

bridging the gap

Research Informing Policies & Practices
for Healthy Youth

Field Validation of Secondary Commercial Data Sources on the Retail Food Outlet Environment in the U.S.

International Society for Behavioral Nutrition and Physical Activity
Minneapolis, MN, U.S.A., June 11, 2010

Lisa M. Powell, PhD
University of Illinois at Chicago

Validation Research Team

Lisa Powell

Tamkeen Khan

Oksana Pugach

Diane Barker

Janna Myllyluoma

Punam Ohri-Vachaspati

Euna Han

Christopher Quinn

Kevin Gibbs

Shannon Zenk

Elli Resnick

Frank Chaloupka

18 field workers

6 data entry workers

Presentation Outline

1. Introduction

- Motivation
- Objectives

2. Methods

- Study Design
- Field Instruments
- Classification
- Statistical Analyses

3. Results

4. Implications

Introduction

Motivation and Objectives

Motivation

- Numerous empirical studies have relied on and continue to rely on commercial data sources to measure the availability of food stores and/or restaurants in specific geographic areas.
- Previous validation studies have shown that moderate count error exists but that it does not seem to be systematic.
- However, previous studies have not validated commercial sources conditional on exact classification matches for specific types of food stores and restaurants.
- Count and classification errors may bias study findings:
 - For example, fewer supermarkets in low-income neighborhoods may be associated with higher rates of overweight. But what if the commercial data lists systematically underreport supermarket availability in such neighborhoods

Objectives

- To validate commercial food outlet business list databases and to provide detailed validation statistics for food stores and restaurants by type of outlet, including supermarkets, grocery stores, convenience stores, full-service restaurants and fast food.
- To test the validity of the secondary sources based on overall agreement with the field observations *and* agreement conditional on exact classification match between the field and the list by outlet type.
- To test whether validity differed by neighborhood characteristics such as income, race, ethnicity and urbanicity and by outlet type.

Methods

Study Design, Field Instruments, Classification and Statistical Analyses

Study Design

- Direct field observations including in store instruments by trained field staff
- Commercial outlet data
 - Dun & Bradstreet
 - InfoUSA
- Census tracts drawn from the Chicago MSA and a 50 buffer around the MSA; sample of 331 tracts, stratified by:
 - Income tertiles
 - Low income:
 - Race (>70% white, >70% black, mixed)
 - Ethnicity (>70% Hispanic, Non-Hispanic)
 - Urbancity (urban, suburban, rural)

Field Instruments

SECTION A: Restaurants

INSIDE THE RESTAURANT:

A1. Is the food order...	NO	YES
a) placed at counter?	0	1
b) delivered at counter?	0	1
c) paid at counter?	0	1

A2. Is there...	NO	YES
a) indoor seating area available?	0	1
b) a bar with seating, serving alcohol?	0	1

OUTSIDE THE RESTAURANT:

A3. Is the restaurant...	NO	YES
a) attached to or within a convenience store?	0	1
b) attached to or within a grocery?	0	1
c) attached to another restaurant?	0	1
d) attached to a gas station?	0	1
e) in a Food Court in a mall?	0	1
f) within a general merchandise store?	0	1

A4. Is there a Drive-Up or Drive-Through Window for ordering?	NO	YES
	0	1

SECTION B: Food Stores

INSIDE THE STORE:

B1. Number of Cash Registers

B2. Does the store sell/have...	NO	YES
a) fresh milk?	0	1
b) fresh meats?	0	1
c) a fresh produce section?	0	1

B3. Number of different types of fresh fruits and vegetables

None	0
1 - 9	1
10 -19	2
20 or more	3

B4. Does it have...	NO	YES
a) a pharmacy counter?	0	1
b) a bank?	0	1
c) a bakery counter?	0	1
d) a deli counter?	0	1
e) a butcher or section for unpackaged, fresh meat?	0	1

OUTSIDE THE STORE:

B5. Is the food store...	NO	YES
a) attached to a gas station?	0	1
b) attached to a restaurant?	0	1

Classification of Food Stores found on the Ground

Specialty Food Stores: Bakeries, meat or fish stores, fruit or vegetable stores, candy or nut stores and coffee and tea stores.

Convenience Stores: Non-specialty food stores with no fresh meat, at most a selection of ten fruits and vegetables and two or fewer cash registers.

Supermarkets: At least four cash registers, at least a selection of twenty or more fruits and vegetables, sold fresh meats and had at least two of the three features including a butcher, deli or bakery.

Grocery Stores: Not a specialty food store, a supermarket or a convenience store.

Classification of Restaurants found on the Ground

Specialty restaurants: Coffee shops, donut shops, ice cream parlors, pretzel shops, banquet halls.

Fast food restaurants: Patrons order food at the counter and pay at the counter.

Full-service restaurants: Patrons offered patrons table service (patrons did not order and pay at the counter) and those restaurants that were not specialty restaurants.

SIC Classification of Food Stores from the Commercial Lists

Store classification	D&B list	InfoUSA list
Convenience store	<ul style="list-style-type: none"> • 541102 (convenience stores) 	<ul style="list-style-type: none"> • 541103 (convenience stores)
Supermarket	<ul style="list-style-type: none"> • 541101 (supermarkets) 	<ul style="list-style-type: none"> • 541101 (food markets) • 541102 (snack products) • 541104 (food products-retail) • 541105 (grocers-retail) • 541106 (markets-kosher) • 541107 (grocers-ethnic foods) • 541108 (grocers-health foods) • 541108 (grocers-take-out foods)
Grocery stores	<ul style="list-style-type: none"> • 541100 (grocery Stores) • 541199 (grocery stores, nec) 	<ul style="list-style-type: none"> • Same as supermarket‡
Specialty food stores	<ul style="list-style-type: none"> • 5421 (meat and fish markets) • 5431 (fruit and vegetable market) • 5441 (candy, nut, and confectionary stores) • 5451 (dairy products stores) • 5461 (retail bakeries) • 5499 (miscellaneous food stores) 	<ul style="list-style-type: none"> • 5421 (meat and fish markets) • 5431 (fruit and vegetable market) • 5441 (candy, nut, and confectionary stores) • 5451 (dairy products stores) • 5461 (retail bakeries) • 5499 (miscellaneous food stores)

SIC Classification of Restaurants from the Commercial Lists

Store classification	D&B list	InfoUSA list
Fast-food restaurants	<ul style="list-style-type: none"> • 581203 (fast food restaurants and stand) except for 58120304 (coffee shop), 	<ul style="list-style-type: none"> • 581206 (food-carry out) • 581219 (sandwiches) • 581208 (delicatessens) • 581229 (deli-bakery)
Full service restaurants	<ul style="list-style-type: none"> • 581200 (eating places) • 581201 (ethnic food restaurants) • 581204 (lunchroom and cafeterias) • 581205 (family restaurants) • 581206 (pizza restaurants) • 581207 (seafood restaurants) • 581208 (steak and barbecue restaurants) • 581209 (eating places, nec) except for 58129903 (caterers), and 58129906 (contract food services) 	<ul style="list-style-type: none"> • 581201 (Luau's) • 581208 (Restaurants) • 581215 (Box lunches) • 581224 (Barbecue) • 581232 (Chow mein & Chop suey) • 581233 (Jewish foods) • 581240 (Chuck wagon dinners) • 581245 (Escargots) • 581247 (Restaurants with liquor) • 581250 (Wedding rehearsal restaurants) • 581251 (Restaurants-family dining) • 581202 (Greek food products) • 581211 (Chicken dinners) • 581222 (Pizza) • 581227 (Italian food products) • 581235 (Fine dining) • 581244 (Oyster bars) • 581246 (Chili parlors) • 581249 (Restaurants-reservations) • 581252 (Cajun products)
Specialty restaurants	<ul style="list-style-type: none"> • 581202 (ice cream, soft drink, and soda fountain stand) • 58120304 (coffee shop) • 58129903 (caterers) • 58129906 (contract food services) 	<ul style="list-style-type: none"> • 581203 Ice cream parlors • 581205 Hamburgers & hot dog stands • 581210 Commissary contractors • 581213 Cafeterias • 581220 Restaurant management • 581221 Refreshment stands • 581225 Beverages-Non alcoholic-retail • 581226 Theaters-dinner • 581231 Food service -Industrial • 581234 Reception Centers • 581236 Tea Rooms • 581237 Dessert Preparations • 581238 Cafeterias-Industrial-Operators • 581239 Catering-Kosher • 581241 Milk Bars • 581242 Mobile Concessions • 581243 Luncheonette Food Supply • 581248 Juice Bars • 581253 Meal Preparation Services • 581204 Wedding cakes • 581207 Foods-Institutional • 581212 Caterers • 581214 Cafes • 581223 Banquet rooms • 581230 Restaurants-foods delivery

Statistical Analyses

- Matches were based on US_Streets and US_RoofTop geocoded addresses which covered 97% of the outlets in InfoUSA and 94% in D&B.
- **Sensitivity**: the percentage of establishments observed on the ground that was listed in the commercial databases
- **Positive Predictive Value (PPV)**: the percentage of establishments listed in the commercial databases that was observed on the ground.
- **Concordance**: the proportion of all establishments on the ground *and/or* listed that were both observed and listed.
- Tested differences by tract characteristics using Fisher's exact test

Results

bridging the gap

Results: Outlet Counts

Food Stores

	All Stores	Supermarket and Grocery	Convenience	Specialty
Ground Total	1130	338	558	234
D&B List	1145	518	282	345
InfoUSA List	1118	395	364	359

Restaurants

	All Restaurants	Fast Food	Full-service	Specialty
Ground Total	3111	1480	1178	453
D&B List	2666	519	1937	210
InfoUSA List	2682	358	2044	280

bridging the gap

Results: Food Store Validation Statistics

	D&B			InfoUSA		
	All Food Stores	Super/ Grocery	Conv	All Food Stores	Super/ Grocery	Conv
Sensitivity	0.52	0.62	0.50	0.64	0.75	0.60
Exact Sensitivity		0.46	0.38		0.54	0.50
PPV	0.49	0.48	0.76	0.64	0.66	0.85
Exact PPV		0.29	0.69		0.44	0.77
Concordance	0.34	0.43	0.46	0.47	0.60	0.57
Exact Concordance		0.22	0.32		0.32	0.44

Results: Food Store Validation Statistics

	D&B			InfoUSA		
	All Food Stores	Super/ Grocery	Conv	All Food Stores	Super/ Grocery	Conv
Sensitivity	0.52	0.62	0.50	0.64	0.75	0.60
Exact Sensitivity		0.46	0.38		0.54	0.50
PPV	0.49	0.48	0.76	0.64	0.66	0.85
Exact PPV		0.29	0.69		0.44	0.77
Concordance	0.34	0.43	0.46	0.47	0.60	0.57
Exact Concordance		0.22	0.32		0.32	0.44

Results: Food Store Validation Statistics

	D&B			InfoUSA		
	All Food Stores	Super/ Grocery	Conv	All Food Stores	Super/ Grocery	Conv
Sensitivity	0.52	0.62	0.50	0.64	0.75	0.60
Exact Sensitivity		0.46	0.38		0.54	0.50
PPV	0.49	0.48	0.76	0.64	0.66	0.85
Exact PPV		0.29	0.69		0.44	0.77
Concordance	0.34*	0.43*	0.46*	0.47	0.60	0.57
Exact Concordance		0.22*	0.32*		0.32	0.44

Results: Sensitivity to Exclusion of Gas Station SIC codes

	D&B			InfoUSA		
	All Food Stores	Super/ Grocery	Conv (Conv w/o Gas)	All Food Stores	Super/ Grocery	Conv (Conv w/o Gas)
Sensitivity	0.52	0.62	0.50 (0.26)	0.64	0.75	0.60 (0.27)
Exact Sensitivity		0.46	0.38 (0.14)		0.54	0.50 (0.18)
PPV	0.49	0.48	0.76 (0.53)	0.64	0.66	0.85 (0.68)
Exact PPV		0.29	0.69 (0.49)		0.44	0.77 (0.57)
Concordance	0.34	0.43	0.46 (0.24)	0.47	0.60	0.57 (0.27)
Exact Concordance		0.22	0.32 (0.12)		0.32	0.44 (0.16)

Restaurant Validation Statistics

	D&B			InfoUSA		
	All Restaurant	Fast Food	Full-Service	All Restaurant	Fast Food	Full-Service
Sensitivity	0.56	0.59	0.64	0.66	0.72	0.71
Exact Sensitivity		0.22	0.59		0.17	0.65
PPV	0.64	0.70	0.64	0.77	0.80	0.80
Exact PPV		0.61	0.36		0.62	0.39
Concordance	0.43	0.54	0.54	0.55	0.70	0.69
Exact Concordance		0.19	0.29		0.15	0.32

Restaurant Validation Statistics

	D&B			InfoUSA		
	All Restaurant	Fast Food	Full-Service	All Restaurant	Fast Food	Full-Service
Sensitivity	0.56	0.59	0.64	0.66	0.72	0.71
Exact Sensitivity		0.22	0.59		0.17	0.65
PPV	0.64	0.70	0.64	0.77	0.80	0.80
Exact PPV		0.61	0.36		0.62	0.39
Concordance	0.43	0.54	0.54	0.55	0.70	0.69
Exact Concordance		0.19	0.29		0.15	0.32

Restaurant Validation Statistics

	D&B			InfoUSA		
	All Restaurant	Fast Food	Full-Service	All Restaurant	Fast Food	Full-Service
Sensitivity	0.56	0.59	0.64	0.66	0.72	0.71
Exact Sensitivity		0.22	0.59		0.17	0.65
PPV	0.64	0.70	0.64	0.77	0.80	0.80
Exact PPV		0.61	0.36		0.62	0.39
Concordance	0.43*	0.54*	0.54*	0.55	0.70	0.69
Exact Concordance		0.19*	0.29*		0.15	0.32

Results: Sensitivity to Fast Food SICs and Chain Status

	D&B			InfoUSA		
	Fast Food	Fast Food Pizzeria	Fast Food Pizzeria Chain	Fast Food	Fast Food Pizzeria	Fast Food Pizzeria Chain
Exact Sensitivity	0.22	0.25	0.42	0.17	0.28	0.44
Exact PPV	0.61	0.59	0.72	0.62	0.60	0.77
Exact Concordance	0.19	0.21	0.36	0.15	0.23	0.39

Results: Sensitivity to Fast Food SICs and Chain Status

	D&B			InfoUSA		
	Fast Food	Fast Food Pizzeria	Fast Food Pizzeria Chain	Fast Food	Fast Food Pizzeria	Fast Food Pizzeria Chain
Exact Sensitivity	0.22	0.25	0.42	0.17	0.28	0.44
Exact PPV	0.61	0.59	0.72	0.62	0.60	0.77
Exact Concordance	0.19	0.21	0.36	0.15	0.23	0.39

Results: Sensitivity to Fast Food SICs and Chain Status

	D&B			InfoUSA		
	Fast Food	Fast Food Pizzeria	Fast Food Pizzeria Chain	Fast Food	Fast Food Pizzeria	Fast Food Pizzeria Chain
Exact Sensitivity	0.22	0.25	0.42	0.17	0.28	0.44
Exact PPV	0.61	0.59	0.72	0.62	0.60	0.77
Exact Concordance	0.19	0.21	0.36	0.15	0.23	0.39

Results: Food Store Concordance by Tract Characteristics

	D&B			InfoUSA		
	All Stores	Supermarket Grocery	Convenience	All Stores	Supermarket Grocery	Convenience
By Income						
Low	0.40 ^b	0.27 ^a	0.29	0.46	0.28	0.38
Middle	0.38	0.17	0.37	0.52	0.35	0.52
High	0.26	0.22	0.30	0.43	0.34	0.41
By Race						
White	0.40	0.26	0.42 ^{ab}	0.53 ^a	0.34	0.53 ^{ab}
Black	0.36	0.19 ^c	0.32 ^c	0.39 ^c	0.23 ^c	0.35
Mixed	0.44	0.33	0.23	0.50	0.31	0.36
By Ethnicity						
Hispanic	0.39	0.33	0.14 ^a	0.51	0.38 ^a	0.24 ^a
Non-Hispanic	0.41	0.25	0.32	0.45	0.25	0.40

Results: Food Store Concordance by Tract Characteristics

	D&B			InfoUSA		
	All Stores	Supermarket Grocery	Convenience	All Stores	Supermarket Grocery	Convenience
By Income						
Low	0.40 ^b	0.27 ^a	0.29	0.46	0.28	0.38
Middle	0.38	0.17	0.37	0.52	0.35	0.52
High	0.26	0.22	0.30	0.43	0.34	0.41
By Race						
White	0.40	0.26	0.42 ^{ab}	0.53 ^a	0.34	0.53 ^{ab}
Black	0.36	0.19 ^c	0.32 ^c	0.39 ^c	0.23 ^c	0.35
Mixed	0.44	0.33	0.23	0.50	0.31	0.36
By Ethnicity						
Hispanic	0.39	0.33	0.14 ^a	0.51	0.38 ^a	0.24 ^a
Non-Hispanic	0.41	0.25	0.32	0.45	0.25	0.40

bridging the gap

Results: Food Store Concordance by Tract Characteristics

	D&B			InfoUSA		
	All Stores	Supermarket Grocery	Convenience	All Stores	Supermarket Grocery	Convenience
By Urbanicity						
Urban	0.34 ^a	0.22 ^a	0.32 ^a	0.47 ^b	0.32	0.44 ^{a,b}
Suburban	0.43 ^c	0.36	0.46	0.51 ^c	0.34	0.52 ^c
Rural	0.29	0.27	0.35	0.32	0.39	0.27

Results: Food Store Concordance by Tract Characteristics

	D&B			InfoUSA		
	All Stores	Supermarket Grocery	Convenience	All Stores	Supermarket Grocery	Convenience
By Urbanicity						
Urban	0.34 ^a	0.22 ^a	0.32 ^a	0.47 ^b	0.32	0.44 ^{a,b}
Suburban	0.43 ^c	0.36	0.46	0.51 ^c	0.34	0.52 ^c
Rural	0.29	0.27	0.35	0.32	0.39	0.27

Results: Food Store Concordance by Tract Characteristics

	D&B			InfoUSA		
	All Stores	Supermarket Grocery	Convenience	All Stores	Supermarket Grocery	Convenience
By Urbanicity						
Urban	0.34 ^a	0.22 ^a	0.32 ^a	0.47 ^b	0.32	0.44 ^{a,b}
Suburban	0.43 ^c	0.36	0.46	0.51 ^c	0.34	0.52 ^c
Rural	0.29	0.27	0.35	0.32	0.39	0.27

Results: Restaurant Concordance by Tract Characteristics

	D&B			InfoUSA		
	All Restaurants	Fast Food Pizza	Fast Food Pizza Chain	All Restaurant	Fast Food Pizza	Fast Food Pizza Chain
By Income						
Low	0.42	0.18	0.47	0.55	0.18 ^b	0.37
Middle	0.45	0.25	0.44	0.57	0.23	0.37
High	0.41	0.19	0.30	0.55	0.26	0.40
By Race						
White	0.43 ^a	0.26 ^{a,b}	0.47	0.59 ^{a,b}	0.24 ^a	0.39
Black	0.36 ^c	0.15	0.42	0.49 ^c	0.11 ^c	0.42
Mixed	0.45	0.18	0.49	0.56	0.20	0.34
By Ethnicity						
Hispanic	0.44	0.15 ^a	0.38	0.53	0.25	0.40
Non-Hispanic	0.42	0.19	0.48	0.55	0.17	0.37

bridging the gap

Results: Restaurant Concordance by Tract Characteristics

	D&B			InfoUSA		
	All Restaurants	Fast Food Pizza	Fast Food Pizza Chain	All Restaurant	Fast Food Pizza	Fast Food Pizza Chain
By Income						
Low	0.42	0.18	0.47	0.55	0.18 ^b	0.37
Middle	0.45	0.25	0.44	0.57	0.23	0.37
High	0.41	0.19	0.30	0.55	0.26	0.40
By Race						
White	0.43 ^a	0.26 ^{a,b}	0.47	0.59 ^{a,b}	0.24 ^a	0.39
Black	0.36 ^c	0.15	0.42	0.49 ^c	0.11 ^c	0.42
Mixed	0.45	0.18	0.49	0.56	0.20	0.34
By Ethnicity						
Hispanic	0.44	0.15 ^a	0.38	0.53	0.25	0.40
Non-Hispanic	0.42	0.19	0.48	0.55	0.17	0.37

bridging the gap

Results: Restaurant Concordance by Tract Characteristics

	D&B			InfoUSA		
	All Restaurants	Fast Food Pizza	Fast Food Pizza Chain	All Restaurants	Fast Food Pizza	Fast Food Pizza Chain
By Urbanicity						
Urban	0.43 ^b	0.21 ^{a,b}	0.36 ^{a,b}	0.55 ^b	0.23 ^b	0.39
Suburban	0.40 ^c	0.04 ^c	0.04 ^c	0.52 ^c	0.01 ^c	0.00
Rural	0.22	0.04	0.05	0.28	0.02	0.02

Results: Restaurant Concordance by Tract Characteristics

	D&B			InfoUSA		
	All Restaurants	Fast Food Pizza	Fast Food Pizza Chain	All Restaurants	Fast Food Pizza	Fast Food Pizza Chain
By Urbanicity						
Urban	0.43 ^b	0.21 ^{a,b}	0.36 ^{a,b}	0.55 ^b	0.23 ^b	0.39 ^{a,b}
Suburban	0.40 ^c	0.04 ^c	0.04 ^c	0.52 ^c	0.01 ^c	0.00
Rural	0.22	0.04	0.05	0.28	0.02	0.02

Fast Food Restaurant Sensitivity by Chain

Chain Name	# On the Ground	# in D&B Fast Food	# in InfoUSA Fast Food
McDonald's	74	48	0
Burger King	33	29	0
Wendy's	23	16	0
KFC	17	10	0
Taco Bell	27	14	0
White Castle	9	3	8
Subway	111	56	87
Quiznos	17	11	16
Jimmy John's	16	6	14
Pizza Hut	21	0	0
Dominos	17	0	0
Little Ceasars	13	0	0

bridging the gap

Fast Food Restaurant Sensitivity by Chain

Chain Name	# On the Ground	# in D&B Fast Food	# in InfoUSA Fast Food
McDonald's	74	48	0
Burger King	33	29	0
Wendy's	23	16	0
KFC	17	10	0
Taco Bell	27	14	0
White Castle	9	3	8
Subway	111	56	87
Quiznos	17	11	16
Jimmy John's	16	6	14
Pizza Hut	21	0	0
Dominos	17	0	0
Little Ceasars	13	0	0

bridging the gap

Results: Fast Food Restaurants by Chain

Chain Name	# On the Ground	# in D&B Fast Food Pizza	# in InfoUSA Fast Food Pizza
McDonald's	74	48	0
Burger King	33	29	0
Wendy's	23	16	0
KFC	17	10	0
Taco Bell	27	14	0
White Castle	9	3	8
Subway	111	56	87
Quiznos	17	11	16
Jimmy John's	16	6	14
Pizza Hut	21	8	9
Dominos	17	6	16
Little Ceasars	13	6	8

bridging the gap

Implications

What we Learned ...

- Previous research by Paquet et al. qualified validation results as Poor (0-0.3), Fair (0.31-0.50), Moderate (0.51-0.7), Good (0.71-0.90) and Excellent (>0.91).

Food Stores:

- Sensitivity was moderate; PPV was moderate to good; Concordance was fair to moderate
- But based on exact match stats were 30-50% lower; overall concordance was fair
- Among store types: Supermarket/Grocery had better sensitivity; Convenience stores had better PPV
- Specialty stores (not shown) had worst agreement but these store types are of less interest in nutrition and obesity-related studies
- Overall agreement higher in InfoUSA compared to D&B
- Important to include Gas station SIC codes to account for convenience stores for both D&B and InfoUSA lists
- Very few differences by tract characteristics which suggests that error is not largely systematic.

What we Learned ...

Restaurants:

- Sensitivity was moderate to good; PPV was moderate to good; Concordance was moderate (better than food stores)
- But based on exact classification match agreement for fast food restaurant was POOR for Both D&B and InfoUSA.
 - Non-chain fast food more likely to be missing from lists and when they were included they were more likely to be listed as full-service
 - InfoUSA only classifies the following SIC codes as limited service: 581206 (food-carry out); 581208 (delicatessens); 581219 (sandwiches); 581229 (deli-bakery); does not include any traditional fast food!
- Including pizzeria SIC codes improved fast food agreement
- Results suggest need to submit list of fast food chains when purchasing the data; but still will not capture informal fast food market
- Fast food agreement higher in D&B; full-service restaurant higher in InfoUSA
- Fast food agreement lower in African American tracts and *substantially lower* in rural tracts for both lists. Also lower in low-income tracts in InfoUSA

bridging the gap