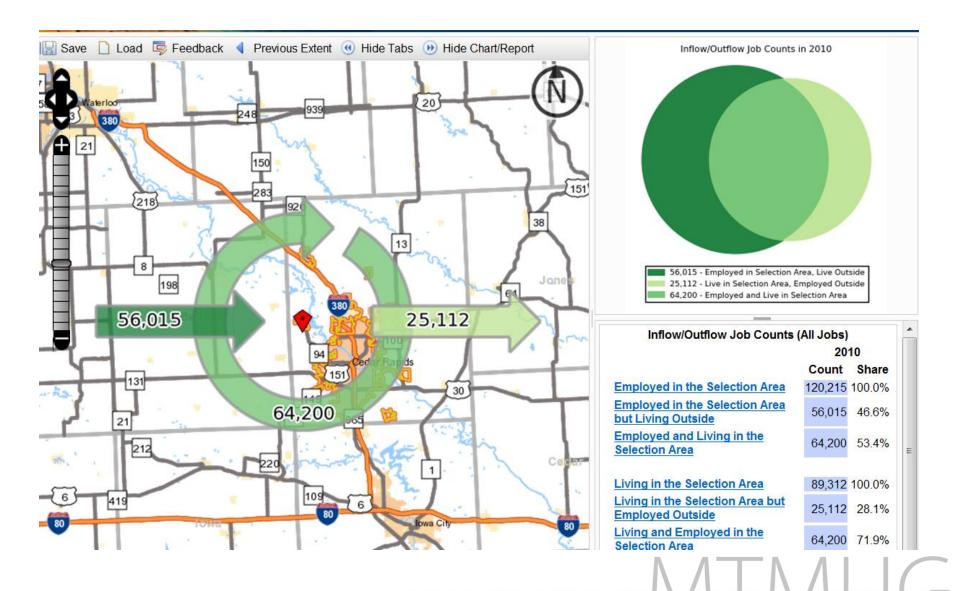
# How Does Trip Rate Compare?



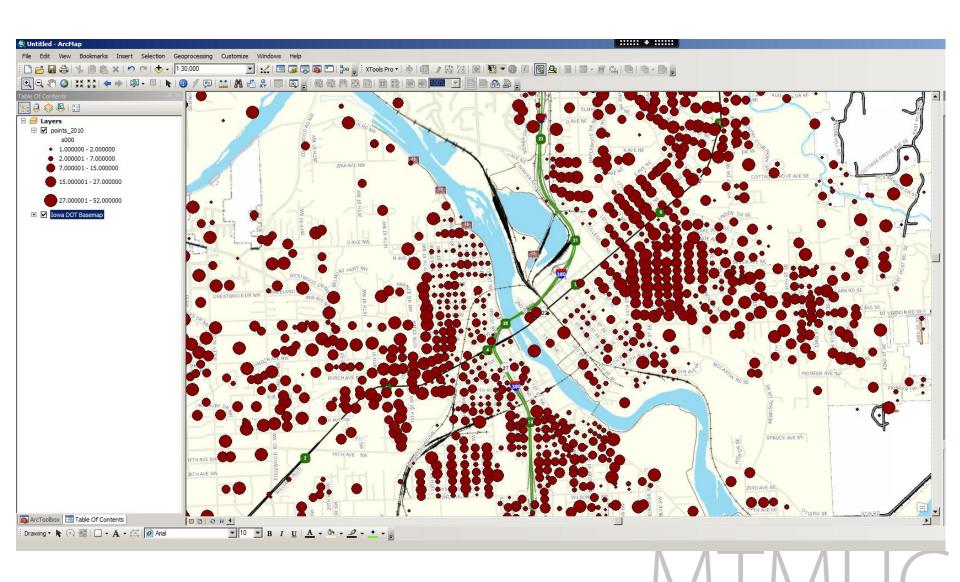
### Trip Distribution – Data Sources

- CTPP (5-year ACS): Soon?
  - At Traffic Analysis District (TAD) level (20,000 pop)
- LEHD
  - On-The-Map interface
  - Block-Level
  - Analysis exportable to SHP
  - Some "Fuzzy" Data Use with Caution

#### **ON-THE-MAP: Commute Flows**



#### Where do Downtown CR Workers Live?



#### Thank You!

- Questions?
- Additional Discussion