

Identifying Growth Potentials Across Markets and Categories



Wagner Kamakura
Duke University



Two studies

How are we doing?

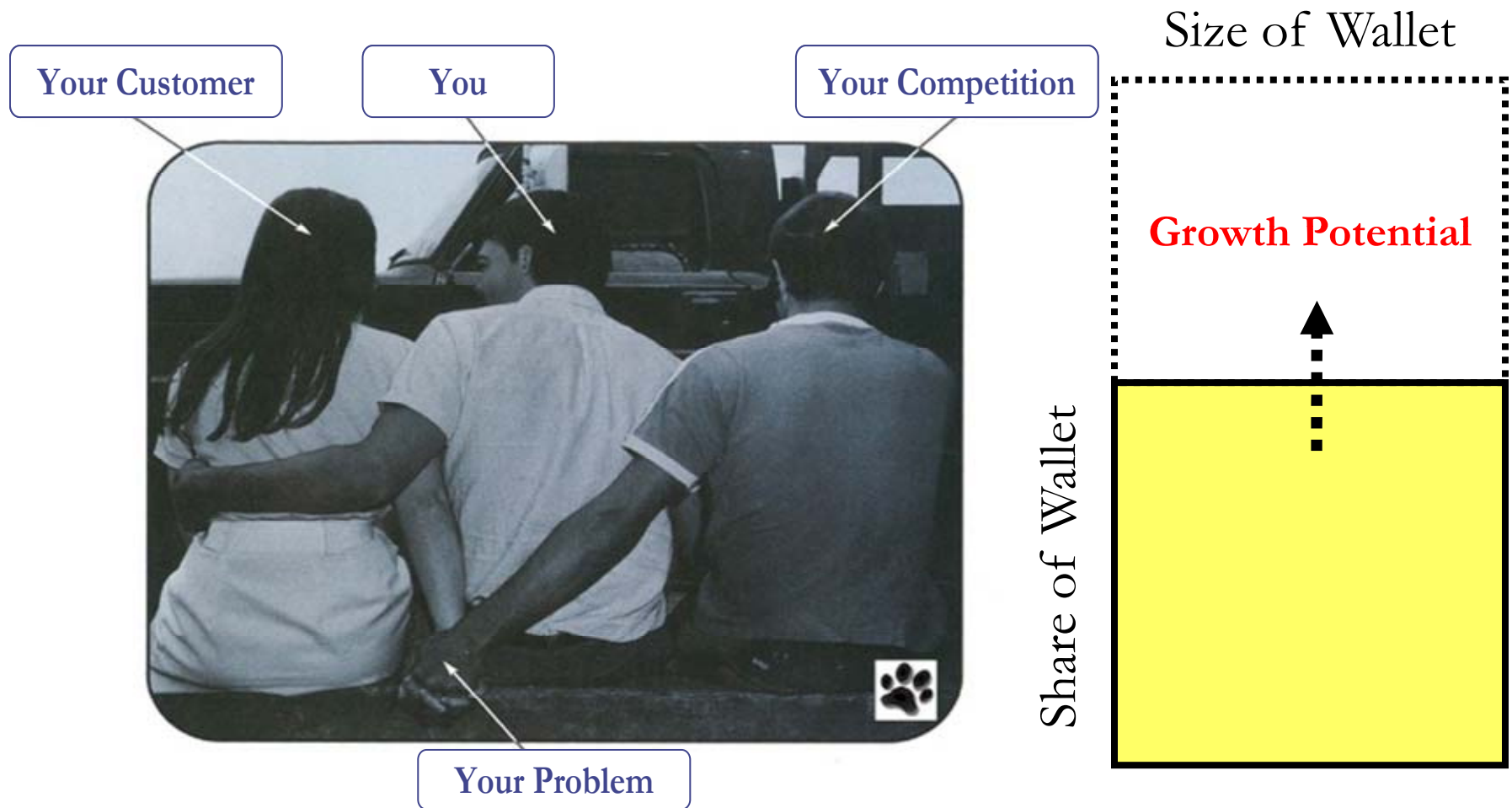
- Relative to our competitors?
 - Share of Wallet
- Relative to our other products and markets?
 - Internal benchmarking

Size and share of Customer Wallet



Du, Kamakura & Mela (*JM*, 2007)

Gauging Growth Potential through Size and Share of Customer Wallet





Customers are polygamous

Our Savings Account

Customer A

\$5,000

Customer B

\$5,000

- What is the customer worth? (Wallet size)
- How can we get a large share of their business (SOW)

What's share of wallet?



Share of garage

What's share of wallet?



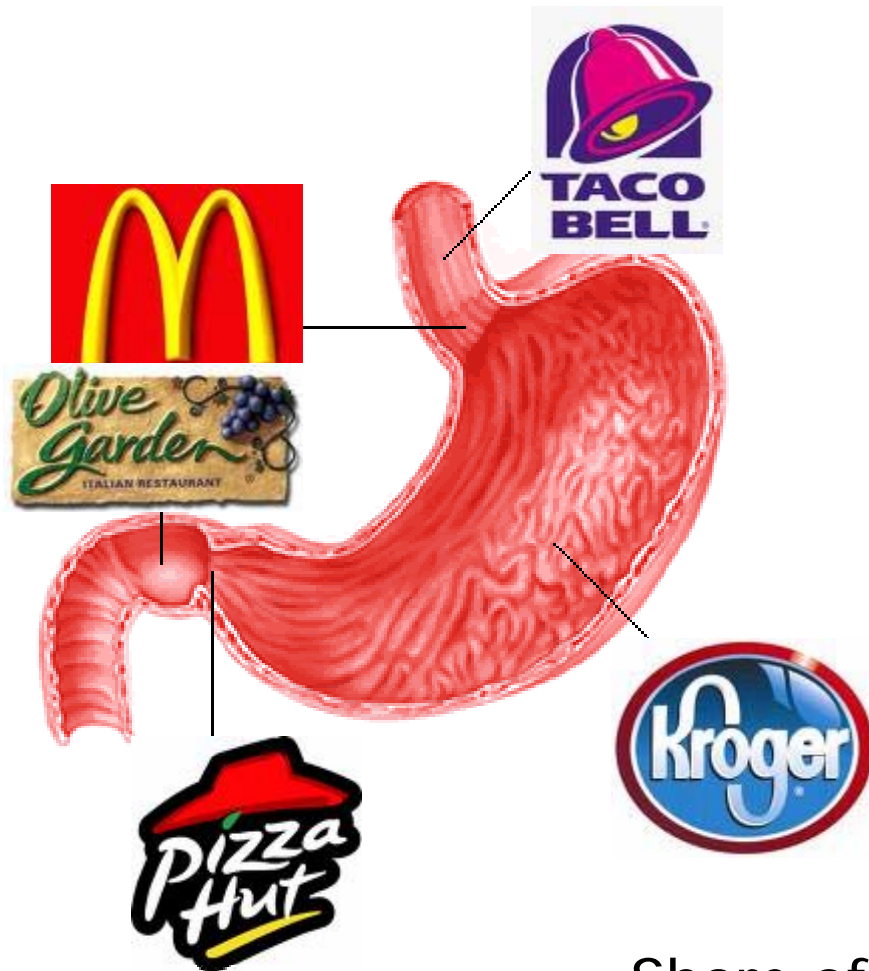
Share of closet

What's share of wallet?



Share of pantry

What's share of wallet?



Share of stomach



Customer Segmentation by wallet size and share

	<u>Low Share</u>	<u>High Share</u>
<u>Large Wallet</u>		
<u>Small Wallet</u>		

Customer Segmentation by wallet size and share

	<u>Low Share</u>	<u>High Share</u>
<u>Large Wallet</u>	High priority aggressive investments	High alert defensive investments
<u>Small Wallet</u>	Low priority exploratory investments	Relationship maintenance

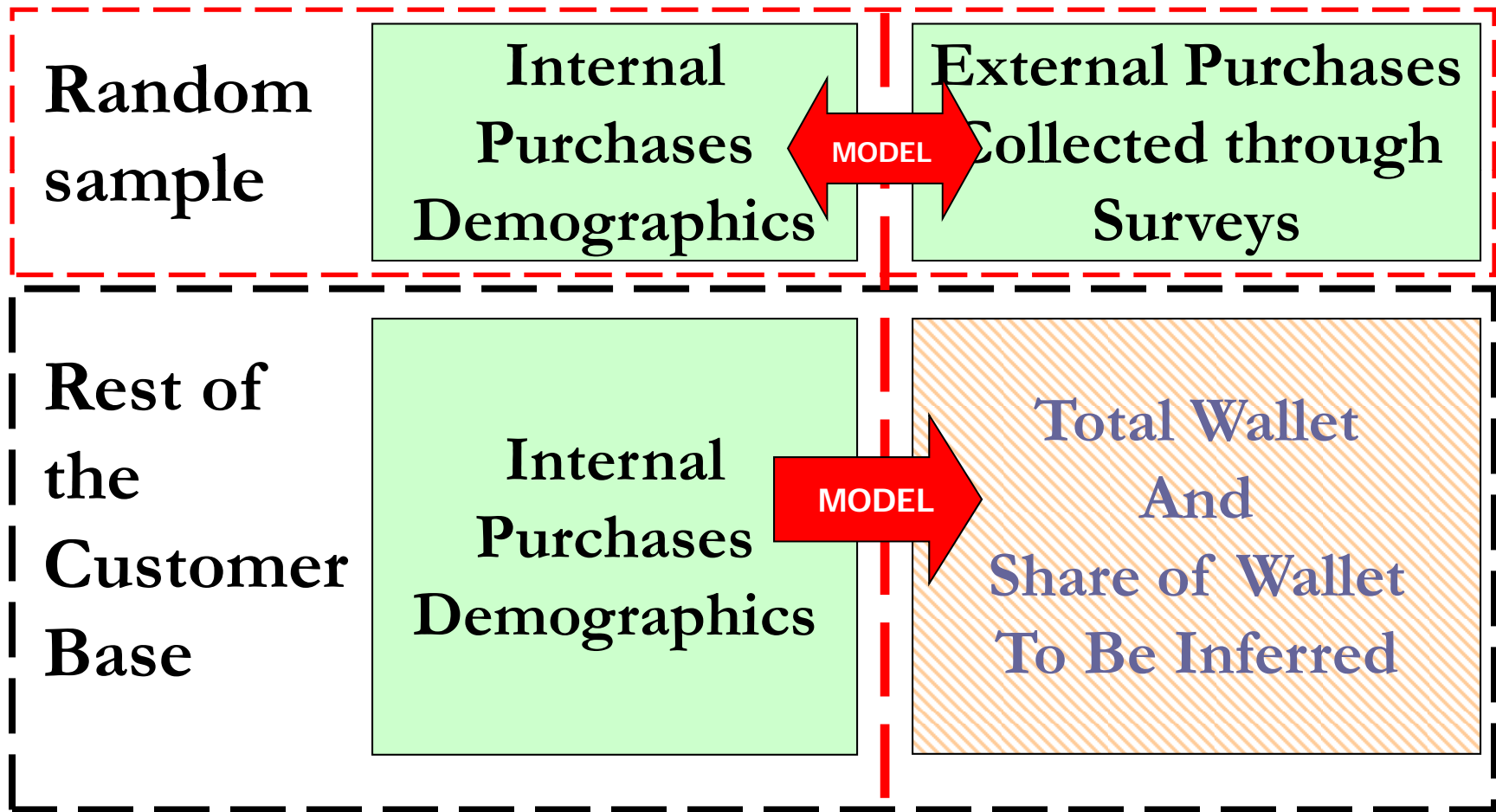
Marketing

Service

Operations

How can we know their wallet size?

Inside the Firm | Outside the Firm





Modeling problem

- What do we know from each customer in the **sample**?
 - Multiple (10) products
 - Wallet size (truncated)
 - Firm's share of this wallet (can be zero)
 - Demographics & Transaction Data
- What do we want to do with the model?
 - Project wallet size and our SOW for each customer in the **customer base** using only the internal data

Calibration Data: Summary Statistics

34,000 customers (database + survey)

	Penetration	Average Balance	Share = 0	0 < Share < 1	Share = 1	Average Share
Non-I Checking	60.4%	\$5,698	29.6%	18.8%	51.5%	54.6%
I-Checking	52.0%	\$10,395	38.9%	17.4%	43.8%	48.1%
Savings	79.6%	\$19,986	43.2%	23.6%	33.2%	36.0%
CDs	19.4%	\$50,017	52.7%	13.3%	34.0%	34.7%
Car loan	46.3%	\$14,105	80.9%	5.9%	13.3%	14.6%
Personal loan	24.8%	\$18,746	81.3%	4.2%	14.5%	16.7%
Line of credit	18.7%	\$10,521	50.0%	7.2%	42.8%	39.9%
Credit card	66.2%	\$4,713	60.8%	26.0%	13.1%	18.1%
Mortgage	52.1%	\$102,595	74.0%	2.0%	24.0%	23.5%
Investments	55.8%	\$111,439	89.0%	6.5%	4.5%	4.7%

Wallet

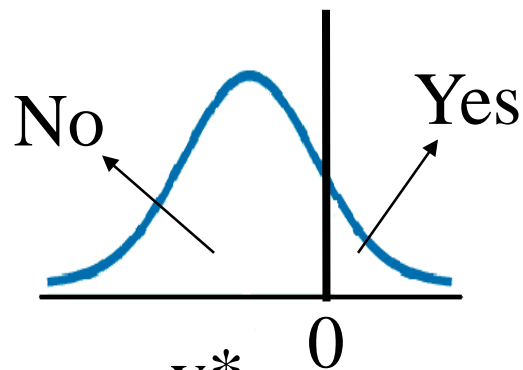
Share of wallet

Model Description – Latent Propensities

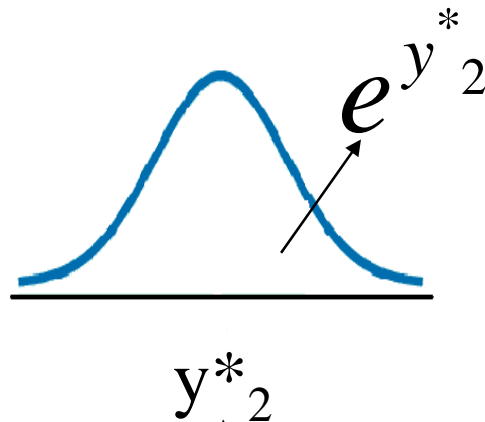
- Customer i
- Product j
- Decision k
- Three propensities:
 - Y_{ij1} = Propensity to own service j
 - Y_{ij2} = Propensity to have a large wallet for j
 - Y_{ij3} = Propensity to use our bank for service j

Model Description – Latent propensities

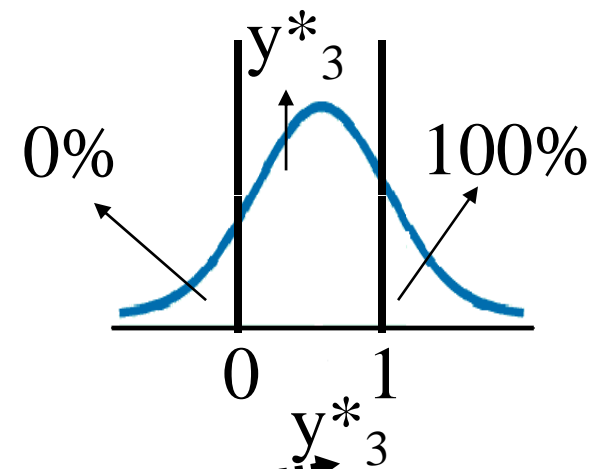
Should I own?



How much \$?



What % at Bank X?



Propensities

Propensities for each product

Product

$$y_{ijOwn}^* = \alpha_{jOwn} + x_i' \beta_{jOwn} + z_i' \gamma_{jOwn} + \varepsilon_{ijOwn}$$

$$y_{ijWallet}^* = \alpha_{jWallet} + x_i' \beta_{jWallet} + z_i' \gamma_{jWallet} + \varepsilon_{ijWallet}$$

$$y_{ijShare}^* = \alpha_{jShare} + x_i' \beta_{jShare} + z_i' \gamma_{jShare} + \varepsilon_{ijShare}$$

Intercepts

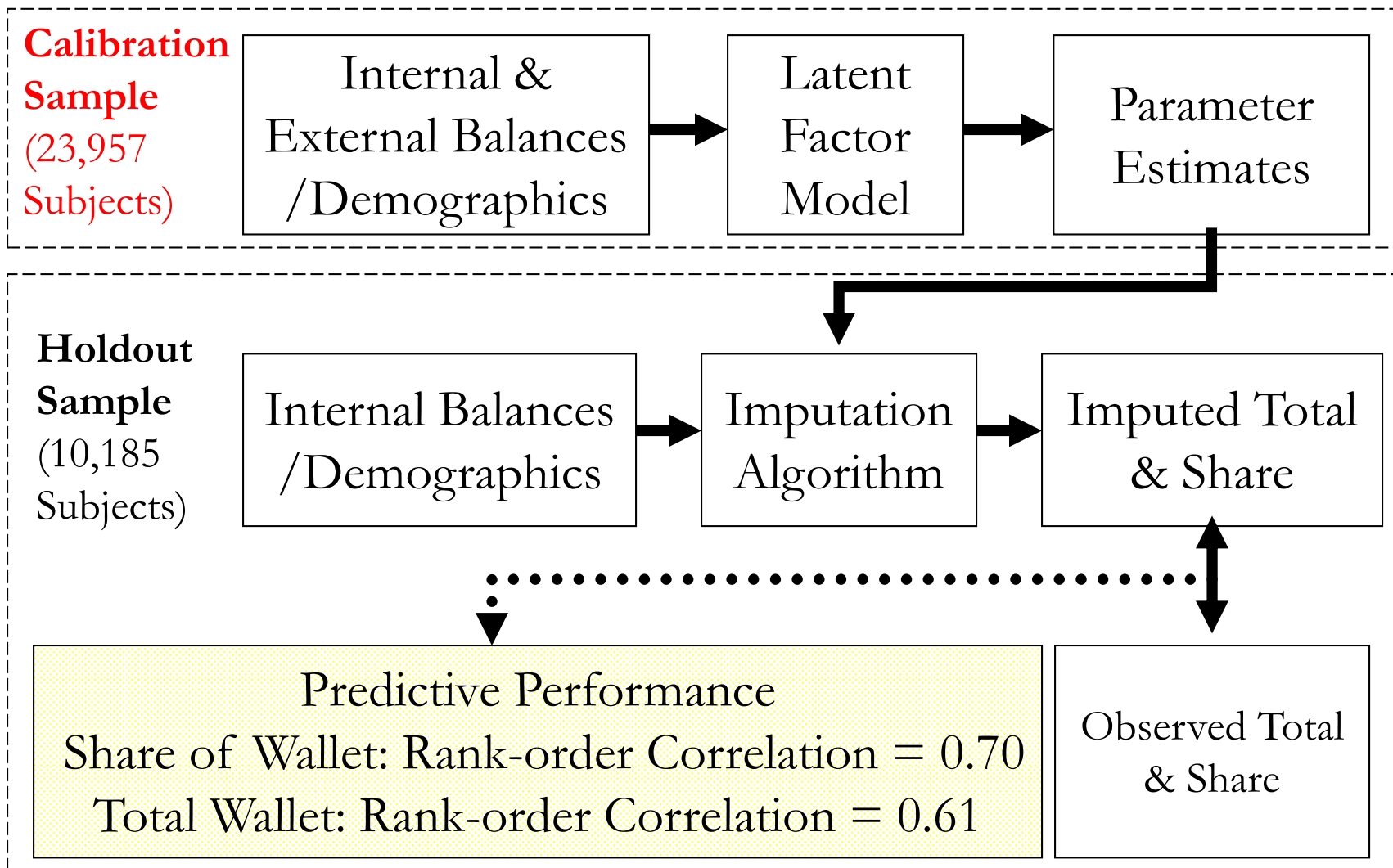
Propensities

Latent
Factors

Customer
Characteristics

Loadings

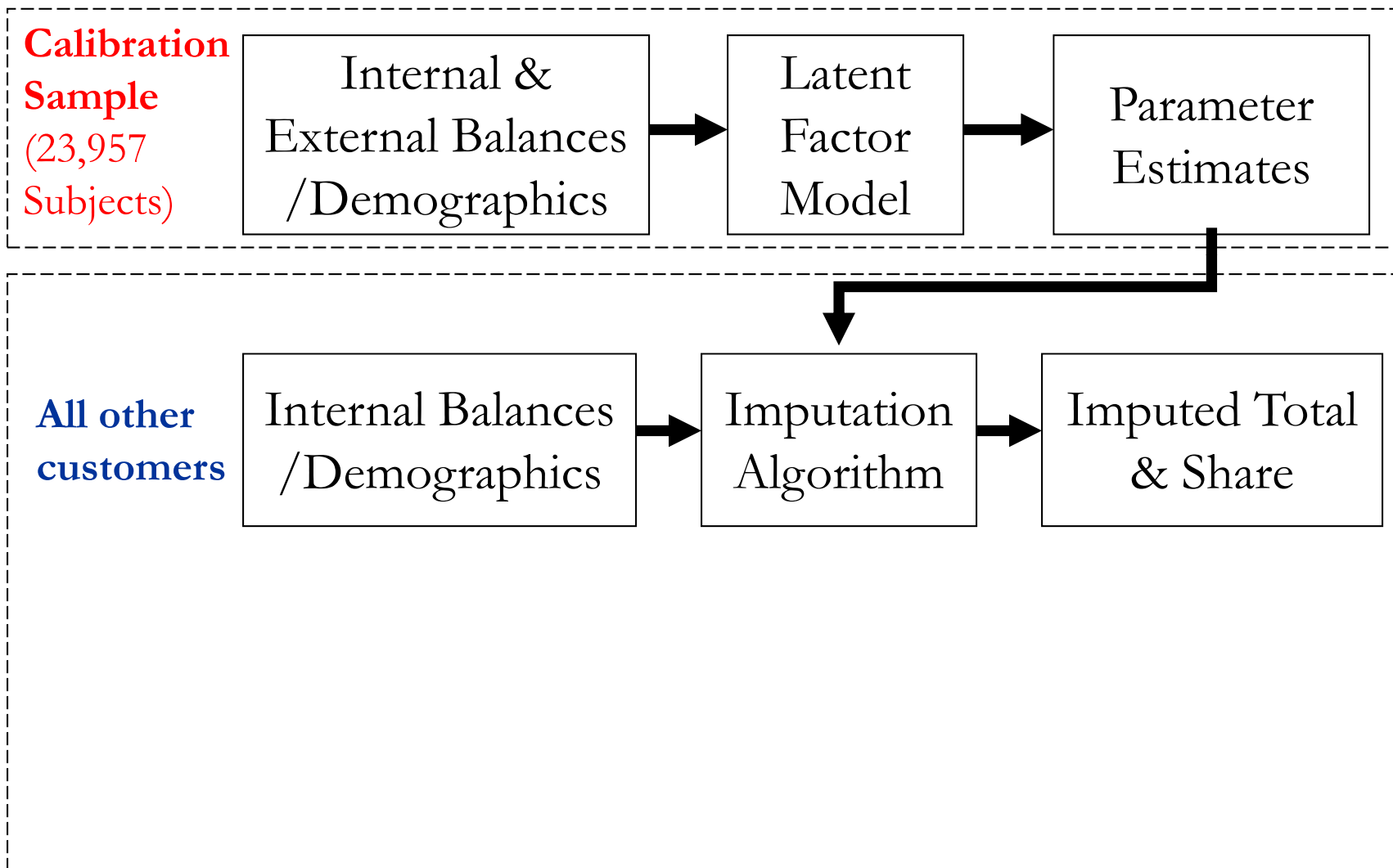
Empirical Implementation



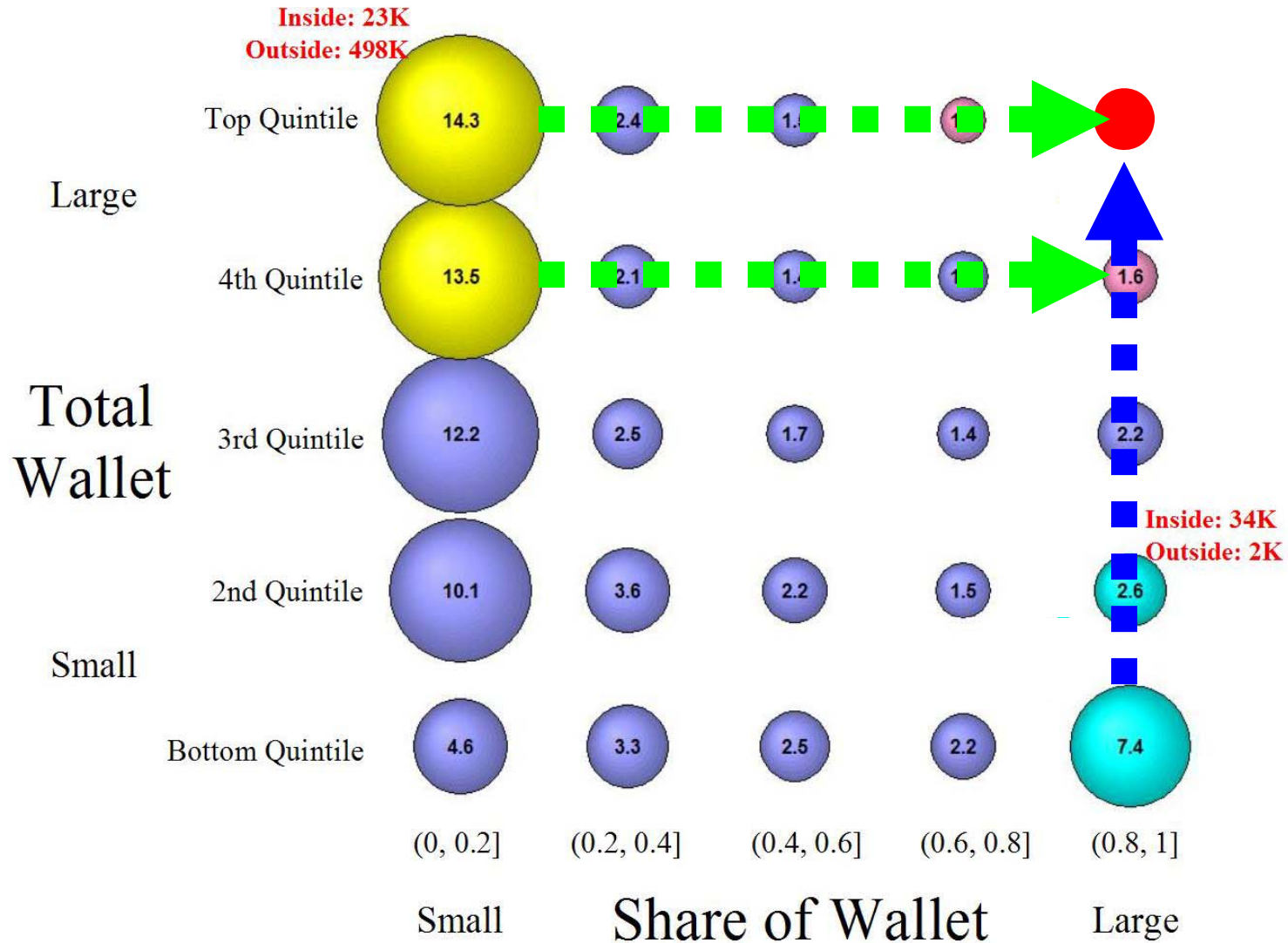
Coefficients for customer characteristics

	Ownership Decision			Total Decision (Conditional on Ownership)			Share Decision (Conditional on Ownership)		
Category	Intercept α_3	$\beta_3(income)$	$\beta_3(tenure)$	Intercept α_2	$\beta_2(income)$	$\beta_2(tenure)$	Intercept α_1	$\beta_1(income)$	$\beta_1(tenure)$
Assets									
Non-interest Checking	0.78	0.07	-0.09	8.01	0.63	0.05	1.04	-0.61	0.56
Interest Checking	0.16	1.12	0.37	7.88	0.76	0.22	0.46	-0.65	0.75
Savings	0.91	0.49	0.06	8.39	0.86	0.38	0.42	-0.54	0.27
CDs	-0.96	0.20	0.24	9.59	0.24	0.29	-0.03	-0.76	0.43
Investments	0.14	0.79	0.09	9.99	1.06	0.49	-2.51	-0.26	0.41
Debts									
Car Loan	-0.09	0.51	-0.20	9.07	0.33	-0.02	-3.64	-0.39	-0.07
Personal Loan	-0.70	0.28	-0.20	9.09	0.29	0.02	-4.75	-1.00	1.68
Line of Credit	-0.96	0.41	0.11	7.88	0.57	0.25	1.18	-1.99	2.01
Credit Card	0.45	0.24	-0.08	7.71	0.41	0.00	-0.31	-0.14	0.19
Mortgage	0.03	0.78	0.11	10.79	0.51	-0.15	-10.53	-0.40	-0.61

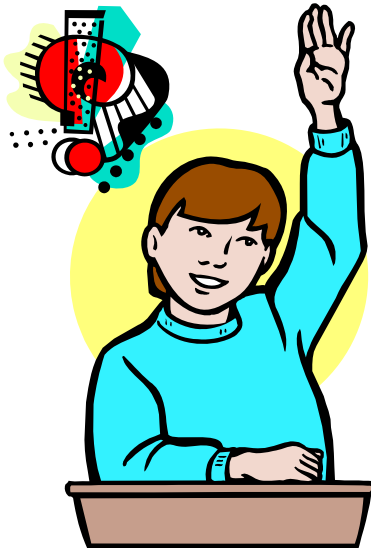
Empirical Implementation



Customer segmentation by wallet size and share of wallet

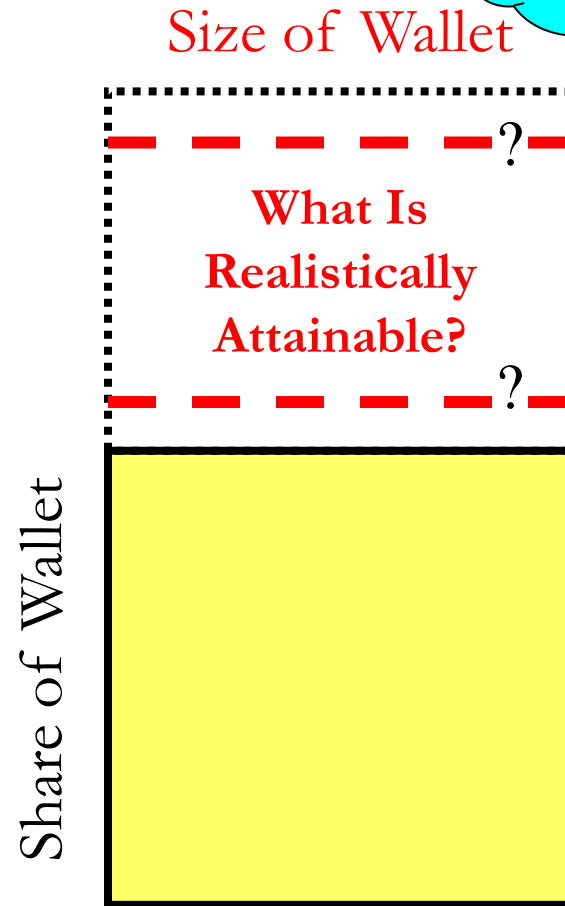
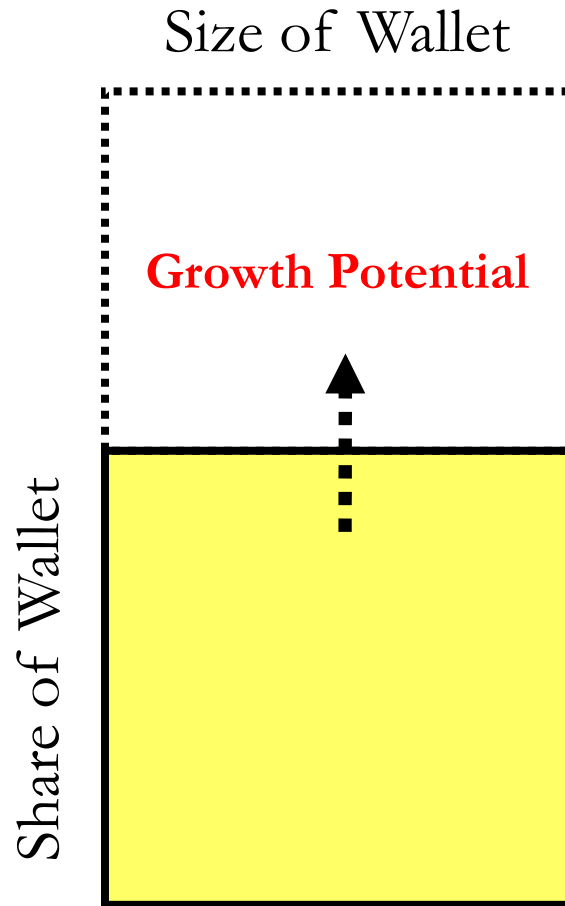


Questions and Comments on SOW?



Limitations of the Share of Wallet Approach

We don't know this!





Conclusions on SOW

- Highly valuable if you know the customer's wallet size
 - Imputations based on a model calibrated on a sample with full information
- In some industries the wallet size is known
 - Credit cards
 - Pharmaceutical drugs
 - Automobiles
- Even when you know the customers' wallet size
 - 100% wallet size might be unrealistic
 - What SOW is attainable?
- What to do if we don't know the wallet size?

Identifying Growth Potentials through Internal Benchmarking across Markets and Categories



Rex Du

University of Houston

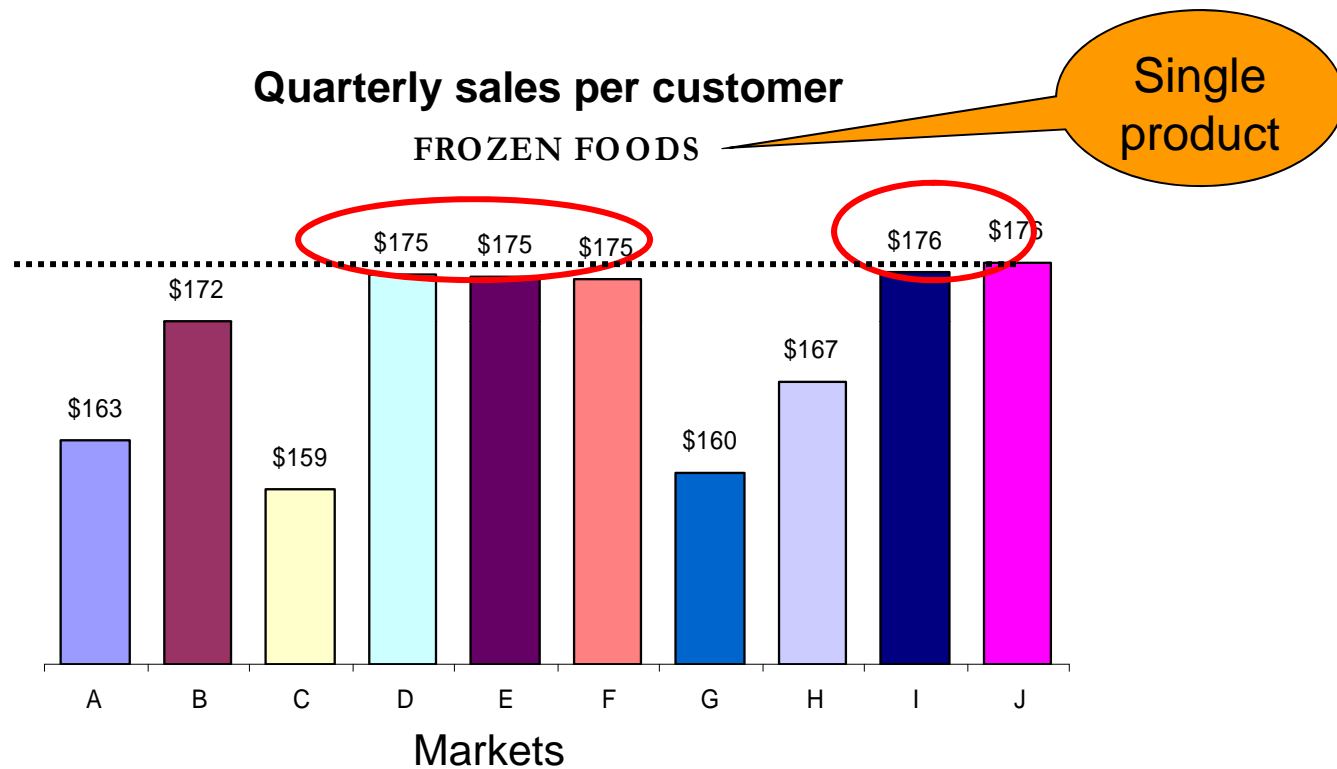
Wagner Kamakura

Duke University

We know how much we have sold, BUT ...,
 how much better could we have done?

CATEGORY / MARKET	BAKERY	COSMETICS	DELI	DRUG	FLORAL	FROZEN FOODS	BABY CARE	BABY FOODS
A	27	6	86	65	12	163	14	18
B	44	15	120	81	23	172	22	22
C	45	13	117	71	22	159	19	23
D	30	10	46	65	13	175	20	24
E	37	13	75	62	14	175	18	20
F	24	9	46	75	10	175	19	20
G	30	9	49	69	11	160	15	14
H	33	8	59	57	12	167	16	26
I	32	11	53	69	11	176	24	29
J	37	13	76	71	18	176	15	19

Gauging Growth Potential through Internal Benchmarking across Markets



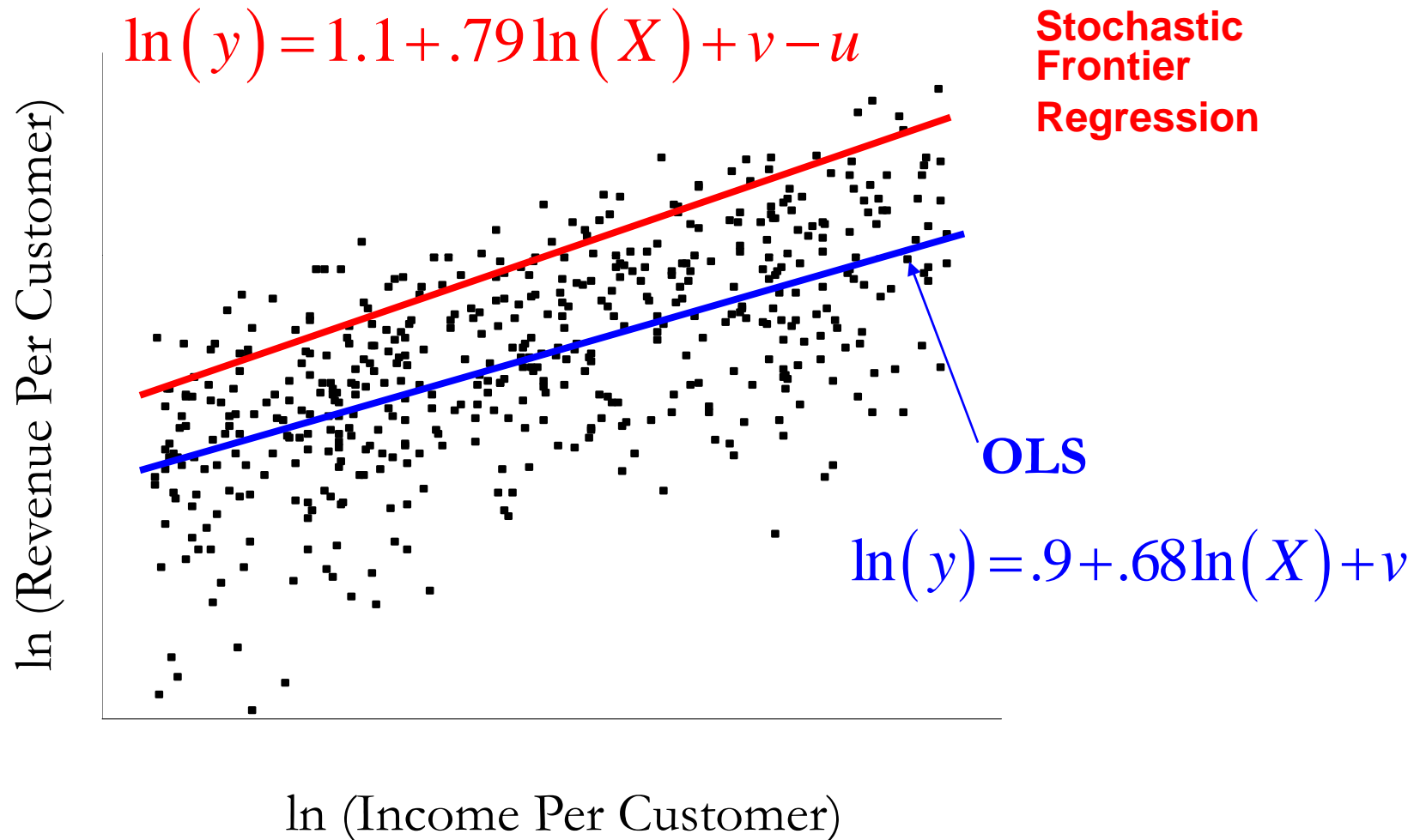
Need to Control for Environmental Factors



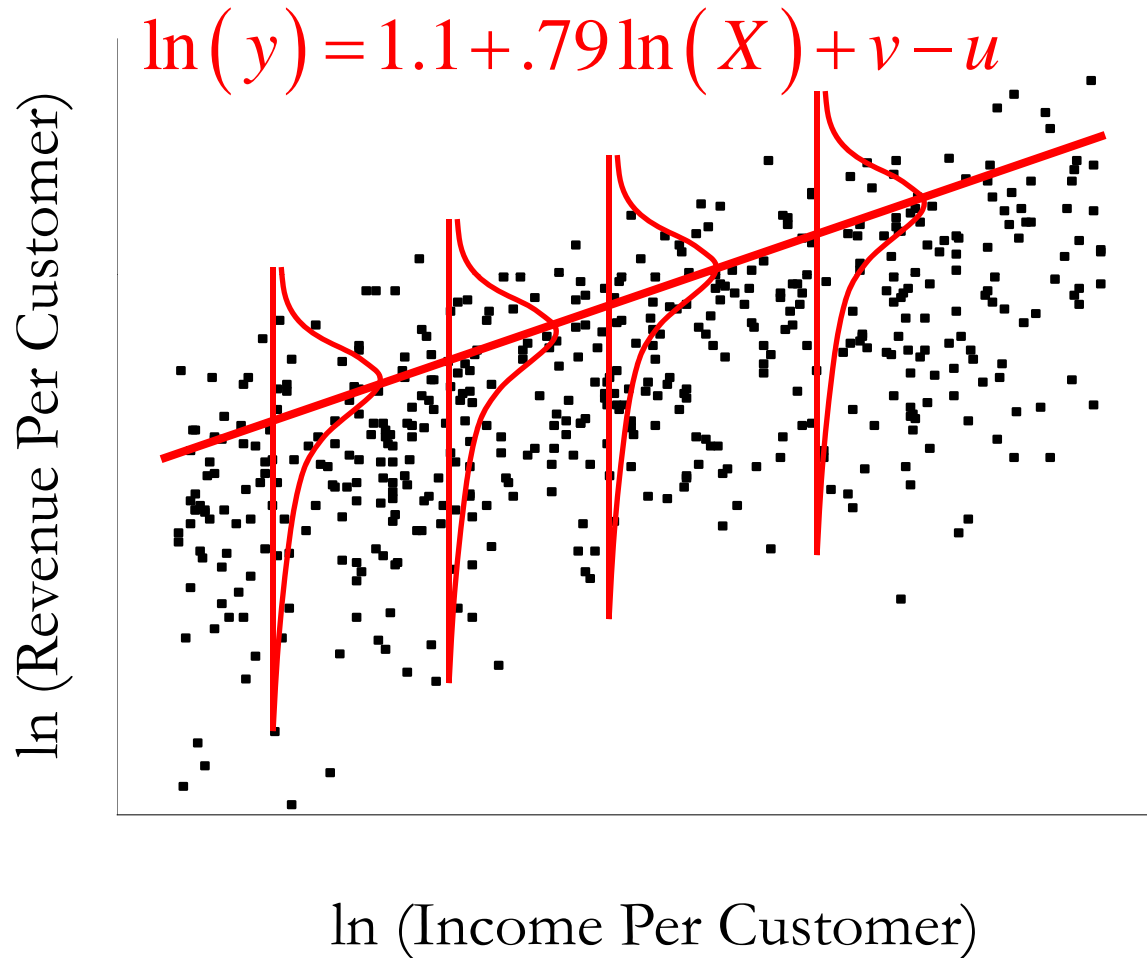
Internal Benchmarking

- *“Identify performance standards and areas for improvement by comparison among similar operations within one’s own organization”* (Camp 1995)
- Advantages of Internal over External Benchmarking
 - Ease of Data Collection
 - Managerial Buy-in
 - Realistic Goal-setting
 - Transfer of Best Practices

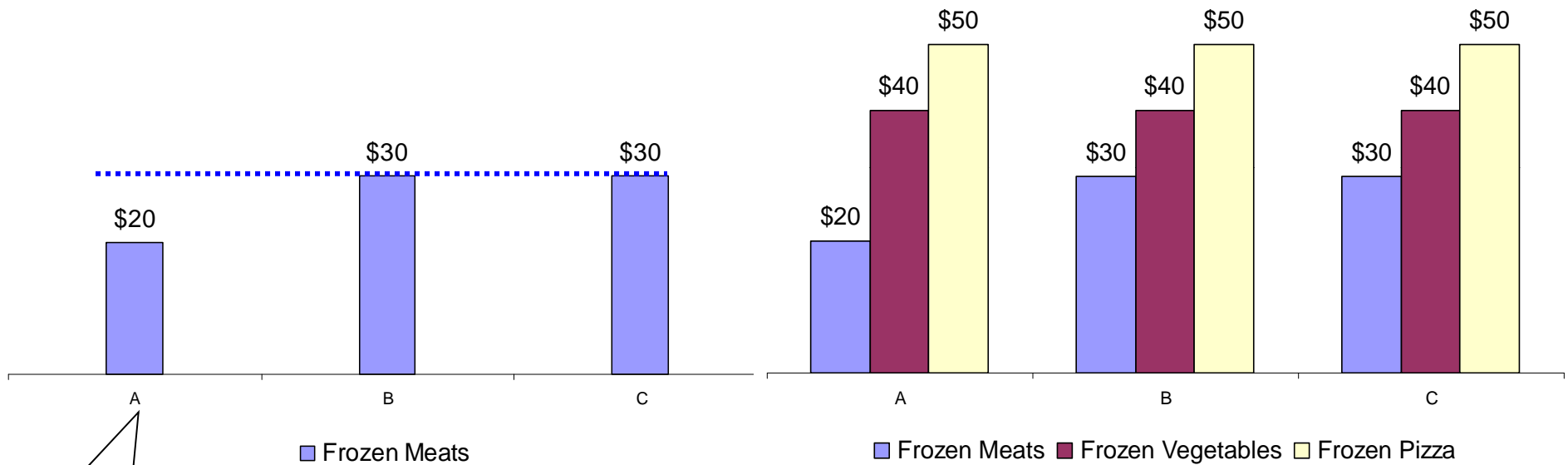
Controlling for Environmental Factors



Stochastic Frontier Regression



Internal Benchmarking across Markets and Categories



Markets

Single Category Across Markets

Across Markets & Categories

Don't Just Compare Apples to Apples, Compare Apples to Oranges As Well

CATEGORY / MARKET	BAKERY	COSMETICS	DELI	DRUG	FLORAL	FROZEN FOODS	BABY CARE	BABY FOODS
A	27	6	86	65	12	163	14	18
B	44	15	120	81	23	172	22	22
C	45	13	117	71	22	159	19	23
D	30	10	46	65	13	175	20	24
E	37	13	75	62	14	175	18	20
F	24	9	46	75	10	175	19	20
G	30	9	49	69	11	160	15	14
H	33	8	59	57	12	167	16	26
I	32	11	53	69	11	176	24	29
J	37	13	76	71	18	176	15	19

In other words ...



INPUT



OUTPUT

Our Proposed Model – Stochastic Frontier Factor Analysis

Revenue per customer in
market \underline{m} for category \underline{c}

**P unobserved environmental
factors in market m**

$$\ln(y_{mc}) = \beta_{0c} + \sum_{k=1}^K \beta_{kc} \ln(X_{km}) + \sum_{p=1}^P \gamma_{pc} Z_{pm} + v_{mc} - u_{mc}$$

K observed
environmental factors in
market m

 $N(0,1)$
 $N(0, \sigma_{uc})$

 $N(0, \sigma_{vc})$

Stochastic Frontier Regression Analysis

Revenue per customer in market \underline{m} for category \underline{c}

$$\ln(y_{mc}) = \beta_{0c} + \sum_{k=1}^K \beta_{kc} \ln(X_{km}) + \sum_{p=1}^P \gamma_{pc} Z_{pm} + v_{mc} - u_{mc}$$

\downarrow
 \uparrow
K observed
environmental factors in
market m

$|N(0, \sigma_{uc})|$
 \downarrow
 $N(0, \sigma_{vc})$

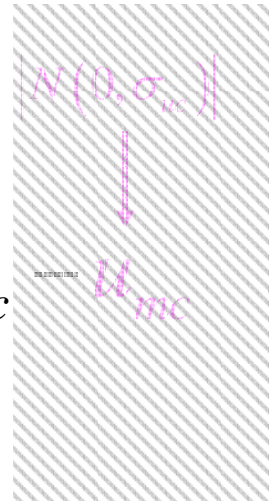
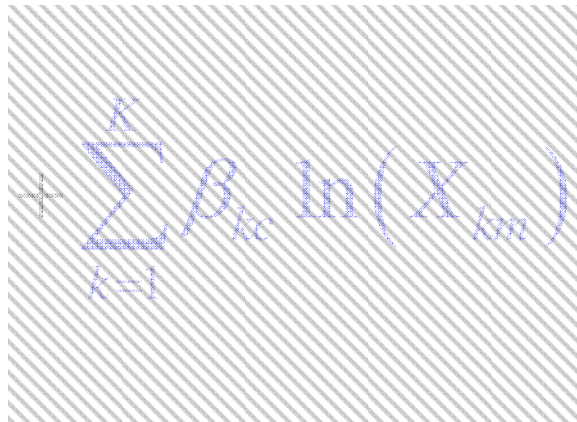
Latent Factor Analysis

Revenue per customer in market \underline{m} for category \underline{c}



$$\ln(y_{mc}) = \beta_{0c} + \sum_{k=1}^K \beta_{kc} \ln(X_{km}) + \sum_{p=1}^P \gamma_{pc} Z_{pm} + v_{mc}$$

P unobserved environmental factors in market m



$N(0, \sigma_{vc})$

Measure Relative Sales Efficiency

$$\varepsilon_{mc} \equiv v_{mc} - \mu_{mc} = \ln(y_{mc}) - \left[\beta_{0c} + \sum_{k=1}^K \beta_{kc} \ln(X_{km}) + \sum_{p=1}^P \gamma_{pc} Z_{pm} \right]$$

$$E_{mc} \equiv y_{mc} / y_{mc}^* = \exp(-u_{mc}) / \exp(-E[u_{mc} | \varepsilon_{mc}])$$

$$E[u_{mc} | \varepsilon_{mc}] = \frac{\sigma_c \lambda_c}{1 + \lambda_c^2} \left[\frac{\phi(a_{mc})}{1 - \Phi(a_{mc})} - a_{mc} \right]$$

Efficiency
measure

$$\lambda_c \equiv \frac{\sigma_{uc}}{\sigma_{vc}}$$

$$\sigma_c^2 \equiv \sigma_{vc}^2 + \sigma_{uc}^2$$

$$a_{mc} \equiv \frac{\lambda_c \varepsilon_{mc}}{\sigma_c}$$

~0 symmetric
>0 frontier

Empirical Illustration

- Major supermarket chain in the U.S.
 - 48 product categories
 - 298 ZIP Code Tabulation Areas in a Southeastern state

- Geodemographic information
 - # of households
 - median age of household head
 - median household size
 - median household income
 - % of white residents
 - % of residents in 5 age groups

- Competitive information
 - Addresses of 1,500+ rival outlets of 3 discount store chains, 4 drug store chains, 2 supercenter chains, 11 supermarket chains



Cross-category Correlations

- COSMETICS and HAIR CARE
 - (0.83)
- LAUNDRY DETERGENTS and BATH TISSUES
 - (0.73)
- BABY CARE and BABY FOODS
 - (0.71)
- BEERS and WINES
 - (0.45)
- DAIRY PRODUCTS and SALADS
 - (0.81)
- FLORAL and DELI
 - (0.72)
- WINES and DELI
 - (0.71)
- TOBACCO and BEEF
 - (0.37)

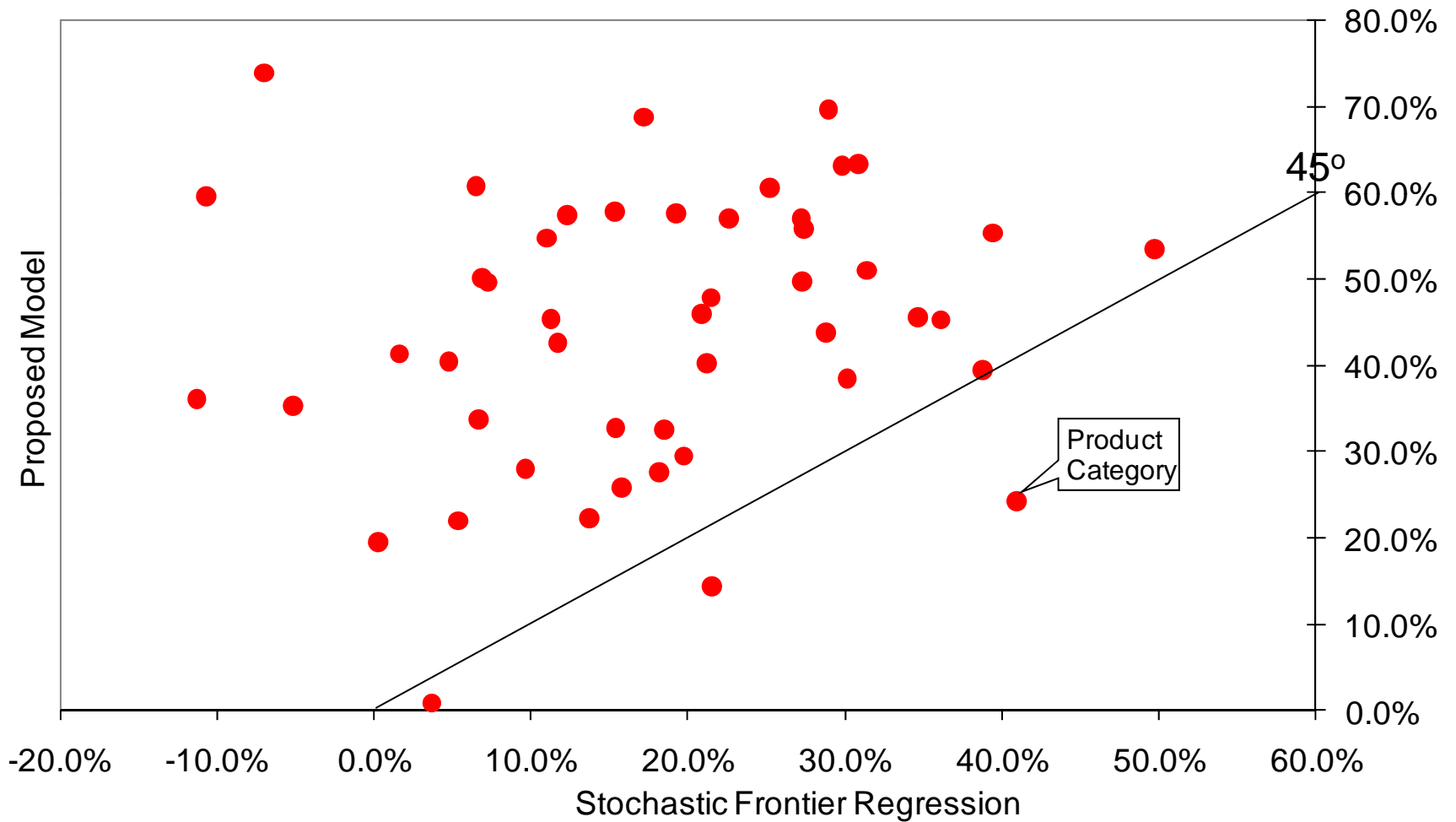
Parameter Estimates

- Median household income (+)
 - Larger for less essential categories
 - Negative for
 - Salad dressing
 - Shelf vegetables
 - Potatoes
 - Packaged meats, sausage & pork
- Median household size (+)
 - Smaller for less essential categories
 - Large negative for wines, nutritional supplements
- 5 latent factors
 - One factor loads on general merchandises; one on non-food categories; one on food categories

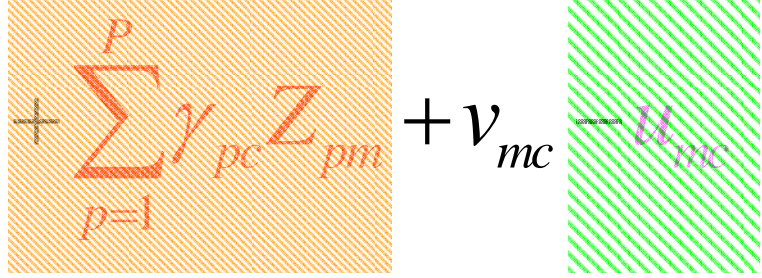
$$\lambda_c \equiv \frac{\sigma_{uc}}{\sigma_{vc}} > 0 \quad \text{for 40 out of 48 categories}$$

Out-of-Sample Predictive Test

% of Variance Explained



Model Comparison

$$\ln(y_{mc}) = \beta_{0c} + \sum_{k=1}^K \beta_{kc} \ln(X_{km}) + \sum_{p=1}^P \gamma_{pc} Z_{pm} + v_{mc} - U_{mc}$$


Variance explained

Stochastic frontier factor analysis
(80.7%)

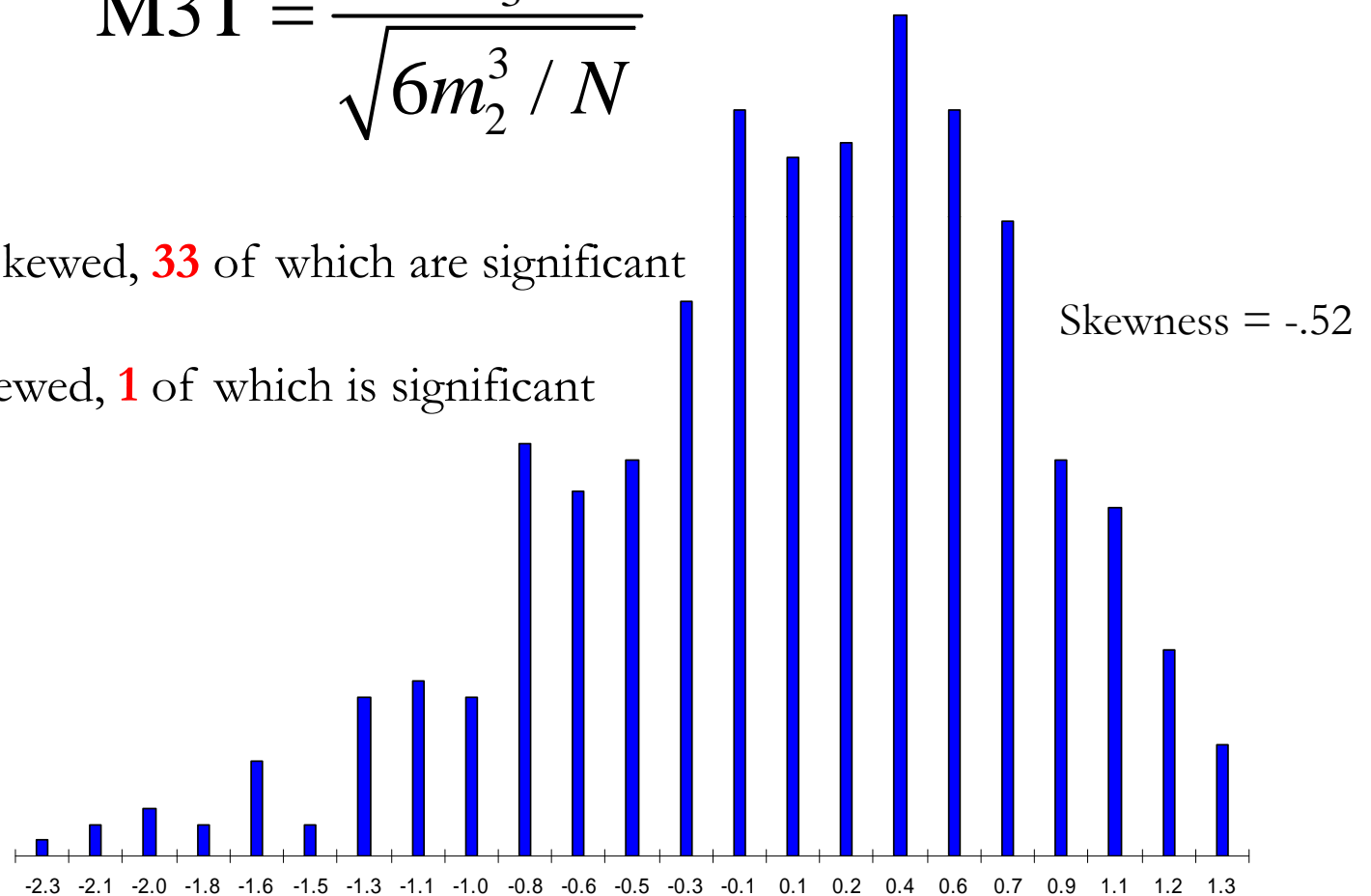
- Stochastic frontier regression (23.2%)
- Latent factor analysis (66.3%)

Validation Test 1: Skewness of OLS Residuals

$$M3T = \frac{m_3}{\sqrt{6m_2^3 / N}}$$

42 are negatively skewed, **33** of which are significant

6 are positively skewed, **1** of which is significant



Validation Test 2: Robustness to unobservables

Regress log **FOOD** products sales efficiencies against competitiveness measures

Stochastic frontier factor analysis

Our model

shortest distance to a discounter	0.00004
shortest distance to a drugstore	-0.00041
shortest distance to a supercenter	0.00043
shortest distance to a competing supermarket	0.00006
shortest distance to one of our stores	-0.00074
food retailing share of shelf space	0.00124

$$R^2 = .09$$

Stochastic frontier regression

Our benchmark

shortest distance to a discounter	0.00072
shortest distance to a drugstore	0.00131
shortest distance to a supercenter	0.00150
shortest distance to a competing supermarket	0.00275
shortest distance to one of our stores	-0.00634
food retailing share of shelf space	0.00659

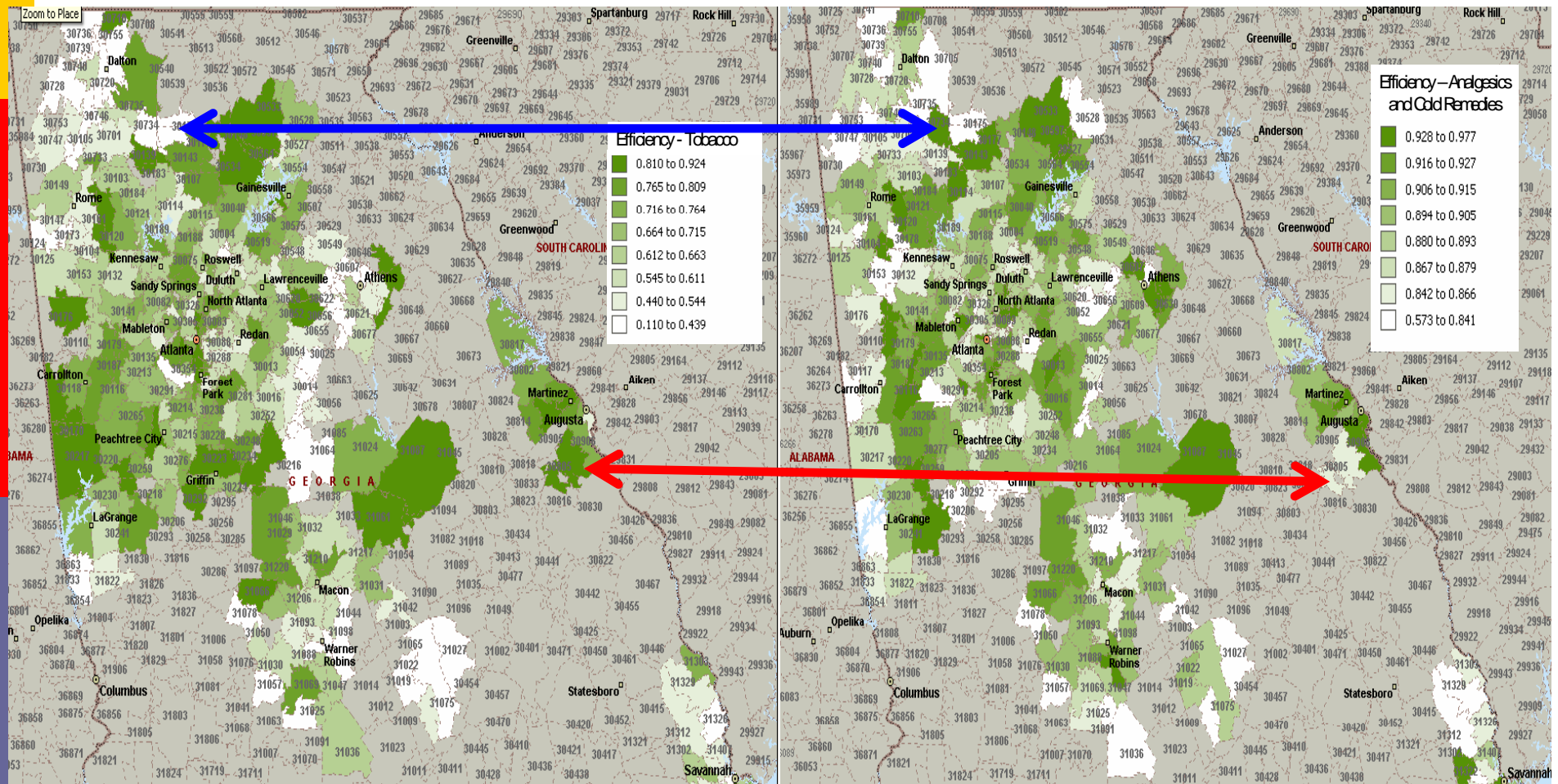
$$R^2 = .25$$

Managerial Implications

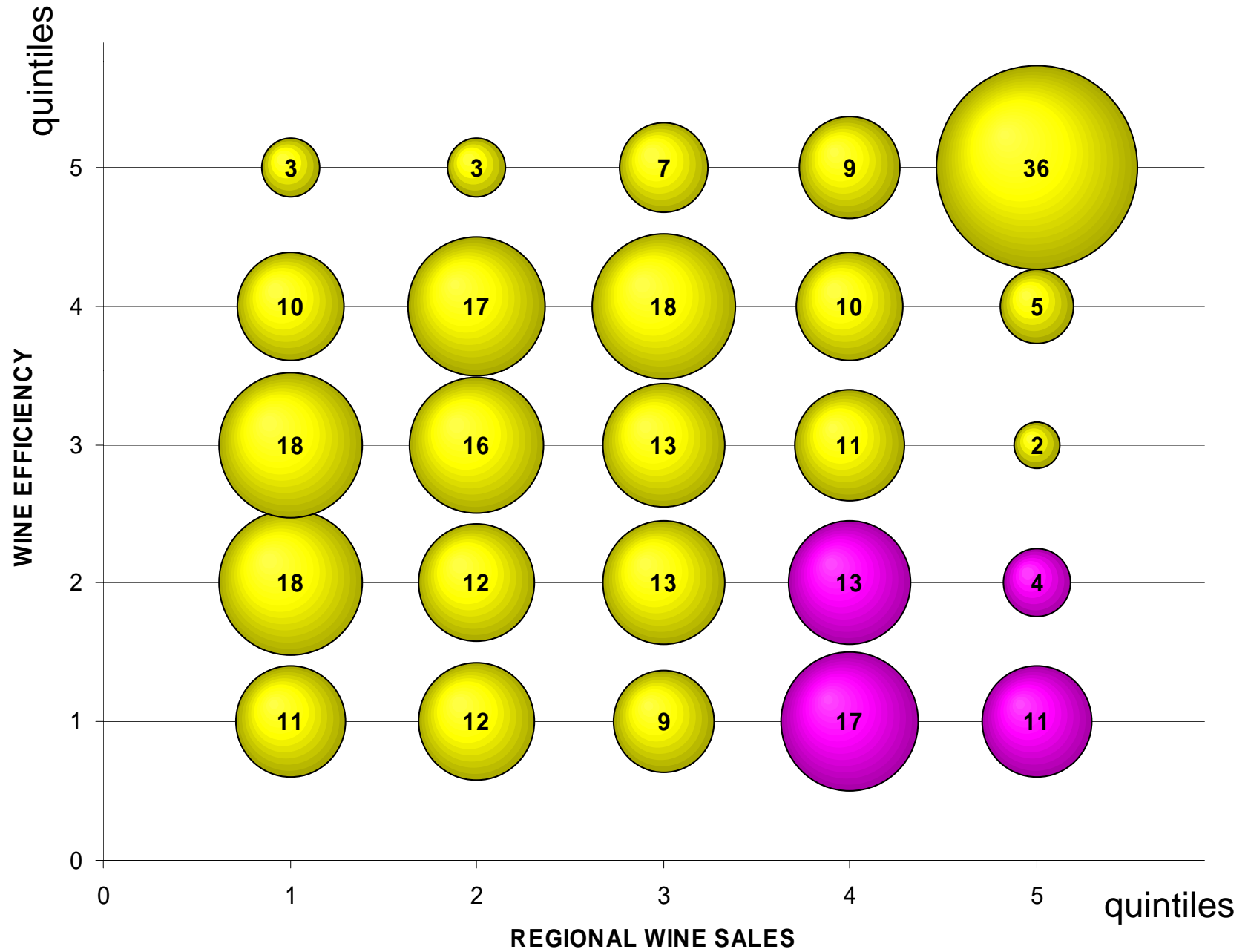


Internal Benchmarking at the
category and cross-category
levels

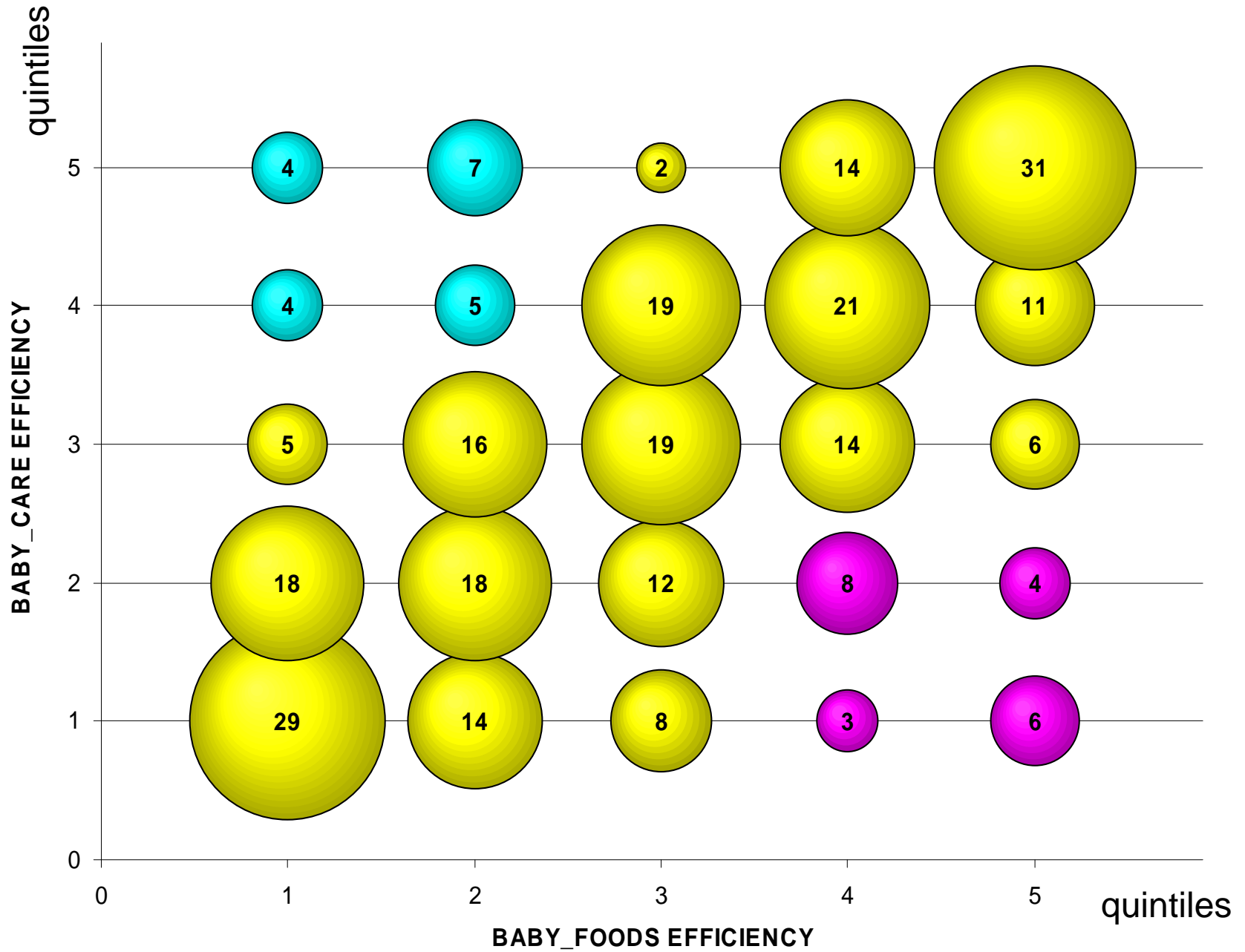
Relative Sales Efficiency Indexes



REGIONAL WINE SALES vs. WINE EFFICIENCY



BABY_CARE vs. BABY FOOD EFFICIENCIES



Closing the Efficiency Gap with Internal Benchmarking

- Identify key levers influencing customers' spending
- Compare and contrast the practices of top performers with those of the bottom performers
- Design market tests to pinpoint the root causes of performance gaps
- Transfer the best practices of the top performers internally and monitor the results



Conclusions

□ Share of Wallet

- How are we doing relative to the competition?
- Clear goal: highest share of the largest wallets
- Limitations
 - Hard to measure wallet size
 - May lead to unrealistic goals

□ Internal benchmarking

- How well are we doing, relative to other products and markets?
- Benefits
 - Identifies “best practices” to be emulated
 - Defines attainable goals
- Limitations
 - Limited to internal know-how
 - Efficiency is not necessarily optimal



Mythos
Griechen