How Much Should We Trust Regional-Exposure Designs?

Jeremy Majerovitz (St. Louis Fed) and Karthik Sastry (Princeton)

July 1, 2023 (SED Meeting, Cartagena)

Regional Exposure Designs Are Common

Common research design: Combine aggregate shock time series with heterogeneous exposure across regions (or other units)

- to learn open-economy fiscal multiplier exploiting heterogeneous geographic incidence of military spending (Nakamura and Steinsson, 2014)
- to learn effect of import competition exploiting different manufacturing product composition of US regions (Autor, Dorn, and Hanson, 2013)
- to learn inverse labor supply elasticity exploiting heterogeneous exposure to industry labor demand shocks (Bartik, 1991)

Our questions: When will our estimator converge to the true causal effect? How do we construct valid confidence intervals?

We Clarify Inference for Regional Exposure Designs

We use the idea of a residual with an approximate factor structure (factor + idiosyncratic) to clarify identification and inference

Identification can come from *shares* or *shocks*

- Shocks more plausible, and is argument most papers (implicitly) make

Under identification from shocks, clustering by region is not valid

- Can two-way cluster (or two-way HAC), or do randomization inference

Application to regional fiscal multipliers (Nakamura and Steinsson 2014)

- Clustering by state rejects 27% of the time in placebo
- Clustering by state 95% CI: $[0.56, 4.39]$
- Randomization Inference 95% CI: $[0.08, 5.34]$

Can improve efficiency with optimal instrument: doubles power in simulation

Related Literature

Econometrics:

- **Non-standard inference in shift-share settings.** Goldsmith-Pinkham, Sorkin, and Swift (2020); Borusyak, Hull, and Jaravel (2022); Borusyak and Hull (2021); Adão, Kolesar, and Morales (2019); Arkhangelsky and Korovkin (2019)
- **Panel standard errors beyond multiple clustering.** Conley (1994), Driscoll and Kraay (1998), Thompson (2011)

Macroeconomics:

- **Correlated fluctuations across domestic/international regions.** Acemoglu, Akcigit, and Kerr (2015); Caliendo, Parro, Rossi-Hansberg, and Sarte (2018)
- **Regional fiscal multipliers:** Nakamura and Steinsson (2014); Chodorow-Reich (2019)
- **Correlated fluctuations among US industries.** Long and Plosser (1987); Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012); Foerster, Sarte, and Watson (2011); Carvalho and Gabaix (2013)

Model

The Regional Exposure Design and Approximate Factor Structure

Consider the following model:

$$Y_{it} = \alpha_t + \gamma_i + \beta \cdot X_{it} + u_{it}$$
$$X_{it} = \omega_t + \zeta_i + \pi \cdot Z_{it} + e_{it}$$

In each period there is a vector-valued aggregate shock $S_t \in \mathbb{R}^K$, for $K \geq 1$. Each region has an exposure $\eta_i \in \mathbb{R}^K$ to each dimension of the shock. We construct the regional-exposure instrument:

$$Z_{it} = \eta_i' S_t$$

Moreover, assume residual has approximate factor structure

$$u_{it} = \lambda_i' F_t + \varepsilon_{it}$$

with factor loadings $\lambda_i \in \mathbb{R}^J$ and factor shocks $F_t \in \mathbb{R}^J$, for $J \geq 1$.

Example: Nakamura and Steinsson (2014)

Nakamura and Steinsson (2014) estimate the following equation:

$$\text{Output Growth}_{it} = \alpha_t + \gamma_i + \beta \cdot \text{Military Procurement Growth}_{it} + u_{it}$$

Panel of fifty states, plus the District of Columbia, with annual data from 1968-2006.

They want to identify β, the regional fiscal multiplier. Two IV strategies:

$$Z_{it} = \eta_i' S_t$$

S_t is national military procurement spending growth

η_i is either:

- A state fixed effect, or
- Military procurement spending as a share of state GDP at start of sample.

We Will Maintain Some Assumptions Throughout

We make some baseline assumptions to make life easier:

- u_{it} and e_{it} have zero mean in each time period and in each region.
- X_{it}, Y_{it}, and Z_{it} have zero mean across regions and time-periods and the econometrician observes a balanced panel of these variables.
- All moments of the form $\mathbb{E}[X_{it}^{a_X} Y_{js}^{a_Y} Z_{kr}^{a_Z}]$, for $(a_X, a_Y, a_Z) \geq 0$ and indices $(i, j, k) \in \{1, \ldots N\}$ and $(t, s, r) \in \{1, \ldots, T\}$, exist and are finite.
- ε_{it} is independent from λ_i and F_t, and has zero mean in each time period and in each region.
- Define λ_i and F_t to each have mean zero.

Why Use a Factor Structure?

We assume that residual has an approximate factor structure:

$$u_{it} = \lambda_i' F_t + \varepsilon_{it}$$

Parsimoniously captures idea that regions comove in response to aggregate shocks.

Can also stand in for spatial/sectoral linkages and/or for "granular" shocks that affect large/influential regions and spill over.

Mirrors structure of the instrument: one paper's instrument is another's residual!

Note: The solutions we propose will be robust to more general correlation structures.

Econometric Issues

Unpacking the Orthogonality Condition

We make two assumptions to let us unpack the orthogonality condition

Assumption

(η_i, λ_i) *is independent from* (S_t, F_t)

Assumption

$\mathbb{E}\left[Z_{it}\varepsilon_{it}\right] = 0$

We then have:

$$\begin{aligned}
\mathbb{E}\left[Z_{it}u_{it}\right] &= \mathbb{E}\left[\eta_i' S_t \cdot \lambda_i' F_t\right] + \mathbb{E}\left[\eta_i' S_t \cdot \varepsilon_{it}\right] \\
&= \mathbb{E}\left[S_t'(\eta_i \lambda_i') F_t\right] \\
&= \mathbf{tr}\left(\mathbb{E}\left[S_t'(\eta_i \lambda_i') F_t\right]\right) \\
&= \mathbf{tr}\left(\mathbb{E}\left[(\eta_i \lambda_i')(F_t S_t')\right]\right) = \mathbf{tr}\left(\mathbb{E}\left[\eta_i \lambda_i'\right]\mathbb{E}\left[F_t S_t'\right]\right)
\end{aligned}$$

Yields two paths to identification: $\mathbb{E}\left[\eta_i \lambda_i'\right] = 0$ or $\mathbb{E}\left[F_t S_t'\right] = 0$

Identification Can Come from Shares or Shocks

There are two primary *sufficient conditions* for identification, $\mathbb{E}\left[Z_{it} u_{it}\right] = 0$

Condition (Identification from Shares)

The regional exposures are uncorrelated with the factor loadings, or $\mathbb{E}[\eta_i \lambda_i'] = 0$.

Condition (Identification from Shocks)

The aggregate shocks are uncorrelated with the factor shocks, or $\mathbb{E}[F_t S_t'] = 0$.

Under identification from shares (and auxiliary assumptions), can show $\hat{\beta} \to \beta$ as $N \to \infty$. Under identification from shocks, need to use $T \to \infty$. Proposition

Connection to the shift-share literature: GPSS (2020) use identification from shares; BHJ (2022) and AKM (2019) use identification from shocks + many-industries asymptotics.

Which Identifying Assumption Is More Plausible?

We argue that identification from shares is usually implausible:

- Places that make fighter jets are different; exposed to different shocks

Literature tends to rely on identification from shocks. Nakamura and Steinsson write:

> *Our identifying assumption is that the United States does not embark on military buildups—such as those associated with the Vietnam War and the Soviet invasion of Afghanistan—because states that receive a disproportionate amount of military spending are doing poorly relative to other states.*

But, we shall see that if identification comes from shocks, inference must change.

Unpacking the Standard Errors

Since we have two-way fixed effects, we use $\tilde{W}$ to denote $W - \bar{W}_i - \bar{W}_t + \bar{W}$. We then have

$$\text{AVAR}\left(\sqrt{N} \cdot \hat{\beta}\right) = \mathbb{E}\left[\tilde{Z}_{it} \tilde{X}_{it}'\right]^{-1} \text{AVAR}\left(\frac{1}{\sqrt{N}} \cdot \frac{1}{T} \sum_{i,t} \tilde{Z}_{it} \tilde{u}_{it}\right) \mathbb{E}\left[\tilde{X}_{it} \tilde{Z}_{it}'\right]^{-1}$$

Define the "meat" of the sandwich

$$\Omega := \text{AVAR}\left(\frac{1}{\sqrt{N}} \cdot \frac{1}{T} \sum_{i,t} \tilde{Z}_{it} \tilde{u}_{it}\right)$$

We can separate the factor and idiosyncratic components with a new assumption:

Assumption

For all i, $(\varepsilon_{it})_{t=1}^{T}$ is independent from $\left((\eta_j)_{j=1}^{N}, (\lambda_j)_{j=1}^{N}, (S_t)_{t=1}^{T}, (F_t)_{t=1}^{T}\right)$.

Unpacking the Standard Errors

Can now decompose the meat:

$$\Omega = \underbrace{\text{AVAR}\left(\frac{1}{\sqrt{N}} \cdot \frac{1}{T} \sum_{i,t} \tilde{Z}_{it} \tilde{\lambda}_i' \tilde{F}_t\right)}_{\text{Factor component}} + \underbrace{\text{AVAR}\left(\frac{1}{\sqrt{N}} \cdot \frac{1}{T} \sum_{i,t} \tilde{Z}_{it} \tilde{\varepsilon}_{it}\right)}_{\text{Idiosyncratic component}}$$

Key question is the factor component: which covariances are non-zero?

Under identification from shares, $\tilde{Z}_{it} \tilde{\lambda}_i' \tilde{F}_t$ will be uncorrelated across regions.

Under identification from shocks, $\tilde{Z}_{it} \tilde{\lambda}_i' \tilde{F}_t$ will be uncorrelated across time. Lemmas

This determines how we can cluster.

Clustering by Region Is Valid Under Identification from Shares

Proposition (Clustering by Region Is Valid Under Identification from Shares)

Assume identification from shares, and that $\left(\eta_i, \lambda_i, (\varepsilon_{it})_{t=1}^{T}, (e_{it})_{t=1}^{T}\right)$ *is independently and identically drawn across regions. Then clustering by region consistently estimates* $AVAR\left(\sqrt{N} \cdot \hat{\beta}\right)$ *as* $N \to \infty$.

So, if identification comes from shares, then current standard practice is fine. But most papers implicitly argue for identification from shocks...

Clustering by Region Is Biased Under Identification from Shocks

Proposition (Clustering by Region Is Biased Under Identification from Shocks)

Assume identification from shocks, that $\left(\eta_i, \lambda_i, (\varepsilon_{it})_{t=1}^{T}\right)$ *is independently and identically drawn across regions, that* ε *is independent of* Z*, and that* (S_t, F_t) *is independently and identically drawn across time. Define*

$$\Omega^{CR} := \mathbb{E}\left[\left(\frac{1}{T}\sum_t \tilde{Z}_{it}\tilde{u}_{it}\right)\left(\frac{1}{T}\sum_t \tilde{Z}_{it}\tilde{u}_{it}\right)'\right]$$

and assume this expectation exists and is finite. Then, as $N \to \infty$*, the asymptotic bias of the clustered estimate of* Ω *is given by:*

$$\frac{1}{N}(\Omega^{CR} - \Omega) \to -\frac{1}{T}\mathbb{E}\left[(\tilde{S}_t'\mathbb{E}\left[\tilde{\eta}_i\tilde{\lambda}_i'\right]\tilde{F}_t)^2\right] - O\left(\frac{1}{T^2}\right)$$

Region-clustered standard errors too small (ignoring $O\left(1/T^2\right)$). Bias scales with N/T.

Proposed Econometric Solutions

Two-Way Clustering Is a Good Idea

The meat term is $\Omega = \lim \frac{1}{NT^2} \sum_{i,t} \sum_{j,s} \mathbb{E}\left[\tilde{u}_{it} \tilde{u}_{js} \tilde{Z}_{it} \tilde{Z}_{js}\right]$

Most estimators will take the form: $\hat{\Omega} = \frac{1}{NT^2} \sum_{i,t} \sum_{j,s} f(i,t,j,s) \mathbb{E}\left[\hat{u}_{it} \hat{u}_{js} \tilde{Z}_{it} \tilde{Z}_{js}\right]$

- No clustering: $f(i,t,j,s) = \mathbf{1}(i = j \text{ AND } t = s)$
- Clustering by region: $f(i,t,j,s) = \mathbf{1}(i = j)$
- Two-way clustering: $f(i,t,j,s) = \mathbf{1}(i = j \text{ OR } t = s)$
- Two-way HAC (Thompson 2011):
  $f(i,t,j,s) = \max\{K(t,s), \mathbf{1}(i = j \text{ OR } t = s)\}$
  with kernel $K(t,s) = \max\left(1 - \frac{|t-s|}{L+1}, 0\right)$ for some bandwidth L

We show that two-way clustering is valid under either identification condition (with either iid draws across regions or across time). Proposition

If shocks are autocorrelated, then two-way HAC may work well.

You Can Also Use Randomization Inference

Traditional inference: Residual is stochastic

Randomization inference: Shocks are stochastic; residuals held fixed (equivalently, condition on residuals)

Procedure: For a given null hypothesis, compute test statistic in sample, simulate new draws of the shocks S_t, and get simulated distribution of test statistic under the null (holding residuals fixed). Compare in-sample test statistic to simulated distribution.

Recently advocated by Borusyak and Hull (2021).

- *Advantages:* Can handle any correlation structure of residual; weak instrument robust; finite-sample valid inference
- *Disadvantages:* Need to specify DGP for shocks, S_t; finite-sample validity relies on homogeneous effects

Implementing Randomization Inference

Algorithm (Randomization Inference with One-Dimensional Shock)

To test a null hypothesis $\beta = \beta_0$*,*

1. Estimate Gaussian AR(1) process for S_t *(three parameters):*

$$S_t = \alpha + \rho S_{t-1} + \sigma \xi_t$$

where $\xi_t \sim N(0,1)$*.*

2. Simulate S_t *with random shocks* $\{\xi_t\}_{t=1}^{T}$*. Let superscript "sim" denote simulated values.*

3. Construct simulated instrument $Z_{it} = \eta_i \cdot S_t^{sim}$

4.Compare in-sample test statistic,

$$\mathcal{T} := \frac{1}{NT} \sum_{i,t} Z_{it} \left(Y_{it} - X_{it} \beta_0 \right)$$

to simulated distribution thereof.

The Factor Structure Implies an Optimal Instrument

If the residual is not homoskedastic and iid, then we can improve efficiency by reweighting to get an optimal instrument.

As shown in Chamberlain (1987, 1992) and Borusyak and Hull (2021), the optimal instrument is:

$$Z^* = \mathbb{E}\left[uu' \mid \eta\right]^{-1} \left(\mathbb{E}\left[X \mid S, \eta\right] - \mathbb{E}\left[X \mid \eta\right]\right)$$

Since our setting is one instrument and one endogenous variable, the relevant part is just $\mathbb{E}\left[uu' \mid \eta\right]^{-1}$.

Implement a *feasible* optimal instrument by

- Run principal components on residuals
- Select number of factors
- Estimate $\mathbb{E}\left[uu' \mid \eta\right]^{-1}$ and reweight instrument
- Combine whole procedure with randomization inference to get confidence intervals

Application: Regional Fiscal Multipliers

The Residual Has a Factor Structure

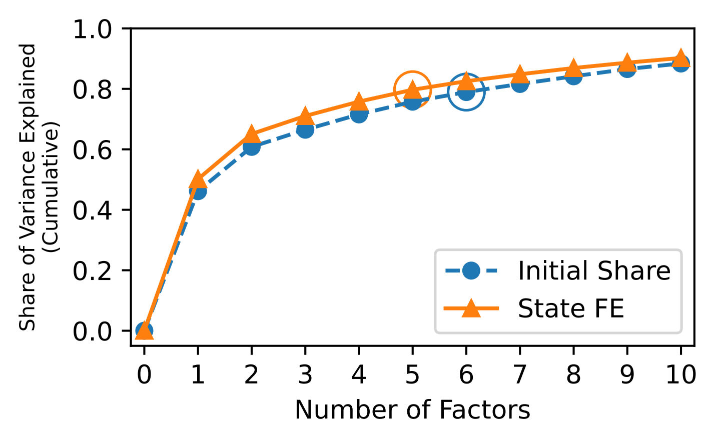

Figure 1: Share of Variance in Residual Explained by Factors

Principal components on regression residual: First two factors explain 60+%

Placebo Test: Traditional Methods Reject Too Often

Generate placebo defense shocks ($\beta = 0$), imposing the true first stage ($\pi = \hat{\pi}$)

Table 1: False Rejection Rates for Placebo Test Based on Nakamura and Steinsson (2014)

	$\text{se}_{\hat{\beta}}$	se_{β_0}
Cluster by State	25.4%	19.8%
Cluster by Year	24.4%	20.8%
Two-way Cluster	21.1%	9.0%
Two-way HAC ($L = 3$)	20.3%	3.0%
Randomization Inference	5% (By Construction)	

All Estimators & Both Strategies

Corrected Confidence Intervals Are Meaningfully Wider

Table 2: 95% Confidence Intervals for Conventional IV Estimate of Regional Fiscal Multiplier in Nakamura and Steinsson (2014)

	Initial Share Strategy (Point Estimate 2.477)	
	$\text{se}_{\hat{\beta}}$	se_{β_0}
Cluster by State	$(0.583, 4.371)$	$(0.906, \infty)$
Two-way Cluster	$(0.370, 4.583)$	$(0.712, \infty)$
Two-way HAC ($L = 3$)	$(0.045, 4.909)$	$(-\infty, \infty)$
Randomization Inference	$(0.08, 5.34)$	

90% Confidence Intervals All Estimates

Efficient Estimation Improves Power, but Gives Similar Confidence Intervals

Table 3: Unweighted Instrument vs. Feasible Optimal Instrument

	Unweighted	Feasible Optimal
Power to Reject $\beta_0 = 0$, Under True $\beta = 1.5$	0.32	0.68
Point Estimate	2.477	1.276
95% Rand. Inference. CI	$[0.08, 5.34]$	$[-0.08, 3.06]$

Feasible optimal instrument has much higher power ex ante, and gives tighter confidence intervals, but similar lower bound because point estimate is lower

Conclusion

In regional exposure designs, identification can come from shares or shocks.

When identification comes from shocks, as is typical, we need to change how we do inference.

Three Suggestions for Practice:

1. Do not cluster by region. Not robust to cross-regional correlation of residuals.
2. Instead, use a more robust method. Can use two-way clustering or two-way HAC, paired with method that computes SE around β_0. Or, do randomization inference.
3. Consider using a feasible optimal instrument. Can improve power substantially.

Consistency Under Identification from Shocks or Shares

Proposition (Convergence of the IV Estimator)

Assume that $\mathbb{E}[\tilde{Z}_{it}\tilde{X}_{it}']$ is finite and full rank (instrument relevance). The following are true:

1. If identification comes from shares and $\left(\eta_i, \lambda_i, (\varepsilon_{it})_{t=1}^{T}, (e_{it})_{t=1}^{T}\right)$ are drawn i.i.d. across regions, then $\hat{\beta} \xrightarrow{p} \beta$ as $N \to \infty$.

2. If identification comes from shocks and $\left(S_t, F_t, (\varepsilon_{it})_{i=1}^{N}, (e_{it})_{i=1}^{N}\right)$ are stationary and strongly mixing across time, then $\hat{\beta} \xrightarrow{p} \beta$ and $T \to \infty$.

Back

Identification Is Closely Tied to Inference

Identification is intimately tied to inference:

Lemma

Let $\omega(i,j,t,s) = \mathbb{E}\left[Z_{it} \cdot \lambda_i' F_t \cdot Z_{js} \cdot \lambda_j' F_s\right]$ *be the factor component covariance between units* (i,t) *and* (j,s)*. The following statements are true:*
1. If identification comes from shares and (η_i, λ_i) *is independent across regions, then* $\omega(i,j,t,s) = 0$ *for all* $i \neq j$*.*
2. If identification comes from shocks and (S_t, F_t) *is independent across time, then* $\omega(i,j,t,s) = 0$ *for all* $t \neq s$*.*

If identification comes from shares, we will want to cluster by region, while if identification comes from shocks then this won't work. Back

Identification Is Closely Tied to Inference (Part 2)

Get similar lemma for double-demeaned version:

Lemma

Let $\tilde{\omega}(i, j, t, s) = \mathbb{E}\left[\tilde{Z}_{it} \cdot \tilde{\lambda}_i' \tilde{F}_t \cdot \tilde{Z}_{js} \cdot \tilde{\lambda}_j' \tilde{F}_s\right]$ *be the demeaned-factor-component covariance between units* (i, t) *and* (j, s)*, using the double demeaned instrument. If Assumptions 1 and 2 hold, then*

1. If identification comes from shares and (η_i, λ_i) *is independent across regions, then* $\tilde{\omega}(i, j, t, s) = O(1/N^2)$ *for all* $i \neq j$*.*

2. If identification comes from shocks and (S_t, F_t) *is independent across time, then* $\tilde{\omega}(i, j, t, s) = O(1/T^2)$ *for all* $t \neq s$*.*

Back

Two-Way Clustering Is Valid Under Either Identification Condition

Proposition (Two-Way Clustering Is Valid Under Either Identification Condition)

Assume that $(\varepsilon_{it})_{t=1}^{T}$ is drawn i.i.d. across regions. Assume further either of the following:
1. Identification from shares holds and (η_i, λ_i) are i.i.d. across regions.
2.Identification from shocks holds and (S_t, F_t) are i.i.d across time.
Then, $\Omega = \Omega^{TWC}$, under the limit where $\frac{N}{T} \to C$, where C is a constant. That is,
$AVAR\left(\sqrt{N} \cdot \hat{\beta}\right) =$
$\lim_{N\to\infty, T\to\infty, \frac{N}{T}\to C} \frac{1}{NT^2} \sum_{i,t} \sum_{j,s} \mathbf{1}\left(i = j \text{ } OR \text{ } t = s\right) \mathbb{E}\left[\tilde{u}_{it}\tilde{u}_{js}\tilde{Z}_{it}\tilde{Z}_{js}\right].$

This shows that, if $\hat{\Omega}^{TWC} \xrightarrow{p} \Omega^{TWC}$, then two-way clustering is valid. Note that providing sufficient conditions such that two-way clustering converges to the expected limit is an active area of research. Back

90% Confidence Intervals

Table 4: 90% Confidence Intervals for Conventional IV Estimate of Regional Fiscal Multiplier in Nakamura and Steinsson (2014)

Initial Share Strategy (Point Estimate 2.477)		
	$\text{se}_{\hat{\beta}}$	se_{β_0}
Cluster by State	$(0.887, 4.066)$	$(1.171, \infty)$
Two-way Cluster	$(0.709, 4.245)$	$(1.095, \infty)$
Two-way HAC ($L = 3$)	$(0.436, 4.518)$	$(0.473, \infty)$
Randomization Inference	$(0.46, 4.72)$	

Back

95% CI: Full Results

Panel A: Initial Share Strategy (Point Estimate 2.477)				
	$\text{se}_{\hat{\beta}}$	se_{β_0}	AR-MD	AR-LM
State	$(0.583, 4.371)$	$(0.906, \infty)$	$(0.784, 4.959)$	$(0.906, \infty)$
Two-way	$(0.370, 4.583)$	$(0.712, \infty)$	$(0.746, 5.440)$	$(0.712, \infty)$
TWHAC	$(0.045, 4.909)$	$(-\infty, \infty)$	$(0.498, 5.879)$	$(-\infty, \infty)$
RI		$(0.08, 5.34)$		

Panel B: State FE Strategy (Point Estimate 1.426)				
	$\text{se}_{\hat{\beta}}$	se_{β_0}	AR-MD	AR-LM
State	$(0.704, 2.149)$	$(0.76, 2.75)$	Empty	$(-\infty, \infty)$
Two-way	$(0.324, 2.528)$	$(0.40, \infty)$	$(-\infty, \infty)^*$	$(-\infty, \infty)^*$
TWHAC	$(0.032, 2.821)$	$(-\infty, \infty)$	$(-\infty, \infty)^*$	$(-\infty, \infty)^*$
RI		$(-4.4, 8.5)$		

Back

Placebo Test: Full Results

	Panel A: Initial Share Strategy				Panel B: State FE Strategy			
	Conventional		Weak-IV Robust		Conventional		Weak-IV Robust	
	$se_{\hat{\beta}}$	se_{β_0}	AR-MD	AR-LM	$se_{\hat{\beta}}$	se_{β_0}	AR-MD	AR-LM
Cluster by State	25.4%	19.8%	27.8%	19.8%	27.0%	17.6%	100.0%	0.0%
Cluster by Year	24.4%	20.8%	28.4%	20.8%	28.8%	21.8%	94.2%	0.0%
Two-way Cluster	21.1%	9.0%	20.9%	9.0%	20.2%	6.8%	15.0%	2.6%
Two-way HAC	20.3%	3.0%	20.7%	3.0%	20.5%	1.4%	8.2%	2.4%
Rand. Inference	5% (By Construction)				5% (By Construction)			

Back