Semiparametric Estimation of Preferences from Diversion Ratios

Chris Conlon (NYU Stern & NBER) and
Julie Holland Mortimer (Boston College & NBER) and
Paul Sarkis (Boston College)

June 19, 2022

Motivation

Diversion Ratios

The diversion ratio is one of the best ways we have to measure competition between products.

- Raise the price of j and count the number of consumers who leave
- The diversion ratio D_{jk} is the fraction of leavers who switch to the substitute k.
- A higher value of D_{jk} indicates closer substitutes.
- Useful because it enters the multi-product Bertrand FOC:

$$\underbrace{p_j\left(1 + 1/\epsilon_{jj}(\mathbf{p})\right)}_{\text{Marginal Revenue}} = c_j + \sum_{k \in \mathcal{J}_f \setminus j} (p_k - c_k) \cdot D_{jk}(\mathbf{p}).$$

- $D_{jk} \equiv \frac{\partial q_k}{\partial p_j} / \left| \frac{\partial q_j}{\partial p_j} \right|$.
- Can also write as $D_{jk} \equiv \frac{\epsilon_{kj}}{|\epsilon_{jj}|} \cdot \frac{q_j}{q_k}$

Diversion Ratios: Conlon Mortimer (RJE 2021)

Our previous work shows most discrete-choice models yield the following representation:

$$D_{jk}^{z_j \to z_j'}(x) = \int D_{jk,i}(x)\, w_i(z_j, z_j', x)\, dF_i$$

- Different interventions $z_j \to z_j'$ (prices, quality, characteristics, assortment) give different weights $w_i(z_j, z_j', x)$ and thus different local average diversion ratios.
- Individual Diversion Ratios $D_{jk,i}(x)$ don't vary with the intervention (only determine how i ranks 2nd and 3rd choices).

In Conlon Mortimer (RJE 2021) we show:

- For any (mixed) logit $D_{jk,i}(x) = \frac{s_{ik}}{1-s_{ij}}$
- For second-choice data, $w_i(x) = \frac{s_{ij}(x)}{s_j(x)}$.
- General LATE interpretation: $w_i\left(z_j, z_j', x\right) = \frac{q_{ij}\left(z_j',x\right)-q_{ij}(z_j,x)}{q_j(z_j,x)-q_j\left(z_j',x\right)}$

What is our paper about?

We consider a problem where we observe some aggregate shares $\mathcal{S} = [\mathcal{S}_1, \ldots, \mathcal{S}_J]$ or sales $\mathcal{Q}_j$, and some elements $(j,k) \in \text{OBS}$ of $\mathcal{D}^T$ a matrix of (second-choice) diversion ratios.

$$
\mathcal{D}^T = \begin{array}{c c}
\begin{array}{ccccc} \text{VZ} & \text{ATT} & \text{TMo} & \text{S} & \text{Other} \end{array} & \\
\begin{pmatrix}
0 & ? & 0.30 & 0.30 & ? \\
? & 0 & 0.45 & 0.15 & 0 \\
? & ? & 0 & 0.45 & ? \\
? & ? & 0.20 & 0 & ? \\
? & ? & 0.05 & 0.10 & 0
\end{pmatrix} & \begin{array}{c} \text{VZ} \\ \text{ATT} \\ \text{TMo} \\ \text{S} \\ \text{Other} \end{array}
\end{array}, \quad
\begin{bmatrix} 0.35 \\ 0.30 \\ 0.20 \\ 0.15 \\ 0.05 \end{bmatrix} = \mathcal{S}
$$

Can we fill in the missing elements?

How do we fill in missing elements?

Typical Approach: estimate a parametric model.

- Multi-product demand with unrestricted matrices of $(J+1)^2$ cross-elasticities (such as AIDS) is often hopeless with large J. Unrestricted diversion likely equally hopeless.
- Plain logit places strong restrictions: $D_{jk} = \frac{s_k}{1-s_j}$.
- Nested logit $D_{jk} = \frac{s_{k|g}}{Z(\sigma, s_g) - s_{j|g}}$ (same nest) where σ is nesting parameter.
- Mixed Logit: Explain substitution patterns using observed characteristics
    - Typically assume independent normal RC
    - Two products with similar x_1 and high substitution $\rightarrow$ larger σ_1.
    - Two products with similar x_2 and low substitution $\rightarrow$ smaller σ_2.
- McFadden and Train (2000) show a mixed logit $u_{ij} = \beta_i x_j + \varepsilon_{ij}$ is fully flexible
    1. This depends on $f(\beta_i)$ heterogeneity being nonparametric
    2. And a sufficient set of characteristics X to explain $\mathcal{D}$

    Much work on (1), not as much attention to (2).

How do we fill in missing elements?

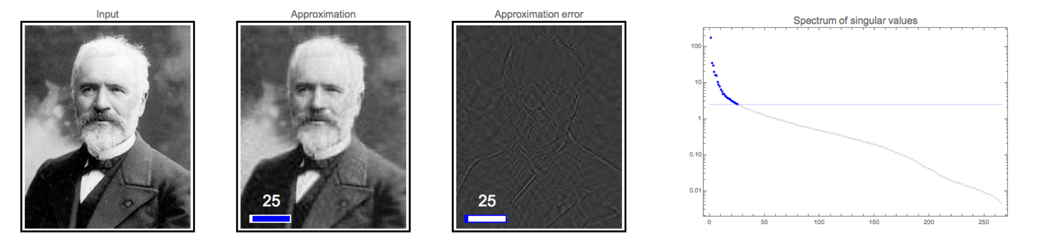
Our paper: Consider a low-rank approximation to $\mathcal{D}$

- Limit the rank of $\mathcal{D}$ directly in product space instead of controlling complexity with product characteristics and parametric restrictions on random coefficients.
- Allow for sparsity in indiv. shares and substitution patterns, with possibility of generating extreme patterns for top substitutes if necessary.

Works well in other domains (CS for image recovery/compression), and we show it has a sensible economic interpretation.

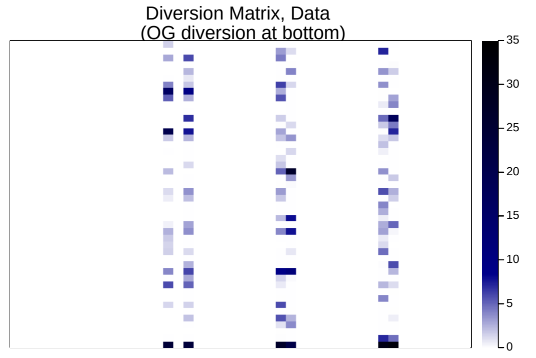
Low Rank Approximations: Image Compression

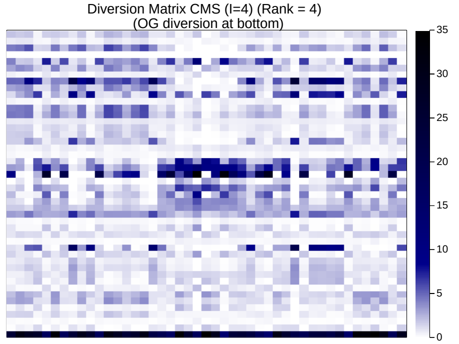
Image of Camille Jordan (1838-1922)



$$A \approx U_{266\times 25} \cdot \Sigma_{25\times 25} \cdot V_{25\times 266}$$

Completing the Matrix: $\mathcal{D}^T$





When might we want to do this?

- We have access to aggregate market shares and some (but not all) second-choice data (microBLP; Grieco, Murry, Yurukoglou (2022)).
- We are interested in estimating substitution patterns across all sets of products but have data on only a subset
    - shares of largest cellular phone providers, and number porting or switching data for merging parties only.
    - survey data on "If this Tesco were to close where would you shop" (as UK CMA asks).
    - win-loss data from merging parties only (Qiu, Sawada, Sheu (2022))
- In many cases we may not have sufficient variation in prices (or other covariates) to estimate a demand system.
- Product characteristics may not accurately capture substitution across products.

Setup and Model

Linear Algebra Notation

- Individual i's share for each choice given by $\mathbf{s_i} = [s_{i0}, s_{i1}, \ldots, s_{iJ}]$.
- Aggregate shares by $\sum_{i=1}^{I} \pi_i \cdot \mathbf{s_i} = \mathbf{s}$.
- The matrix of individual diversion ratios is given by $\mathbf{D}_i = \mathbf{s_i} \cdot \left[\frac{1}{(1-\mathbf{s_i})}\right]^T$.

We write the $(J+1) \times (J+1)$ matrix of second-choice diversion as:

$$
\begin{aligned}
D_{jk} &= \sum_{i=1}^{I} \pi_i \cdot \frac{s_{ik}}{1 - s_{ij}} \cdot \frac{s_{ij}}{s_j} \\
\mathbf{D} &= \left(\sum_{i=1}^{I} \pi_i \cdot \mathbf{s_i} \cdot \left[\frac{1}{(1-\mathbf{s_i})} \right]^T \cdot \operatorname{diag}(\mathbf{s_i}) \right) \cdot \operatorname{diag}(\mathbf{s})^{-1} \\
&= \left(\sum_{i=1}^{I} \pi_i \cdot \mathbf{s_i} \cdot \left[\frac{\mathbf{s_i}}{(1-\mathbf{s_i})} \right]^T \right) \cdot \operatorname{diag}(\mathbf{s})^{-1}
\end{aligned}
$$

Notation continued

Under relatively general conditions, second-choice diversion can be written as:

$$\text{diag}(\mathbf{s}) \cdot \mathbf{D} = \sum_{i=1}^{I} \pi_i \cdot \begin{bmatrix} | \\ \mathbf{s_i} \\ | \end{bmatrix} \cdot \begin{bmatrix} — & \frac{\mathbf{s_i}}{1-\mathbf{s_i}} & — \end{bmatrix}$$

- Each individual diversion ratio is of rank one since it is the outer product of $\mathbf{s_i}$ with itself (and some diagonal "weights").
- The (unrestricted) matrix of diversion ratios $\mathbf{D}$ is $(J+1) \times (J+1)$.
- Logit restricts $\mathbf{D}$ to be of rank one. Nested logit of rank $\leq G$ (the number of non-singleton nests). Mixed logit to $\text{rank}(\mathbf{D}) \leq I$ (but bound is likely uninformative).

Setting

- Assume that we observe aggregate market shares $\mathcal{S}_j$ and some subset of the diversion matrix $\mathcal{D}_{jk}$ for $(j,k) \in \text{OBS}$.
- Goal: Can we obtain an estimate for the remainder of the matrix $\mathcal{D}$?
    - Related to CS literature on matrix completion methods.
    - Useful tip from linear algebra: nuclear norm: $||A||_* = \sum_i \sigma_i(A)$ where $\sigma_i(A)$ are singular values. This works like a continuous approximation to rank.
    - We don't need to do nuclear norm penalization since discrete choice provides enough structure.
- Our low rank approximation is still consistent with utility maximization under discrete choice.
    - Theoretical interpretation as indirect utilities, not just mech. rank reduction (ie: PCA).

Our Semiparametric Problem

$$\min_{s_{ij},\pi_i,\overline{q}_0} \sum_{(k,j)\in \text{OBS}} \left(\mathcal{D}_{kj} - D_{kj}\right)^2 + \lambda_1 \cdot \sum_j \left(\mathcal{S}_j - \sum_i \pi_i \cdot s_{ij}\right)^2 + \lambda_2 \left\|\pi_i\right\|^2$$

$$\text{subject to} \quad D_{kj} = \sum_{i=1}^{I} \pi_i \cdot \frac{s_{ij}}{1-s_{ik}} \cdot \frac{s_{ik}}{s_k}$$

$$\mathcal{S}_j = \frac{\mathcal{Q}_j}{\overline{q}_0 + \sum_{k\in\mathcal{J}_t} \mathcal{Q}_k}$$

$$0 \le s_{ij}, \pi_i, s_j, D_{kj} \le 1, \quad \sum_{i=1}^{I} \pi_i = 1, \quad \sum_j s_{ij} = 1$$

- Use cross validation to select # of types I.
- With $\lambda_2 > 0$ we penalize *HHI* of w_i and becomes elastic net
- Can recover outside good share $\overline{q}_0$

Reformulated Problem

$$\min_{s_{ij},\pi_i} \sum_{(k,j)\in\text{OBS}} \frac{1}{\mathcal{S}_k^2} \cdot \left(\mathcal{S}_k \cdot \mathcal{D}_{kj} - \tilde{D}_{kj}\right)^2 + \lambda_1 \cdot \sum_j \left(\mathcal{S}_j - \sum_i \pi_i \cdot s_{ij}\right)^2 + \lambda_2 \left\|\pi_i\right\|^2$$

$$\text{subject to} \quad \tilde{D}_{kj} = \sum_{i=1}^{I} \pi_i \cdot \frac{s_{ij}}{1 - s_{ik}} \cdot \frac{s_{ik}}{1}$$

$$\mathcal{S}_j = \frac{q_j}{\widehat{q}_0 + \sum_{j\in\mathcal{J}_t} q_j}$$

$$0 \leq s_{ij}, \pi_i, s_j, D_{kj} \leq 1, \quad \sum_{i=1}^{I} \pi_i = 1, \quad \sum_j s_{ij} = 1$$

This reformulated problem isn't globally convex, but much easier to solve than previous problem. Requires that λ_1 large enough so that $\mathcal{S}_j \approx s_j$.

Discussion

- Goal: a good predictive model for unobserved elements of $\mathcal{D}$.
- We are worried about overfitting so we use cross validation (withholding columns of $\mathcal{D}$) to select number of types I.
  - Otherwise we would always prefer the more complicated model
  - Compare models based on out-of-sample fit (RMSE, MAD).
- Model may or may not be sparse $s_{ij} = 0$ for some (i, j)
  - Could be that consumer i doesn't consider j.
  - Or consequence that $s_{ij} \geq 0$ and $\sum_j s_{ij} = 1$ amounts to an L_1 penalty $\sum_j |s_{ij}| \leq 1$
- Model is a semiparametric logit for $V_{ij} \in \mathbb{R}$ (don't rule out $V_{ij} \to \pm\infty$):

$$u_{ij} = V_{ij} + \varepsilon_{ij}, \quad s_{ij} = \frac{e^{V_{ij}}}{1 + \sum_k e^{V_{ik}}}$$

Comparison: Fox, Kim, Ryan, Bajari (QE 2011)

$$\min_{\pi_i \geq 0} \sum_j \left(\mathcal{S}_j - \sum_i \pi_i \cdot \hat{s}_{ij}(\widehat{\beta}_i) \right)^2 \quad \text{subject to} \quad \sum_i \pi_i = 1$$

$$\hat{s}_{ij}(\widehat{\beta}_i) = \frac{e^{\widehat{\beta}_i x_j}}{1 + \sum_{j'} e^{\widehat{\beta}_i x_j'}}$$

- Draw $\beta_i \sim G(\beta_i)$ from a prior distribution.
- Solved in characteristic space with a semi-parametric form for $F(\beta_i)$.
- Often produces very sparse models $\pi_i = 0$ (for all but 50 of 1000 simulated consumers).
- Hard to incorporate fixed parameters (see Heiss, Hetzenecker, and Osterhaus (2021)).

Comparison: Raval et al. (2017, 2020)

- Cut data into bins (zip, income, age, gender)
- Observe shares (hospital demand) within each bin $s_{g(i),j}$
- A separate plain logit for each bin with only ξ_j as the common parameter.
- Use second choices from hospital closures (natural disasters) to compare models.

$$s_{g(i),j} = \frac{e^{\beta_g x_j + \xi_j}}{1 + \sum_{j'} e^{\beta_g x_{j'} + \xi_{j'}}}, \quad D_{kj,i} = \frac{s_{g(i),j}}{1 - s_{g(i),k}}$$

Comparison: Latent Class Logit (Greene and Hensher 2003)

Most similar to what we're doing here.

- Estimate separate β_i for each class.
- Estimate proportion of each class π_i.
- Estimating finite mixtures is tricky and usually requires EM.

$$s_k(\pi, \beta) = \sum_{i=1}^{I} \pi_i \cdot \left(\frac{e^{\beta_i x_{ij} + \xi_j}}{1 + \sum_k e^{\beta_i x_{ik} + \xi_k}} \right)$$

Data

Description of Vending Data

- Same data as Conlon and Mortimer (JPE, 2021).
- 66 Vending Machines in white-collar office buildings in downtown Chicago
- About 35-40 snack products in each
- 6 exogenous product removals (2.5-3.5 weeks long each)
  - Snickers, M&M Peanut, Doritos Nacho, Cheetos, Animal Crackers, Famous Amos
- Empirical challenge: Overall demand is pretty noisy over time.

Description of Cars Data

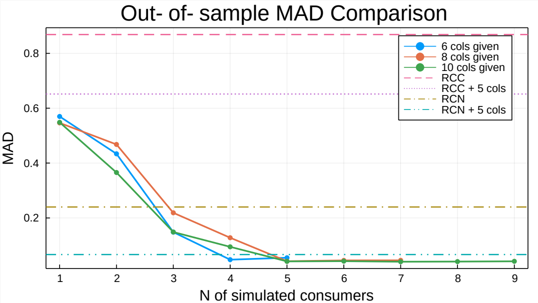
- Subset of data from Grieco, Murry and Yurukoglu (2022).
- Focus on one year of sales from 2015
  - Aggregate sales observed at the model-year level from Ward's Automotive.
  - Second choices from MaritzCX survey (53,328 purchases)
  - Construct $J = 181$ products by consolidating all models below 15,000 annual sales.
    - Consolidated products are: Car/Truck by Low/Mid/High prices (6 products)
- Same Goal: Predict unobserved second-choice data without characteristics.

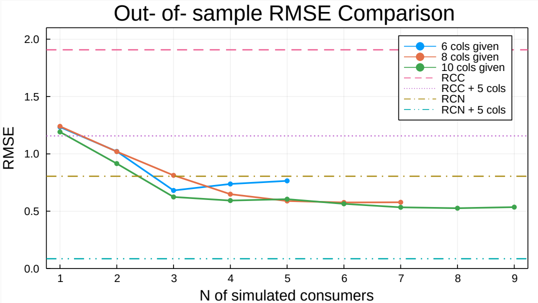
Monte Carlo

Generating Data

- Fit (i) nested logit, (ii) RC logit to the vending data from the JPE paper.
- Generate fake sales and diversion from those parameter estimates.
  - $J = 45$ products; $T = 250$ markets; with 30 randomly selected products in each. Market size $M = 1000$ per market. Nesting parameter is $\rho = 0.25$.
  - Categories: Salty Snacks, Chocolate, Non-Chocolate Candy, Cookies, Pastry, Other.
- Estimate a variety of misspecified parametric models: RC on nest dummies, RC on characteristics (Salt, Sugar, Nut Content), and our semiparametric estimator.
  - Include $m \ll J$ columns of $\mathcal{D}_{kj}$ as extra moments.
- Compare out-of-sample predicted Diversion Ratios.
  - MAD: Median $\left(|\mathcal{D}_{k,j} - \hat{D}_{k,j}|\right)$ for $(k,j) \in \{\text{Validation}\}$.
  - RMSE: $\sqrt{\frac{1}{n}\sum_{(k,j)\in\{\text{Validation}\}} |\mathcal{D}_{k,j} - \hat{D}_{k,j}|^2}$

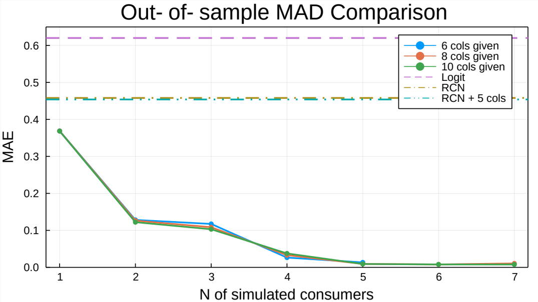
Monte Carlo: DGP is Nested Logit

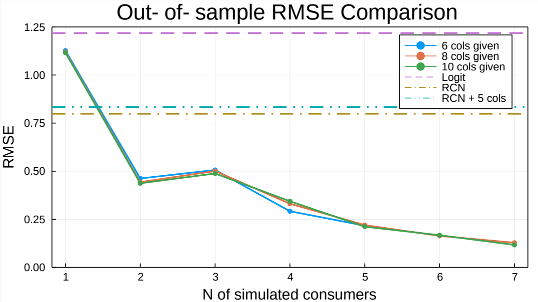




- RCC is mis-specified
- Diversion Moments improve efficiency of RCN
- $I \geq 4$ does a pretty good job.

Monte Carlo: DGP is RC on chars





- RCN is mis-specified
- $I \geq 2$ does a pretty good job.

Empirical Example

Measuring Observed Diversion

We need a measure of $\mathcal{D}_{jk} \equiv \frac{\partial q_k}{\partial p_j} / \left|\frac{\partial q_j}{\partial p_j}\right|$. Require:

1. A control week is valid IFF it is from the same machine.
2. A control week is valid IFF total sales do not increase, and do not decrease more than expected sales of removed product. $Q_s \in [Q_t, Q_t + E[q_j|Q_t]]$
3. $\mathcal{D}_{jk}$ is restricted to Unit Simplex. (Dirichlet Prior).
4. This cleans up noisy estimates from infrequently stocked substitutes.

Use `Turing.jl`

$$\left(\frac{\Delta q_1}{\Delta q_j}, \ldots, \frac{\Delta q_J}{\Delta q_j}, \frac{\Delta q_0}{\Delta q_j}\right) \sim \text{Multinomial}\left(\Delta q_j; \mathcal{D}_{j,1}, \ldots, \mathcal{D}_{j,K}, \mathcal{D}_{j,0}\right)$$

$$\left(\mathcal{D}_{j,1}, \ldots, \mathcal{D}_{j,K}, \mathcal{D}_{j,0}\right) \sim \text{Dirichlet}\left(\mathcal{S}_1, \ldots, \mathcal{S}_K, \mathcal{S}_0\right)$$

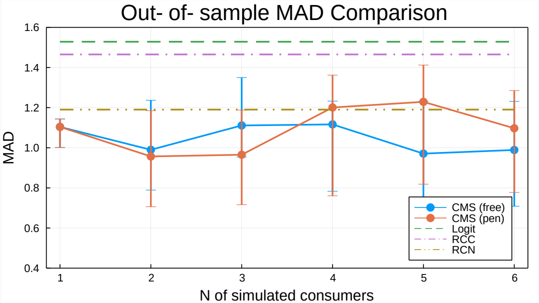
Experimental Diversion Estimates: M&M Peanuts Removal

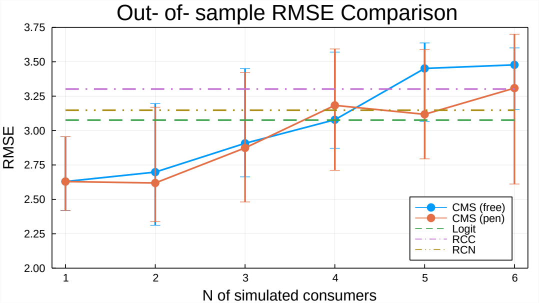
Product	$\mathcal{S}$	$\mathcal{D}_{jk}$
Outside Good	30.12	33.83
Snickers	3.96	17.58
M&M Milk Chocolate	1.16	10.69
Planters (Con)	1.92	8.99
Butterfinger	0.50	5.41
Raisinets	1.60	5.21
Twix Caramel	2.42	4.96
Sun Chip	2.12	1.98
Cliff (Con)	0.28	1.68
Nabisco (Con)	0.39	1.63
Choc Hershey (Con)	0.22	1.45

Use this as data $\mathcal{S}$ and $\mathcal{D}$ for our estimates. (CI: are fairly precise)

Results

Cross Validation: Model Selection



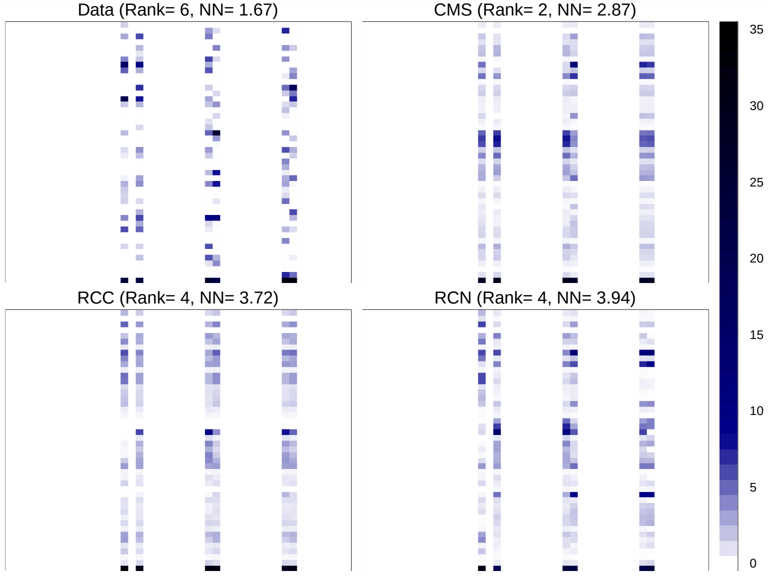


Out-of-sample fit (mostly) beats in-sample fit of parametric models.

Error bars are across all holdout experiments/ Dots are cross-validated means.

Seems to select $I = 2$ or $I = 3$ (bias-variance tradeoff).

Diversion Matrix: Fit Comparison



Top Substitutes: M&M Peanuts

Product	$\mathcal{S}$	$\mathcal{D}_{jk}$	Logit	RCC	RCN	CMS(I=2)	CMS(I=3)
Outside Good	30.12	33.21	23.25	24.74	26.30	31.81	30.87
Snickers	3.96	17.74	3.24	4.69	6.29	8.55	8.47
Twix Caramel	2.42	7.23	2.48	2.85	4.83	7.82	9.58
KarNuts (Con)	1.70	5.91	1.74	3.13	1.51	1.80	1.71
Baked (Con)	2.39	5.04	2.12	1.30	1.37	2.66	2.02
Nature Valley (Con)	2.00	4.78	2.30	2.55	2.01	1.57	1.12
Twizzlers	0.33	4.37	1.86	2.07	1.74	1.28	1.15
Reeses Peanut Butter Cups	0.59	4.03	1.78	2.78	3.48	3.19	4.14
M&M Milk Chocolate	1.16	3.05	2.02	2.64	4.03	5.24	5.77
Rice Krispies Treats	0.27	2.62	1.01	0.90	0.94	2.82	2.47
Planters (Con)	1.92	2.36	2.22	4.33	1.45	3.75	4.40

Top Substitutes: Doritos and Cheetos

Product	$\mathcal{S}$	$\mathcal{D}_{jk}$	Logit	RCC	RCN	CMS(I=2)	CMS(I=3)
Outside Good	30.12	35.57	23.59	14.99	18.03	28.39	32.32
Baked (Con)	2.39	8.85	2.15	3.34	3.54	2.01	3.54
FritoLay (Con)	1.48	7.74	1.61	2.47	2.64	1.68	3.86
Snickers	3.96	7.52	3.28	2.72	2.08	6.15	1.29
Ruffles (Con)	2.80	6.77	2.62	3.96	4.31	1.92	4.19
Frito	1.86	6.74	2.23	2.75	3.66	1.38	3.88
Dorito Blazin Buffalo Ranch	0.61	3.07	1.56	2.40	2.59	1.57	1.69
Cheez-It Original	2.00	2.70	1.94	3.09	3.19	1.46	3.35
Rold Gold (Con)	2.56	2.66	3.09	6.82	5.09	1.81	1.52
Rice Krispies Treats	0.27	2.59	1.02	1.16	0.72	2.22	3.02
KarNuts (Con)	1.70	2.49	1.76	1.34	1.10	1.29	1.12

Underlying Diversion: Individual shares: $\lambda_2 = 1$

Category	Model/Rank:		I = 1	I = 2		I = 3			I = 4			
	Weight on individual:		100.0%	50.8%	49.2%	33.6%	34.1%	32.3%	25.8%	25.1%	24.7%	24.4%
	Product	Logit Sj	i = 1	i = 1	i = 2	i = 1	i = 2	i = 3	i = 1	i = 2	i = 3	i = 4
SALTY SNACKS	Snyders (Con)	2.21	0.70	0.34	1.50	0.00	2.61	0.33	0.00	2.44	0.26	1.24
	Cheetos	2.52	0.50	0.00	0.00	0.19	0.76	0.29	0.39	1.03	0.07	0.01
	Ruffles (Con)	0.98	1.88	0.61	4.91	0.00	5.62	3.06	0.00	6.41	2.97	0.00
	Dorito Nacho	2.05	0.98	0.93	0.51	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Rold Gold (Con)	0.94	1.86	2.35	0.41	2.48	1.37	0.67	0.00	2.38	0.25	8.44
	Baked (Con)	1.99	2.08	2.39	1.08	1.28	5.64	0.00	2.41	4.62	0.00	0.00
	Salty Other (Con)	2.78	0.22	0.17	0.25	0.09	0.42	0.25	0.00	0.38	0.16	0.93
	Sun Chip	1.81	4.76	0.00	18.60	0.00	0.00	21.83	0.00	5.81	20.63	0.00
	Cheez-It	1.77	1.47	0.76	3.09	0.00	4.70	1.47	0.49	4.65	1.35	0.00
	Jays (Con)	1.48	0.17	0.16	0.13	0.08	0.33	0.17	0.00	0.30	0.10	0.73
	Frito	2.03	1.41	0.98	2.33	0.00	6.36	0.00	0.00	6.26	0.00	0.62
	FritoLay (Con)	1.49	1.71	1.35	2.41	0.03	6.25	0.48	0.11	6.16	0.41	1.59
	Smartfood	1.51	0.56	0.53	0.49	0.34	0.87	0.61	0.36	0.85	0.49	0.58
	Lays	1.45	0.54	0.40	0.77	0.00	2.10	0.05	0.00	1.98	0.01	0.80
	Cheetos Flamin	0.96	0.55	0.35	0.91	0.00	2.43	0.06	0.00	2.44	0.06	0.11
	Dorito Blazin	1.45	1.47	0.00	5.49	0.00	0.00	6.80	0.00	1.82	6.37	0.00
	Popcorn (Con)	2.06	0.51	0.58	0.22	0.63	0.00	0.69	0.41	0.12	0.58	0.95
	Ritz Bits	0.51	0.16	0.21	0.00	0.19	0.09	0.11	0.00	0.17	0.04	0.63
CHOCOLATE CANDY	M&M Peanut	3.21	4.78	7.08	0.00	8.74	1.98	0.02	10.19	0.00	0.00	2.22
	Snickers	3.53	6.08	8.63	0.00	9.94	0.00	0.53	9.52	0.68	0.46	9.10
	Twix Caramel	2.29	5.19	8.10	0.00	10.88	0.00	0.00	0.00	0.00	0.00	32.60
	Raisinets	1.47	1.47	2.19	0.00	2.74	0.00	0.00	3.44	0.00	0.00	0.00
	M&M Milk Choc	1.80	3.63	5.32	0.00	6.58	0.41	0.02	6.91	0.00	0.00	3.42
	Choc Mars (Con)	2.13	1.03	1.57	0.00	2.04	0.00	0.00	0.00	0.35	0.00	6.49
	Reeses PB Cups	1.68	1.62	3.50	0.00	4.73	0.00	0.00	3.58	0.00	0.00	3.96
	Butterfinger	1.10	2.72	3.33	1.01	3.73	1.18	1.80	3.62	0.99	1.50	3.14
	Choc Herhsey (Con)	1.22	2.87	1.46	6.28	1.61	1.34	7.09	1.26	3.34	6.48	2.22

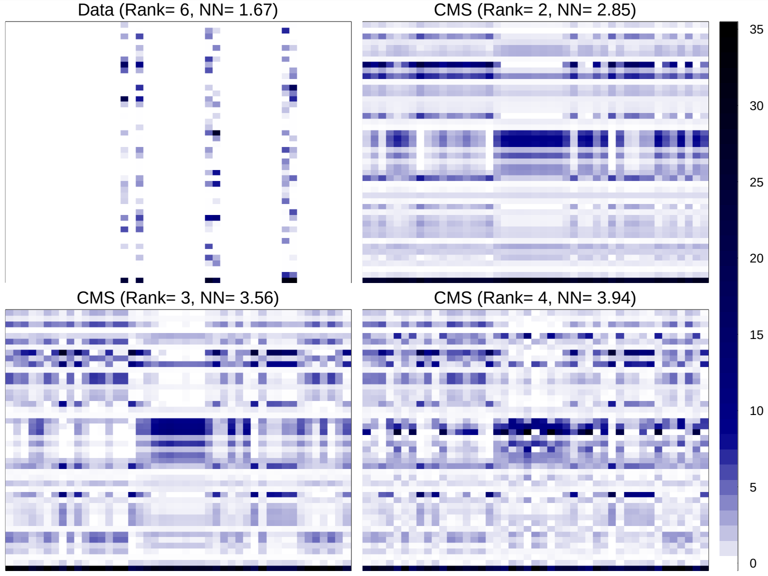
Category	Model/Rank:		I = 1	I = 2		I = 3			I = 4			
	Weight on individual:		100.0%	50.8%	49.2%	33.6%	34.1%	32.3%	25.8%	25.1%	24.7%	24.4%
	Product	Logit Sj	i = 1	i = 1	i = 2	i = 1	i = 2	i = 3	i = 1	i = 2	i = 3	i = 4
NONCHOC. CANDY	Skittles Original	1.03	0.12	0.14	0.01	0.07	0.23	0.11	0.00	0.22	0.04	0.51
	Nonchoc Other (Con)	1.06	0.41	0.60	0.00	0.69	0.17	0.03	0.00	0.27	0.00	2.31
	Twizzlers	1.66	1.16	1.14	1.08	0.91	2.10	0.61	1.75	1.27	0.67	0.00
COOKIES	ZAnimal Cracker	1.90	0.29	0.37	0.05	0.62	0.02	0.13	0.53	0.00	0.17	0.25
	CC Fam Amos	1.58	1.57	0.00	3.35	0.01	0.00	9.61	0.05	0.00	16.91	0.00
	Ruger Wafer (Con)	1.60	0.54	0.62	0.25	0.47	1.05	0.08	0.00	1.18	0.00	2.20
	Grandmas CC	1.15	0.84	0.39	1.90	0.36	0.18	2.58	0.55	0.67	2.43	0.00
	Rasbry Knotts	0.68	1.10	0.40	2.75	0.44	0.34	3.30	0.63	1.15	3.08	0.00
	Choc Fam Amos	0.91	1.35	1.10	1.86	1.37	0.00	2.95	1.39	0.00	2.94	1.00
	Nabisco (Con)	1.23	1.44	1.14	2.08	1.29	0.00	3.14	1.97	0.05	2.86	0.00
PASTRY	Pop-Tarts (Con)	2.42	0.27	0.37	0.00	0.38	0.21	0.09	0.00	0.24	0.02	1.48
	Rice K Treats	0.85	2.25	2.61	1.24	2.06	4.30	0.33	3.44	3.19	0.25	0.00
OTHER	Nature Valley (Con)	2.13	1.42	1.29	1.58	0.38	5.00	0.00	1.40	3.94	0.00	0.00
	Planters (Con)	1.63	4.81	3.54	7.77	4.37	0.00	10.57	6.08	0.85	9.94	0.00
	KarNuts (Con)	1.65	1.25	1.77	0.00	1.73	1.36	0.00	2.56	0.41	0.00	0.00
	Farleys Fruit Snax	0.99	0.58	0.45	0.80	0.03	2.27	0.00	0.27	2.02	0.00	0.18
	Cherry Fruit Snax	0.52	0.09	0.12	0.00	0.07	0.15	0.05	0.00	0.16	0.00	0.34
	Cliff (Con)	3.91	1.03	1.22	0.41	1.35	0.00	1.26	1.84	0.00	1.04	0.00
	Outside Good	25.34	28.58	29.44	24.49	27.09	38.12	18.85	34.84	31.17	17.44	11.96

Underlying Diversion: Individual shares: $\lambda_2 = 0$

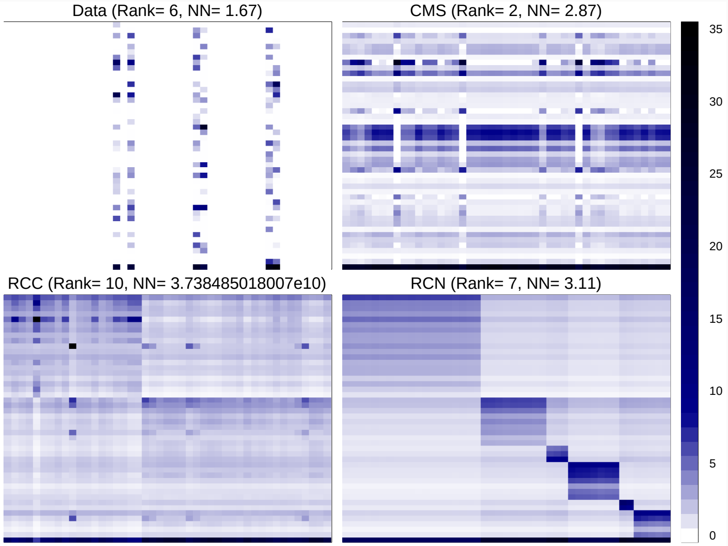
	Model/Rank:		I = 1	I = 2		I = 3			I = 4			
	Weight on individual:		100.0%	81.2%	18.8%	62.8%	31.4%	5.8%	73.5%	24.1%	2.4%	0.02%
	Product	Logit Sj	i = 1	i = 1	i = 2	i = 1	i = 2	i = 3	i = 1	i = 2	i = 3	i = 4
SALTY SNACKS	Snyders (Con)	2.21	0.70	0.64	0.70	0.00	0.00	2.17	0.00	2.38	0.00	0.23
	Cheetos	2.52	0.50	0.00	0.00	0.24	0.25	0.80	0.31	0.00	0.00	0.00
	Ruffles (Con)	0.98	1.88	0.98	4.82	0.00	2.53	4.84	0.03	5.05	2.54	0.00
	Dorito Nacho	2.05	0.98	0.90	1.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Rold Gold (Con)	0.94	1.86	2.35	0.00	2.50	0.13	1.40	0.65	1.95	0.14	4.41
	Baked (Con)	1.99	2.08	2.49	0.40	1.36	0.00	3.94	2.35	3.76	0.00	0.43
	Salty Other (Con)	2.78	0.22	0.29	0.00	0.05	0.00	0.59	0.00	0.71	0.00	0.30
	Sun Chip	1.81	4.76	0.00	22.96	0.00	19.29	5.19	0.09	5.71	18.70	0.00
	Cheez-It	1.77	1.47	1.06	2.67	0.00	0.91	3.90	0.00	4.00	0.93	0.43
	Jays (Con)	1.48	0.17	0.23	0.00	0.04	0.00	0.47	0.00	0.57	0.00	0.24
	Frito	2.03	1.41	1.28	1.65	0.00	0.00	4.55	0.00	4.70	0.00	0.04
	FritoLay (Con)	1.49	1.71	1.57	1.94	0.15	0.36	4.49	0.53	4.50	0.45	0.34
	Smartfood	1.51	0.56	0.67	0.00	0.31	0.16	1.02	0.00	1.17	0.18	0.69
	Lays	1.45	0.54	0.56	0.31	0.00	0.00	1.62	0.00	1.68	0.00	0.30
	Cheetos Flamin	0.96	0.55	0.50	0.55	0.00	0.00	1.83	0.00	1.92	0.03	0.00
	Dorito Blazin	1.45	1.47	0.01	6.47	0.00	5.77	1.82	0.00	2.06	5.62	0.00
	Popcorn (Con)	2.06	0.51	0.63	0.00	0.61	0.33	0.34	0.22	0.54	0.36	0.80
	Ritz Bits	0.51	0.16	0.22	0.00	0.17	0.00	0.23	0.00	0.32	0.03	0.26
CHOCOLATE CANDY	M&M Peanut	3.21	4.78	6.46	0.00	8.95	0.00	1.18	16.65	0.00	0.00	0.00
	Snickers	3.53	6.08	8.00	0.00	9.75	0.70	0.00	14.91	0.00	0.11	0.13
	Twix Caramel	2.29	5.19	7.32	0.00	11.03	0.00	0.00	4.04	0.00	0.00	18.60
	Raisinets	1.47	1.47	2.03	0.00	2.67	0.00	0.21	2.74	0.03	0.05	2.06
	M&M Milk Choc	1.80	3.63	4.90	0.00	6.47	0.00	0.67	5.31	0.36	0.04	7.24
	Choc Mars (Con)	2.13	1.03	1.46	0.00	2.03	0.00	0.11	0.00	0.31	0.00	4.69
	Reeses PB Cups	1.68	1.62	2.98	0.00	4.68	0.00	0.36	2.95	0.25	0.00	7.38
	Butterfinger	1.10	2.72	3.21	0.80	3.69	1.13	1.67	2.06	1.73	1.16	5.62
	Choc Herhsey (Con)	1.22	2.87	1.58	7.20	1.64	5.92	2.92	0.91	3.23	5.77	2.53

	Model/Rank:		I = 1	I = 2		I = 3			I = 4			
	Weight on individual:		100.0%	81.2%	18.8%	62.8%	31.4%	5.8%	73.5%	24.1%	2.4%	0.02%
	Product	Logit Sj	i = 1	i = 1	i = 2	i = 1	i = 2	i = 3	i = 1	i = 2	i = 3	i = 4
NONCHOC. CANDY	Skittles Original	1.03	0.12	0.18	0.00	0.03	0.00	0.36	0.00	0.43	0.00	0.17
	Nonchoc Other (Con)	1.06	0.41	0.59	0.00	0.65	0.00	0.32	0.00	0.41	0.00	1.54
	Twizzlers	1.66	1.16	1.20	0.86	0.90	0.59	1.71	1.96	1.41	0.66	0.00
COOKIES	ZAnimal Cracker	1.90	0.29	0.35	0.14	0.33	0.39	0.00	0.16	0.00	0.00	0.00
	CC Fam Amos	1.58	1.57	0.00	3.66	0.04	26.42	0.00	0.23	0.00	28.73	0.00
	Ruger Wafer (Con)	1.60	0.54	0.68	0.00	0.46	0.00	0.94	0.00	1.07	0.00	1.31
	Grandmas CC	1.15	0.84	0.46	2.07	0.34	2.21	0.89	0.47	0.92	2.25	0.05
	Rasbry Knotts	0.68	1.10	0.47	3.18	0.45	2.89	1.12	0.76	1.19	2.87	0.00
	Choc Fam Amos	0.91	1.35	1.09	2.12	1.47	2.65	0.38	1.19	0.52	2.63	1.35
	Nabisco (Con)	1.23	1.44	1.22	2.04	1.24	2.28	1.11	0.99	1.20	2.21	1.44
PASTRY	Pop-Tarts (Con)	2.42	0.27	0.39	0.00	0.34	0.00	0.36	0.00	0.46	0.00	0.87
	Rice K Treats	0.85	2.25	2.64	0.80	2.06	0.16	3.22	2.78	3.03	0.24	1.59
OTHER	Nature Valley (Con)	2.13	1.42	1.47	1.02	0.42	0.00	3.54	1.63	3.22	0.00	0.00
	Planters (Con)	1.63	4.81	3.51	9.13	4.39	8.99	2.79	4.91	2.79	8.74	3.10
	KarNuts (Con)	1.65	1.25	1.68	0.00	1.71	0.00	1.05	2.51	0.87	0.01	0.36
	Farleys Fruit Snax	0.99	0.58	0.57	0.45	0.04	0.00	1.69	0.16	1.69	0.00	0.08
	Cherry Fruit Snax	0.52	0.09	0.14	0.00	0.05	0.00	0.25	0.00	0.30	0.00	0.12
	Cliff (Con)	3.91	1.03	1.29	0.00	1.30	0.51	0.67	0.62	0.82	0.48	2.03
	Outside Good	25.34	28.58	29.75	22.86	27.43	15.42	33.28	27.87	32.77	15.06	29.27

Diversion Matrix: Estimates Comparison $\lambda_2 = 1$

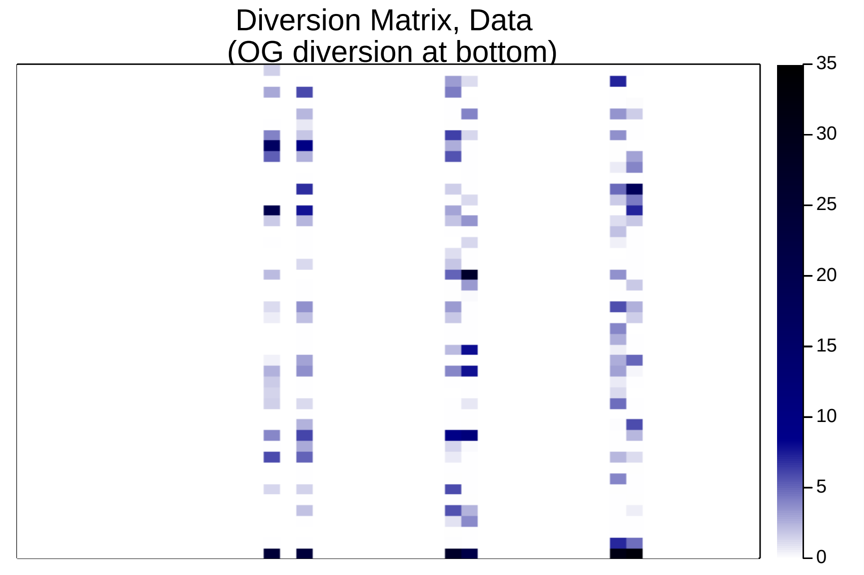


Diversion Matrix: Estimates Comparison $\lambda_2 = 0$



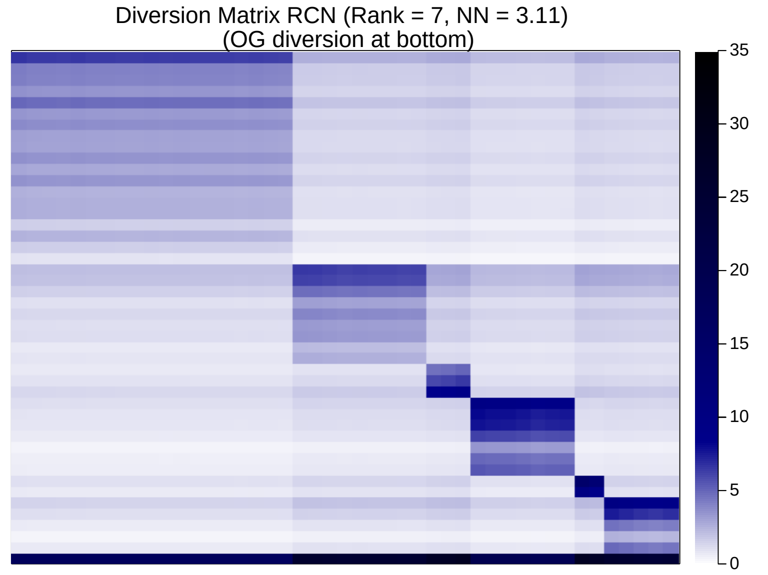
▸ Close look: RCC ▸ Close look: RCN ▸ Close look: Data only ▸ Close look: CMS2 ▸ Close look: CMS4

Diversion Matrix: Experiment columns only



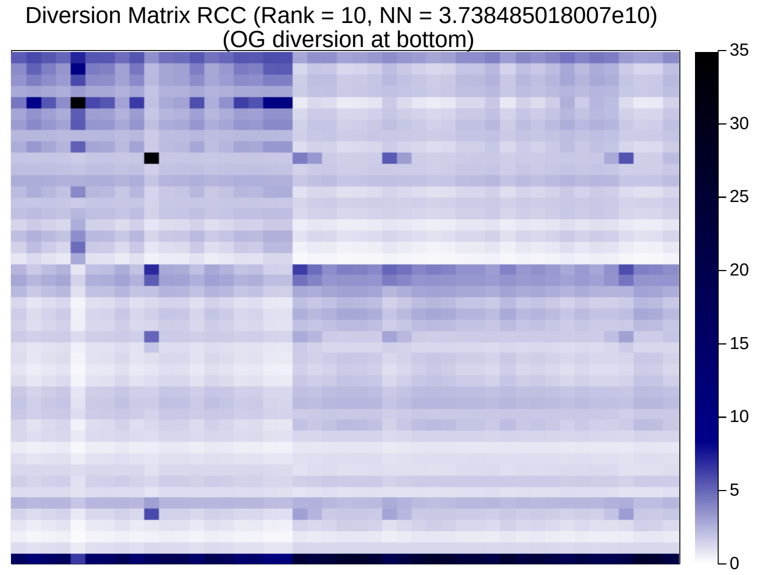
Back

Diversion Matrix: Rand. Coeff. on Nests (RCN)



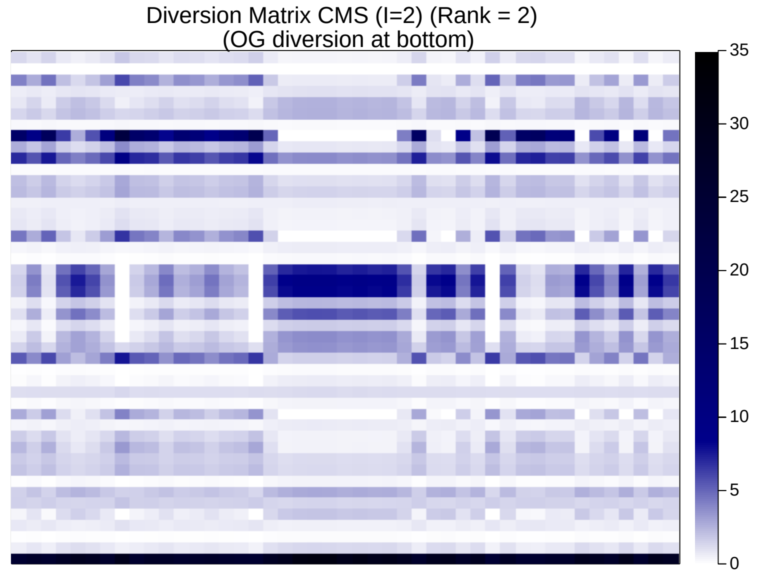
Back

Diversion Matrix: Random Coeff. on Chars. (RCC)



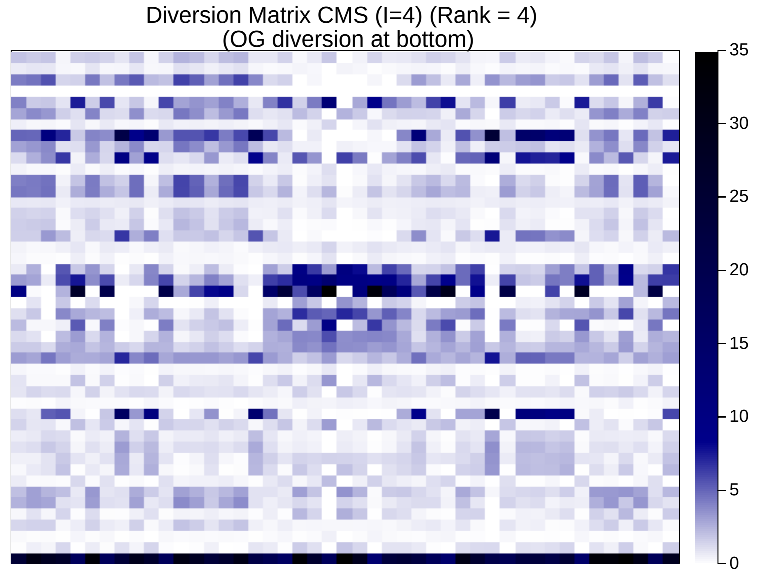
Back

Diversion Matrix: Semiparametric (CMS)



Back

Diversion Matrix: Semiparametric (CMS)



Back

Results on Cars Data

Diversion Matrix: Comparison

Top Substitutes: Ford F Series

Model	Raw	CMS (I=19)	CMS (I=21)	GMY (Parametric)
Ram Pickup	24.6	18.2	14.9	24.3
Gmc Sierra	20.3	10.8	8.7	-
Chevrolet Silverado	15.6	15.4	12.9	33.6
Toyota Tundra	13.0	5.9	4.5	-
Toyota Tacoma	6.3	1.1	0.9	-
Chevrolet Colorado	4.6	0.6	0.4	-
Gmc Canyon	2.3	0.2	0.1	-
Nissan Frontier	1.6	0.4	0.2	-
Jeep Wrangler	1.6	1.4	1.4	-
Truck Mid (Con)	0.7	2.3	2.4	-

Top Substitutes: Ford Mustang

Model	Raw	CMS (I=19)	CMS (I=21)	GMY (Parametric)
Dodge Challenger	19.9	5.5	5.2	7.4
Chevrolet Camaro	12.9	5.8	5.4	8.9
Car Hi (Con)	12.6	10.3	10.8	-
Car Mid (Con)	7.0	2.7	2.5	-
Hyundai Genesis	3.7	0.3	0.3	-
Subaru Wrx	3.3	0.4	0.4	-
Dodge Charger	3.0	4.8	4.8	-
Honda Accord	2.3	1.4	1.5	-
Ford Fiesta	1.9	0.4	0.3	-
Audi A3	1.9	0.2	0.1	-

Extensions

- What about (exogenous) price or quality changes?
  Expression for D_{jk} changes slightly.
- Want to add covariates or endogenous prices?
  Straightforward to run an IV regression:

  $$\log s_{ij} - \log s_{i0} = x_j \beta_i + \xi_j$$

  Test how much we lose using only a basis in $f(x_1, x_2)$.
- Optimal Experimentation: Which product is most informative about $\mathcal{D}$?
  - $\mathcal{D}$ looks like a transition matrix with a network structure
  - Relates to measures of centrality / eigenvalues.
  - Cross elasticities are not a well-behaved network.

Conclusion

- Allowing for flexible unobserved types can give more accurate substitution patterns
    - Particularly true in capturing closeness of best substitutes not captured by product characteristics (e.g. Snickers and Peanut M&M's vs Snickers and Milky Way)
- Using observable substitution patterns (experiments or surveys) and "completing" the $(J+1) \times (J+1)$ matrix with a low-rank approximation looks promising.
- How much information on second choices is "enough"?
- Other cases where we have incomplete substitution patterns?