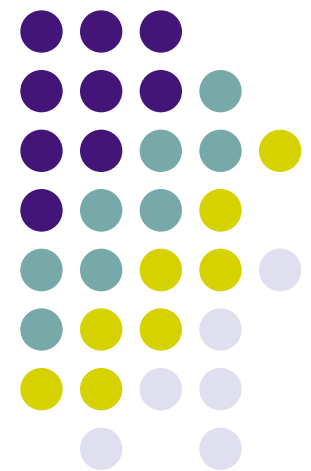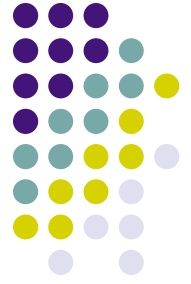# National County-level Commodity Flow Data Generation and Analysis

Jane Lin, Ph.D.,
Associate Professor
Department of Civil and Materials Engineering,
Institute for Environmental Science and Policy
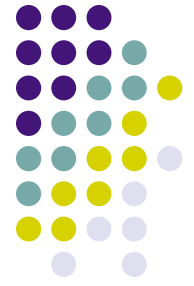University of Illinois at Chicago

Cincinnati OH, April 27, 2010,

UIC University of Illinois at Chicago

1

Outline

- County level commodity flow data generation using public data sources
  - In-/out-bound shipments
  - O-D flow at 2- and 3-digit SCTG
- Validation/case studies
- Spatial analysis of commodity flows
- Summary

Freight Data Issues

- No nationwide freight movements data publicly available down at the county level or lower
  - NCHRP Synthesis 298, "*the state of the practice in truck trip generation data is fairly <u>primitive </u>compared with trip generation data practices used in analyzing passenger vehicle movements*".

Data	Source	Availability
CFS	U.S. Census Bureau, U.S. Department of Commerce, Bureau of Transportation Statistics	Publicly available at state or metropolitan level
TransBorder	Bureau of Transportation Statistics	Publicly available
PIERS	Journal of Commerce	Proprietary
Waybill	Association of American Railroads	Public-used and government-restricted versions
Transearch	Global Insight, Inc.	Proprietary

National County Level Commodity Flow Data Generation Using Public Data Sources

- ● Research Framework

# Summary Status of National County Level Commodity Flow Data
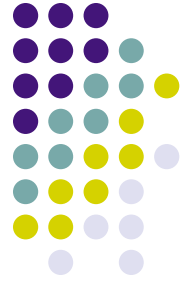
- County-level commodity inbound and outbound flow (completed)
  - 3139 counties
  - 43 commodities, including one "unknown" category
  - 33 3-digit NAICS industries including wholesale, utilities and supporting activities
- County-level commodity O-D flow (just completed)
  - 2-digit SCTG: 43 commodities, highway trucks
  - 3-digit SCTG: Corn (022), Soybeans (034), Dairy (071), Manmade Fibers (242), and Auto Parts (364), highway trucks
- Will be available in Freight Data View soon
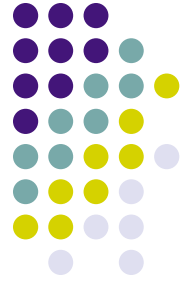
Public Data Sources

- Freight Data:
  - Commodity Flow Survey 2002 (CFS 2002)
  - Freight Analysis Framework 2002 (FAF)
- Auxiliary Data:
  - County Business Pattern 2002 (CBP 2002)
  - Quarterly Census of Employment and Wages (ES202)
  - Leonard's Guide: National Warehouse and Distribution Directory 2002
  - National Transportation Atlas 2002
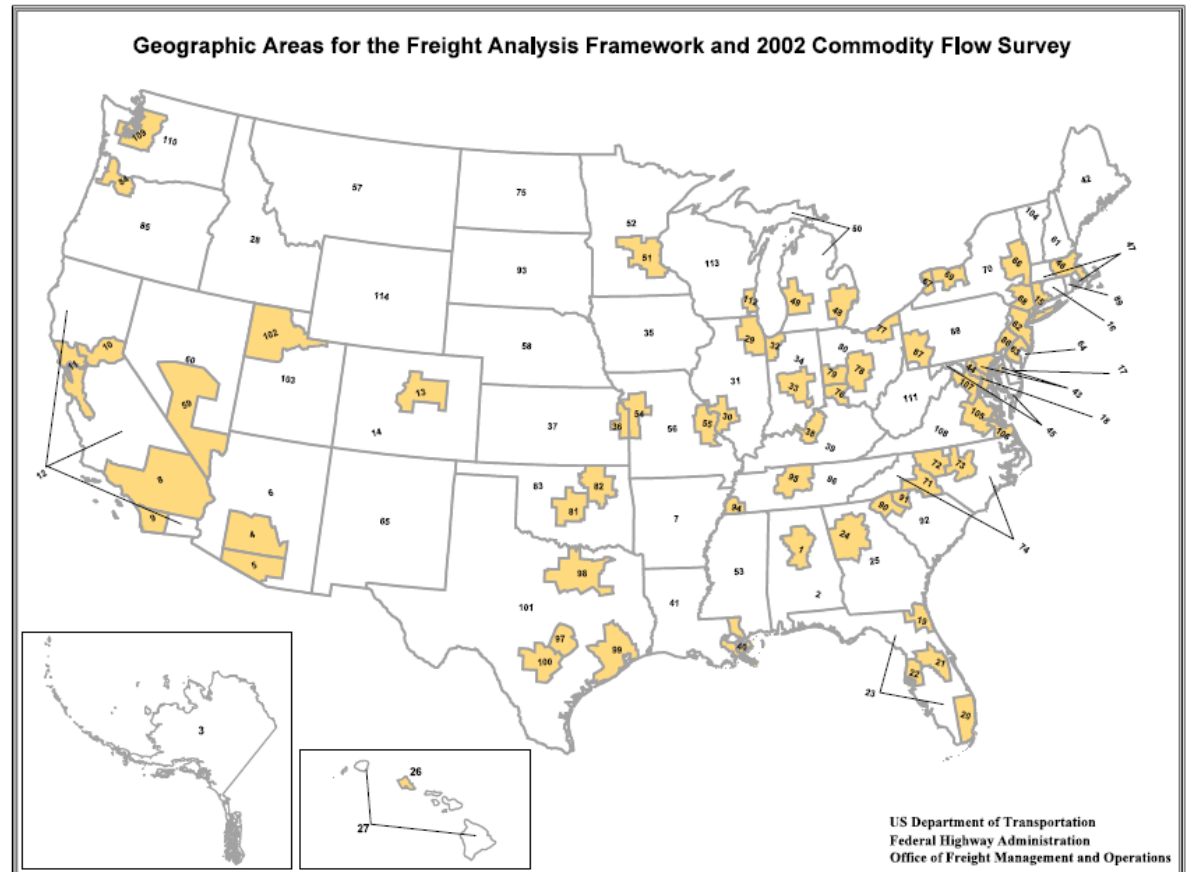
Freight Analysis Framework (FAF)
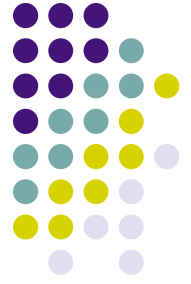
- Freight Analysis Framework (FAF) by Federal Highway Administration (FHWA):
  - Based on year 2002 and forecasted for 2010 through 2035.
  - Four datasets, in terms of commodity values and weights, that include the following at the FAF regional level:
    - Domestic activities,
    - Trans-border activities with Canada and Mexico,
    - Seaport international activities,
    - Air international activities.

FAF Regions
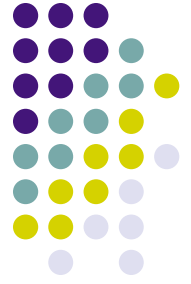
- 114 Domestic Regions

    - 64 Metropolitan Areas
    - 50 other complete states or the rest of the states

- 17 International Gateways

- 7 Foreign Regions

Geographic Areas for the Freight Analysis Framework and 2002 Commodity Flow Survey

US Department of Transportation
Federal Highway Administration
Office of Freight Management and Operations

114 domestic FAF regions

County Level In/outbound Commodity Shipment Estimation
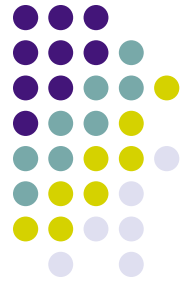
- Approach:
  - *County level outbound shipments*: Disaggregating FAF region totals into counties proportional to the total employment counts of all related industries adjusted by annual payroll for each commodity
  - *County level inbound shipments*: Simultaneous equations to estimate county inbound and outbound shipments
  - Optimum Spatial Disaggregation Method to minimize the estimation errors in counties

Simultaneous Equations Model

- Simultaneously estimating outbound and inbound shipments of each county by commodity, to take into account the endogeneity nature of the inbound and outbound activities:

$$y_1 c_{11} + y_2 c_{21} + x_1 b_{11} + x_2 b_{21} + \ldots + x_k b_{k1} = e_1$$

$$y_1 c_{12} + y_2 c_{22} + x_1 b_{12} + x_2 b_{22} + \ldots + x_k b_{k2} = e_2$$

where

y_1, y_2: Inbound and Outbound shipment by commodity, endogenous variables.

x_1, x_2, ... , x_k: explanatory variables, employment, number of intermodal facilities, etc.

e_1, e_2: Stochastic disturbance terms or random variables.

c: Coefficients of endogenous variables.

b: Coefficients of predetermined variables.

Simultaneous Equations Model

- Explanatory variables considered
  - Employment in 3-digit NAICS Manufacture industries
  - Employment in supporting activity industries
  - Employment in wholesale
  - Number of intermodal facilities
  - Tonnage of Outbound shipment by commodity
  - Tonnage of Inbound shipment by commodity

Simultaneous Modeling Results for Inbound Shipments

SCTG	Description	Model Specification	Adj R²
01	Live Animals and Fish	0.733579*Out_SCTG1 + 20.13982*EMP Food Mfg.	0.94933
02	Cereal Grains (including seed)	0.950475* Out_SCTG2	0.79517
03	Other Agricultural Products, except for Animal Feed	0.949515*Out_SCTG3	0.79046
04	Animal Feed and Products of Animal Origin, n.e.c.	1.029461*log(Out_SCTG4)	0.98350
05	Meat, Fish, and Seafood, and Their Preparations	1.055760*log(Out_SCTG5)	0.97965
06	Milled Grain Products and Preparations, and Bakery Products	0.681403* Out_SCTG6 + 10.91014* EMP Wholesale	0.86249
07	Other Prepared Foodstuffs, and Fats and Oils	1.020763* Out_SCTG7	0.84410
08	Alcoholic Beverages	1.063763* Out_SCTG8	0.82965
09	Tobacco Products	1.081081* Out_SCTG9	0.93913
10	Monumental or Building Stone	0.903294* Out_SCTG10	0.84245
11	Natural Sands	1.034637* Out_SCTG11	0.86430
12	Gravel and Crushed Stone	1.003968* Out_SCTG12	0.91322
13	Non-Metallic Minerals, n.e.c.	0.772316* Out_SCTG13+ 21.50881* EMP Motor vehicle supply	0.81487
14	Metallic Ores and Concentrates	0.987698*log(Out_SCTG14)+0.118004*EMP Primary Metal Mfg.	0.81562
15	Coal	1.688131*log(Out_SCTG15)	0.64818
16	Crude Petroleum Oil	1.219936*log(Out_SCTG16)	0.53984
17	Gasoline and Aviation Turbine Fuel	0.941359*Out_SCTG17	0.93402
18	Fuel Oils	0.945129*Out_SCTG18	0.94576
19	Coal and Petroleum Products, n.e.c.	1.031149*log(Out_SCTG19)	0.99047
20	Basic Chemicals	0.913759*Out_SCTG20	0.84206
21	Pharmaceutical Products	1.075287*log(Out_SCTG21)	0.96743
22	Fertilizers	1.049804*log(Out_SCTG22)	0.97888

* SCTG (Standard Classification of Transported Goods)
* All variables are in 95% significance level.
* In commodities that with log-transformed variables, the dependent variables are also log-transformed.

Simultaneous Modeling Results for Inbound Shipments

SCTG	Description	Model Specification	Adj. R²
23	Chemical Products and Preparations, n.e.c.	0.532276*Out_SCTG23 + 115.6089* EMP Utilities	0.87063
24	Plastics and Rubber	0.803330*Out_SCTG24	0.81680
25	Logs and Other Wood in the Rough	0.962659*Out_SCTG25	0.98582
26	Wood Products	0.601101*Out_SCTG26 + 54.75438* EMP Wholesale	0.90245
27	Pulp, Newsprint, Paper, and Paperboard	0.944358*Out_SCTG28 + 73.97757* EMP Utilities	0.82602
28	Paper or Paperboard Articles	0.627966*Out_SCTG28 + 10.01022* EMP Wholesale	0.92787
29	Printed Products	1.018861*Out_SCTG29	0.89468
30	Textiles, Leather, and Articles of Textiles or Leather	0.659579*Out_SCTG30 + 34.13657*EMP Utilities	0.90345
31	Non-Metallic Mineral Products	0.990029*Out_SCTG31	0.96284
32	Base Metal in Primary or Semi-Finished Forms and in Finished Basic Shapes	0.711823*Out_SCTG32+ 42.40031* EMP Wholesale	0.83481
33	Articles of Base Metal	0.966960*Out_SCTG33	0.95455
34	Machinery	1.296673*Out_SCTG34	0.74727
35	Electronic and Other Electrical Equipment and Components, and Office Equipment	0.597900*Out_SCTG35 + 6.742093* EMP Wholesale	0.90016
36	Motorized and Other Vehicles (including parts)	0.774298*Out_SCTG36 + 9.784049* EMP Wholesale	0.86201
37	Transportation Equipment, n.e.c.	1.021461*log(Out_SCTG37)	0.99005
38	Precision Instruments and Apparatus	1.082404* log(Out_SCTG38)	0.96237
39	Furniture, Mattresses and Mattress Supports, Lamps, Lighting Fittings, and Illuminated Signs	0.507850*Out_SCTG39 + 5.933223*EMP Wholesale	0.91112
40	Miscellaneous Manufactured Products	0.911020*Out_SCTG40	0.90440
41	Waste and Scrap	1.009059*Out_SCTG41	0.83637
42	Unknown	0.891684* Out_SCTG42 + 76.50838* EMP Misc。 Mfg.	0.93540
43	Mixed Freight	0.957848*Out_SCTG43	0.93949

* SCTG (Standard Classification of Transported Goods)
* All variables are in 95% significance level.
* In commodities that with log-transformed variables, the dependent variables are also log-transformed.

Optimum Spatial Disaggregation

- Finding the spatially correlated covariance matrix that minimizes the estimation errors when disaggregated from higher level to lower level geographic scales:
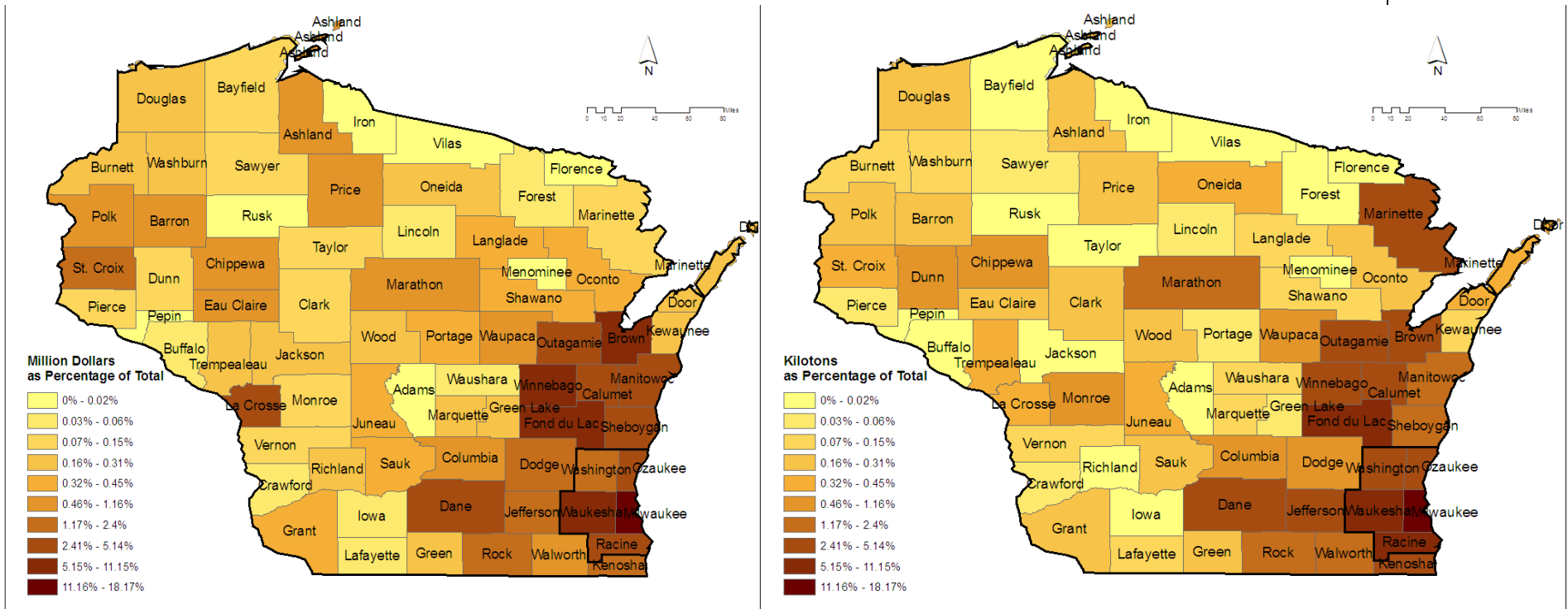
$$\hat{\beta} = (X'C'(C\Omega C')^{-1}CX)^{-1}X'C'(C\Omega C')^{-1}Cy$$
$$\hat{y} = X\hat{\beta} + \Omega C'(C\Omega C')^{-1}(Cy - CX\hat{\beta})$$

where C is the aggregation matrix and Ω is the covariance matrix

- Assumptions:
  - Similarity in model structure between upper and lower levels
  - Similarity in error structure between upper and lower levels

Validation (I) –Wisconsin

FAF regions and counties

FAF Top 10 Commodities outbound shipments in values in Wisconsin

SCTG	Commodity	Value (Million Dollars)
34	Machinery	34,706.46
42	Mixed freight	30,897.32
7	Other foodstuffs	26,245.02
36	Motorized vehicles	22,925.23
35	Electronics	14,257.00
40	Misc. mfg. prods.	12,294.91
33	Articles-base metal	10,533.90
30	Textiles/leather	9,397.40
28	Paper articles	9,143.73
43	Unknown	8,746.17

- Zone 112: Milwaukee-Racine-Waukesha, which includes 5 counties
- Zone 113: The rest of Wisconsin, which includes 67 counties

15

Machinery Outbound Shipments

Proportional weighting results (normalized)

TRANSEARCH (normalized)

Proportional weighting surrogate variables: total employment rates in industries of:
- Machinery manufacturing industries
- Supporting industries providing materials, tools, preprocessing, packaging and selling for machinery manufacturing.
- Requiring the input of machinery products.

16

Validation (II) - Minnesota

- Correlations with the Minnesota TRANSEARCH County Data

SCTG Group	Inbound	Outbound
Agriculture, Animal, Alcoholic, Tobacco Products, etc.	0.49334	0.71590
Monumental, Natural Sands, Gravel, Non-Metallic Minerals, Metallic Ores and Concentrates, etc.	0.68744	0.28613
Coal, Crude Petroleum Oil, Gasoline, etc.	0.70520	0.18206
Basic Chemicals, Fertilizers, Pharmaceutical and Chemical Products, etc.	0.74399	0.85986
Plastics and Rubber	0.82614	0.86305
Logs and Woods, Metal, Precision Instruments and Apparatus, Miscellaneous Manufactured Products, etc.	0.86722	0.67896
Pulp, Newsprint, Paper, Paperboard, etc.	0.85537	0.69349
Printed Products	0.88289	0.88392
Textiles, Leather, and Articles of Textiles or Leather	0.90507	0.70113
Non-Metallic Mineral Products	0.74008	0.61667
Machinery	0.88034	0.95297
Electronic, Vehicles, Transportation Equipment, Furniture, Mattresses and Mattress Supports, Lamps, etc.	0.93511	0.82584
Mixed Freight, Unknown	0.99655	0.99832

SCTG (Standard Classification of Transported Goods) are grouped based on the bridge between SCTG and STCC (Standard Transportation Commodity Classification) , since STCC is used in TRANSEARCH.
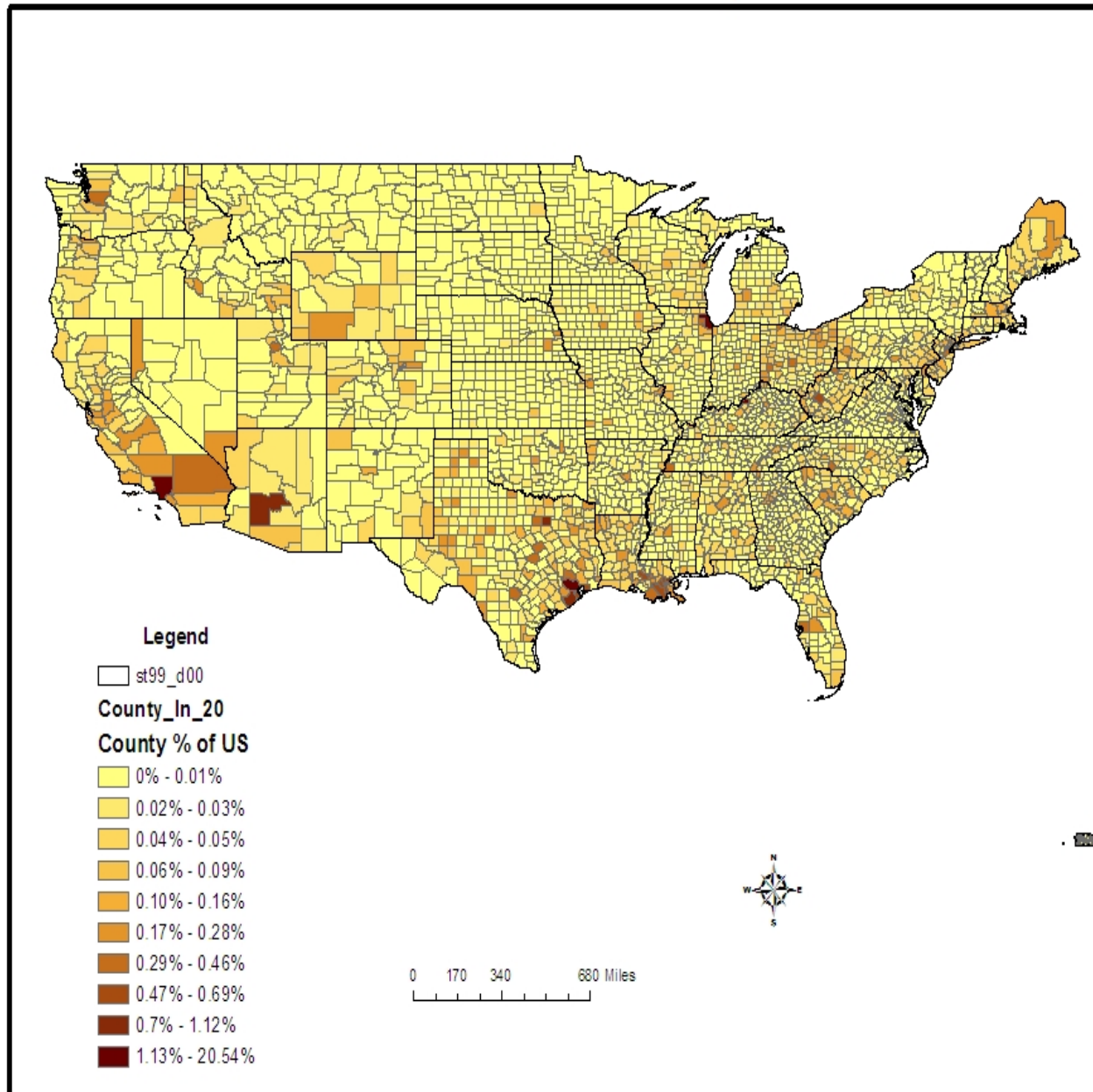
US Machinery Outbound Shipment

Legend

st99_d00

County_Out_34

County % of US

- 0% - 0.01%
- 0.02% - 0.03%
- 0.04% - 0.05%
- 0.06% - 0.09%
- 0.1% - 0.16%
- 0.17% - 0.28%
- 0.29% - 0.46%
- 0.47% - 0.69%
- 0.7% - 1.12%
- 1.13% - 20.54%

0   170   340   680 Miles
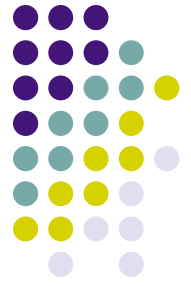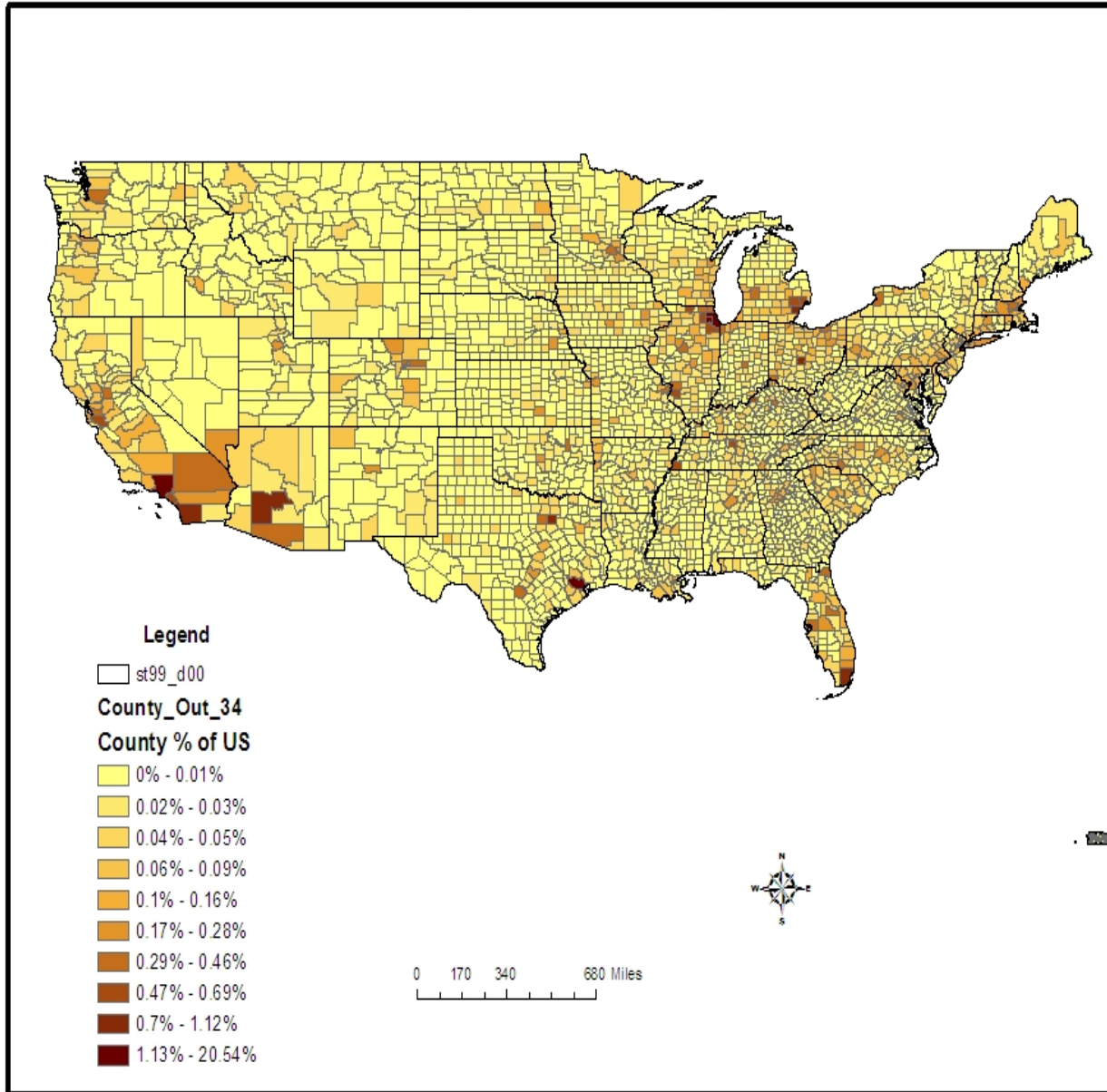
US Machinery Inbound Shipment

Legend

st99_d00

County_In_34
County % of US

- 0% - 0.01%
- 0.02% - 0.03%
- 0.04% - 0.05%
- 0.06% - 0.09%
- 0.1% - 0.16%
- 0.17% - 0.28%
- 0.29% - 0.46%
- 0.47% - 0.69%
- 0.7% - 1.12%
- 1.13% - 20.54%

0   170   340   680 Miles
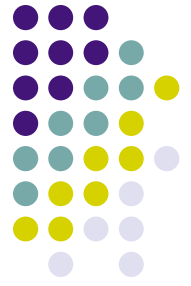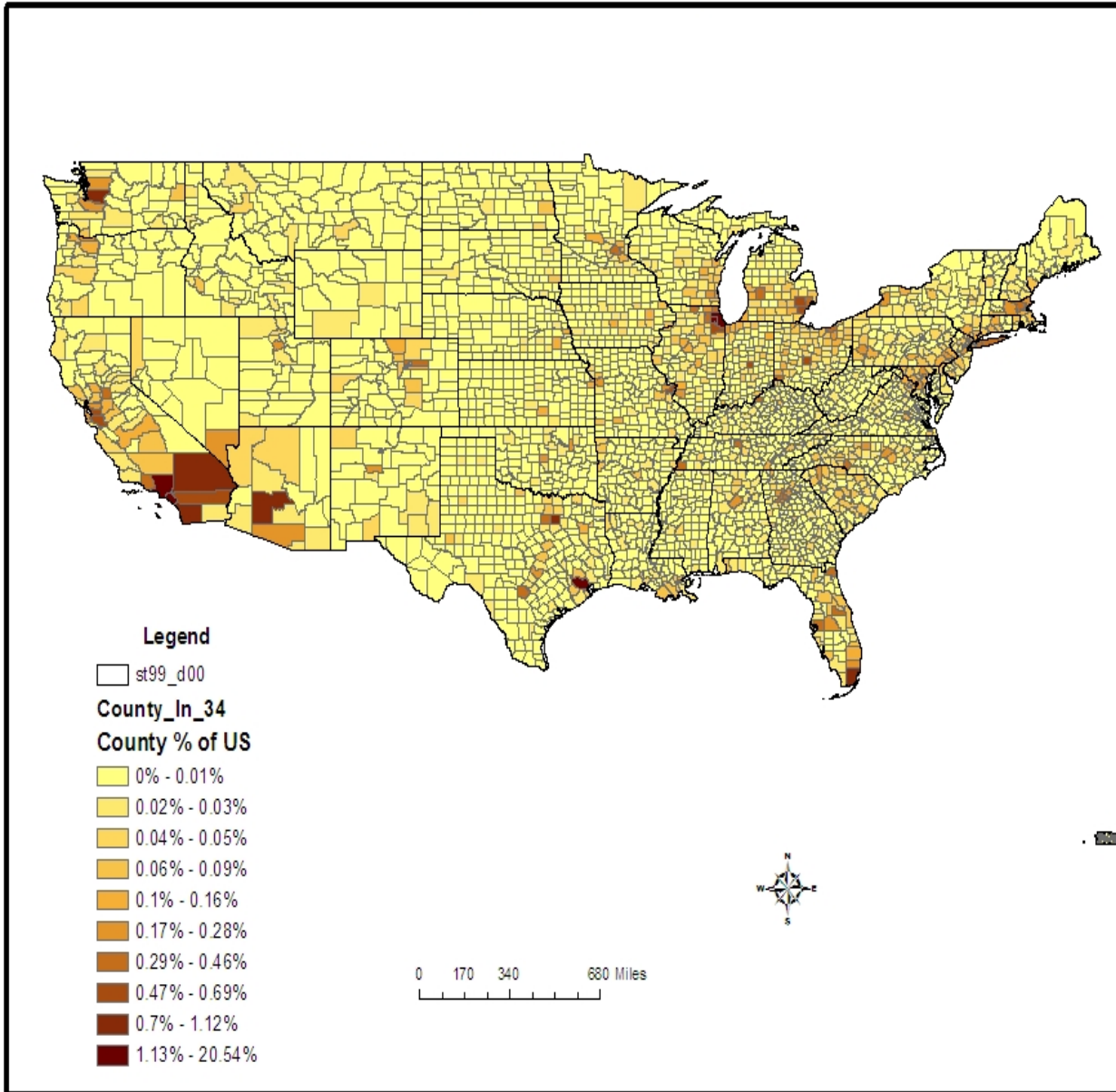
US Basic Chemicals Outbound Shipment

Legend

st99_d00

County_Out_20

County % of US

- 0% - 0.01%
- 0.02% - 0.03%
- 0.04% - 0.05%
- 0.06% - 0.09%
- 0.10% - 0.16%
- 0.17% - 0.28%
- 0.29% - 0.46%
- 0.47% - 0.69%
- 0.70% - 1.12%
- 1.13% - 20.54%

0   170   340   680 Miles
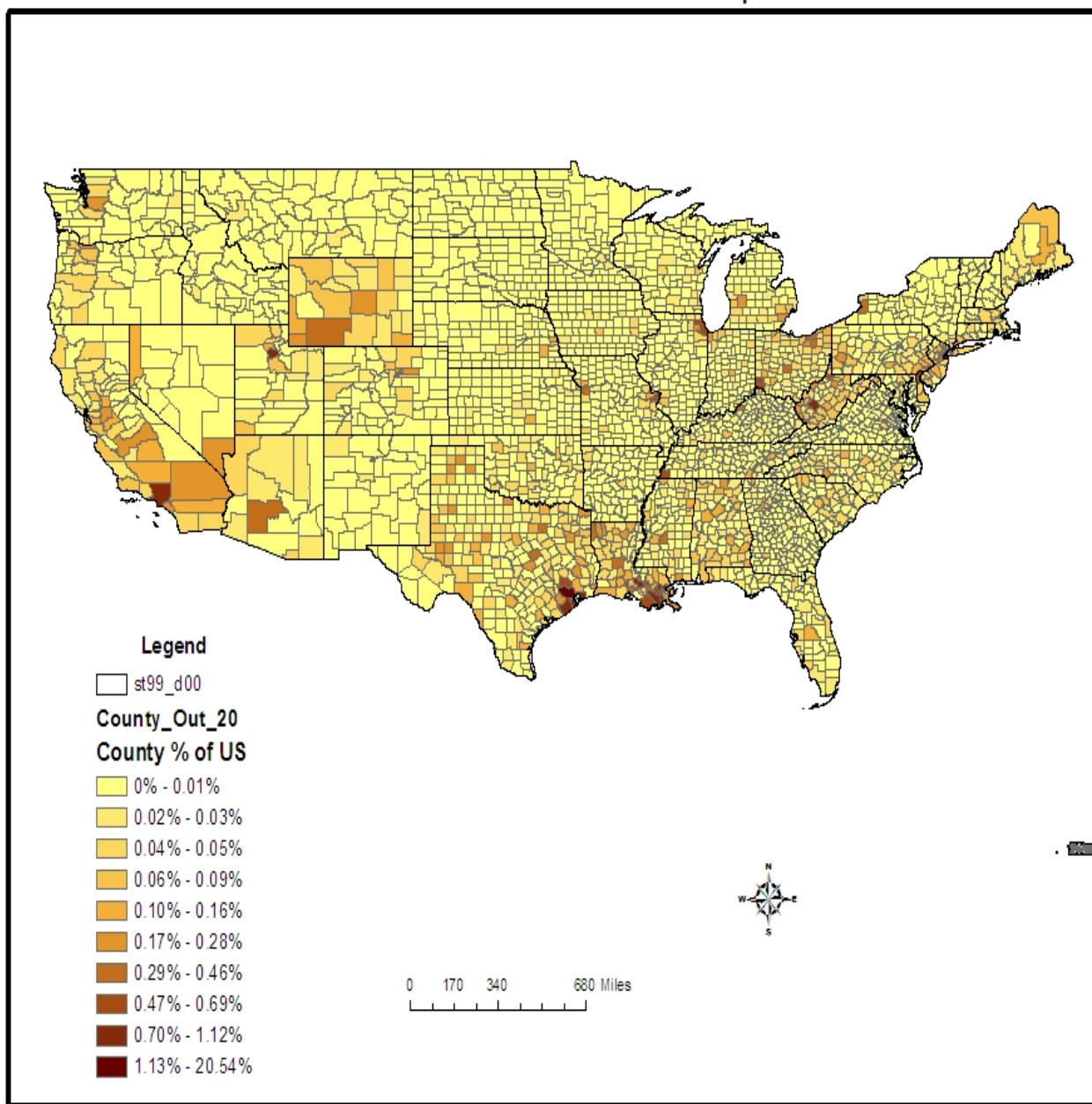
US Basic Chemicals Inbound Shipment

Legend

☐ st99_d00

County_In_20
County % of US

☐ 0% - 0.01%
☐ 0.02% - 0.03%
☐ 0.04% - 0.05%
☐ 0.06% - 0.09%
☐ 0.10% - 0.16%
☐ 0.17% - 0.28%
☐ 0.29% - 0.46%
☐ 0.47% - 0.69%
☐ 0.7% - 1.12%
☐ 1.13% - 20.54%

0   170   340   680 Miles
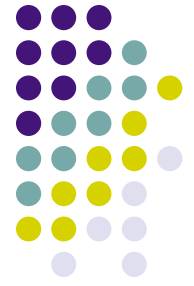
County-level O-D Flow Data

Aggregate FAF region 2-digit
SCTG O-D by commodity by mode

↓

Generate county level 2-digit SCTG
truck O-D (43 commodities)*

↓

Aggregate to county level total
commodity truck O-D

↓

Generate county level 3-digit SCTG
truck O-D by 3-digit SCTG tons
percentage
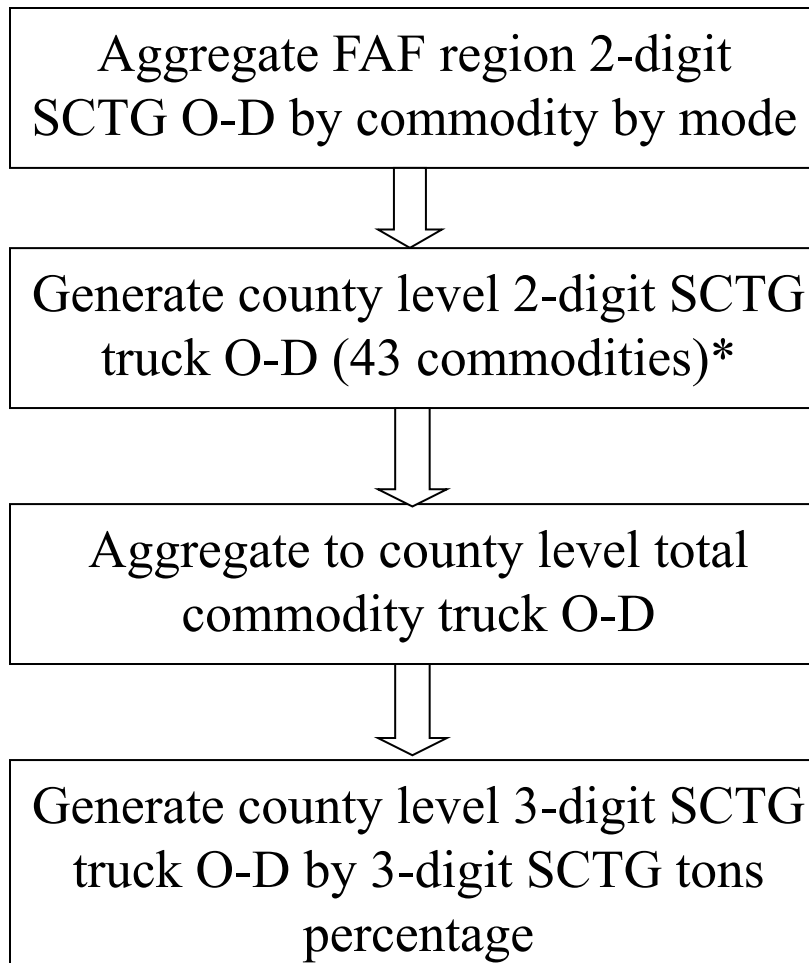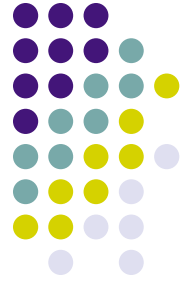
*Outbound shipment using county
employment percentage within the
FAF
* Inbound shipment: county
population percentage within the FAF

Available data

- **2-Digit SCTG Commodity:** FAF2.2 commodity O-D in FAF region by commodity type by mode, include activities in:
  - Domestic activities,
  - Trans-border activities with Canada and Mexico,
  - Seaport international activities,
  - Air international activities.

- **3-Digit SCTG Commodity:** CFS report Table 11 - Shipment Characteristics by Three-Digit Commodity for State of Origin: 2002

- **Employment:** County Business Pattern (CBP) 2002

- **Population:** Census 2000

Spatial Autocorrelation Analysis

- The previous disaggregation analysis shows that commodity outbound and inbound shipments are **not independent across regions.**

- In order to understand the **extent to which such spatial dependence exists**, this study performed an investigation of the spatial autocorrelations between and among the outbound and inbound shipments, at the FAF regional level and the county-level respectively, throughout the country.

- Local Moran's I statistic is a standard measure of spatial autocorrelation.

  - County level
  - FAF region

Spatial Autocorrelation Analysis
----*Local Moran's I*
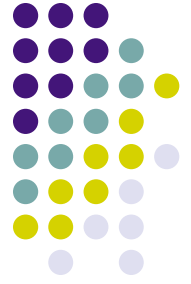
- **Local Moran's I:** capture the local component of spatial autocorrelation.

- Visualizes four spatial autocorrelation combinations in four quadrants:
  - Positive Autocorrelation: High-High and Low-Low for 4 in-/out-bound combinations
  - Negative Autocorrelation: High-Low and Low-High for 4 in-/out-bound combinations

Local Moran's I Results
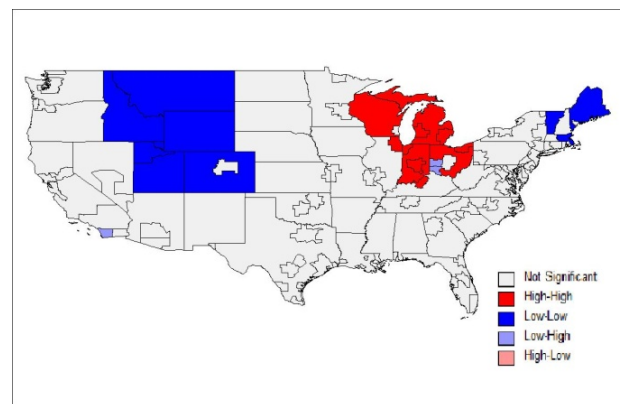
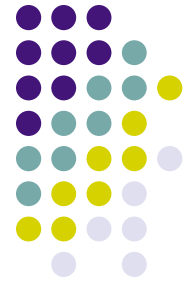- **FAF Live Animals and Fish (SCTG 1)**
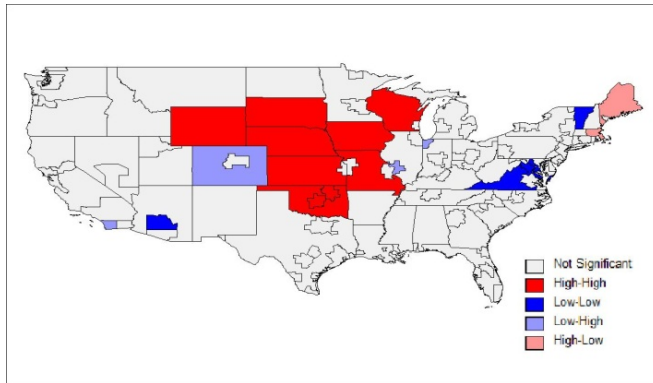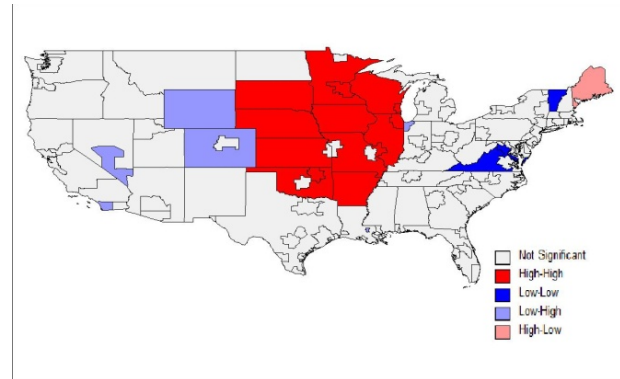
(Outbound-Outbound)
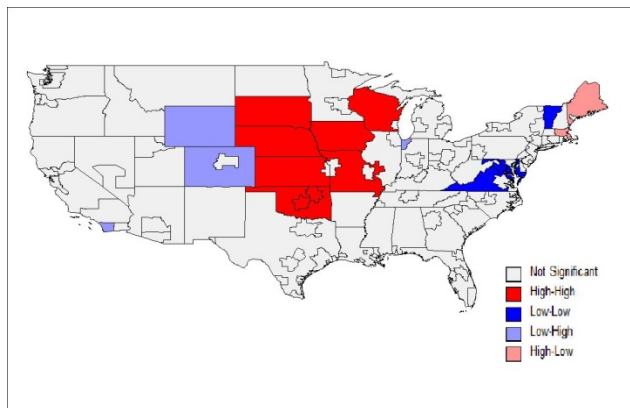
(Inbound-Inbound)
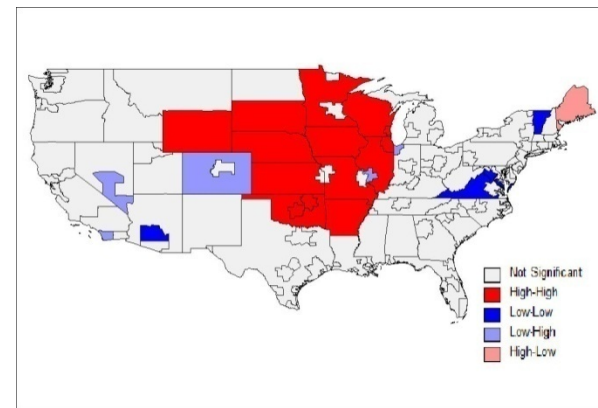
(Outbound-Inbound)

(Inbound-Outbound)

Local Moran's I Results *(con't)*

- **County Live Animals and Fish (SCTG 1)**

(Outbound-Outbound)

(Inbound-Inbound)

(Outbound-Inbound)

(Inbound-Outbound)

27

Local Moran's I Results *(con't)*

- **FAF Motorized and Other Vehicles (SCTG 36)**

(Outbound-Outbound)

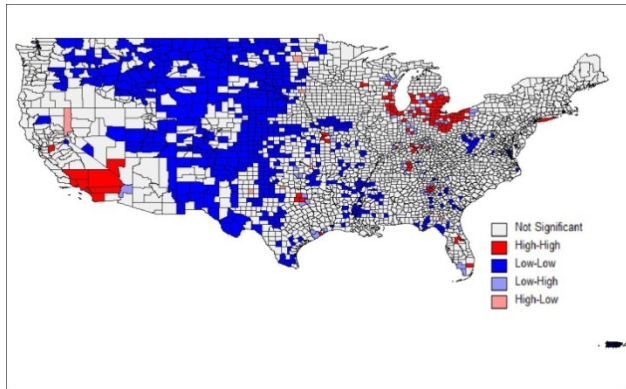(Inbound-Inbound)

(Outbound-Inbound)
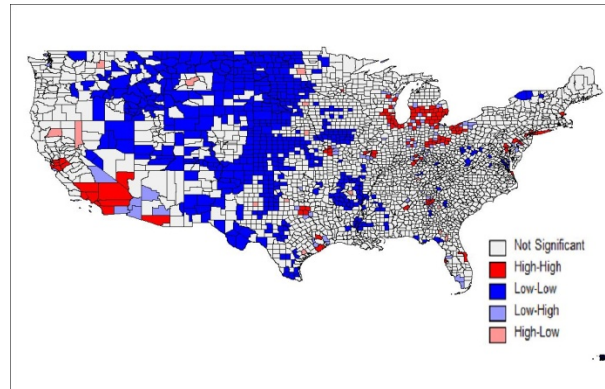

(Inbound-Outbound)

28

Local Moran's I Results *(con't)*
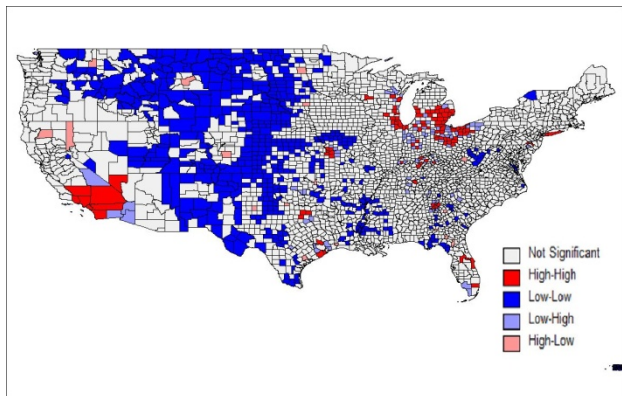
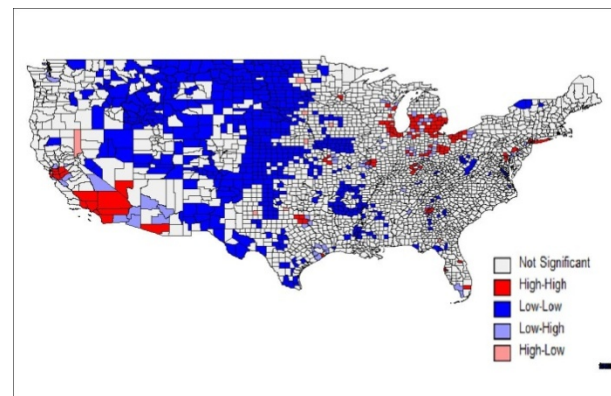- **County Motorized and Other Vehicles (SCTG 36)**
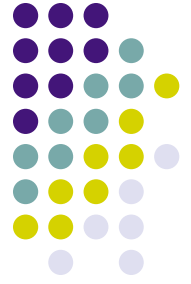

(Outbound-Outbound)

(Inbound-Inbound)

(Outbound-Inbound)

29   (Inbound-Outbound)

Summary

- This study proposed a freight data synthesis framework in generating high geographic resolution freight data, which were validated with available TRANSEARCH data in two midwestern states

- The comparison results showed good agreement between synthesized and existing data

- It is worth pointing out that the above conclusions are based on the comparisons with the TRANSEARCH data. More robust validation should be to compare the estimated results with true regional samples (e.g., from a regional survey, GPS data).

- The spatial autocorrelation analyses at both the FAF region and county level reveal that **commodity shipments are generally spatially concentrated**, consistent with the notion of clusters of economic activities.

- Compared with the FAF region based spatial patterns, it has been shown that the **county-based spatial autocorrelation maps are visually more realistic** with greater spatial information and are able to capture spatial variability at much finer resolutions.