Supplemental Materials for "Minimizing post shock forecasting error using disparate information"

Jilei Lin*, Ziyu Liu†, and Daniel J. Eck‡

Department of Statistics, University of Illinois at Urbana-Champaign

July 8, 2020

Jilei, make sure that the bootstrap algorithm (what is now B_u) is explained correctly. Also add the description for what was the B_c procedure. However, this time refer to it as the fixed donor pool bootstrap. Add the algorithm as well. Make sure to edit the text to make sure that it integrates well with the reported tables. Also, make sure that the predicted response information is included in one table. Remove the best estimator results. Also include an introductory summary of what is in this document and put it here.

1 Supplementary materials for data analysis

The independence of the estimated September, 2008 shock-effects are further tested using likelihood ratio test (LRT) based on their estimated covariance matrix. The estimated covariance matrix is

$$\hat{\boldsymbol{\Sigma}} = \left(\begin{array}{ccc} 4.012 & 0.362 & -0.062 \\ 0.362 & 3.894 & -0.029 \\ -0.062 & -0.029 & 3.927 \end{array} \right).$$

with degrees of freedoms 35. Using hte LRT for independence between blocks of random variables [Marden, 2015, Section 10.2], the LRT test statistic is 0.304 with p-value of 0.581. Therefore, we do not reject the null hypothesis that the three estimated shock-effects are independent.

---
*jileil2@ilinois.edu

†ziyuliu3@illinois.edu

‡dje13@illinois.edu

2   Bootstrap algorithms and fixed donor pool bootstrapping

Algorithm 1: Parametric bootstrap for approximation for mean and variance of shock-effect estimators of α_1.

Input: B – the number of parametric bootstraps

$\{(y_{i,t}, \mathbf{x}_{i,t}): i = 2, \ldots, n+1, t = 0, \ldots, T_i\}$ – the data

$\{T_i^*: i = 1, \ldots, n+1\}$ – the time point just before the shock

$\{\hat{\varepsilon}_{i,t}: t = 1, \ldots, T_i\}$ – the collection of residuals for $t = 1, \ldots, T_i$

$\{\hat{\eta}_i, \hat{\alpha}_i, \hat{\phi}_i, \hat{\theta}_i, \hat{\beta}_i: i = 2, \ldots, n+1\}$ – the OLS estimates

Result: The sample mean, and sample variance of bootstrapped adjustment estimator, inverse-variance weighted estimator, and weighted-adjustment estimator.

---

1 **for** $b = 1 : B$ **do**

2 $\quad$ **for** $i = 2, \ldots, n+1$ **do**

3 $\quad\quad$ Sample with replacement from $\{\hat{\varepsilon}_{i,t}: t = 1, \ldots, T_i\}$ to obtain $\{\hat{\varepsilon}_{i,t}^{(b)}: t = 1, \ldots, T_i\}$

4 $\quad\quad$ Define $y_{i,0}^{(b)} = y_{i,0}$

5 $\quad\quad$ **for** $t = 1, \ldots, T_i$ **do**

6 $\quad\quad\quad$ Compute $y_{i,t}^{(b)} = \hat{\eta}_i + \hat{\alpha}_i 1(t = T_i^* + 1) + \hat{\phi}_i y_{i,t-1}^{(b)} + \theta_i' \mathbf{x}_{i,t} + \beta_i' \mathbf{x}_{i,t-1} + \hat{\varepsilon}_{i,t}^{(b)}$

7 $\quad\quad$ **end**

8 $\quad\quad$ Compute $\hat{\alpha}_i^{(b)}$ based on OLS estimation of the parameters in $\mathcal{M}_1$ (**does this work for the other models?**) with $\{(y_{i,t}^{(b)}, \mathbf{x}_{i,t}): t = 0, \ldots, T_i\}$

9 $\quad$ **end**

10 $\quad$ Compute the bth shock-effect estimate $\hat{\alpha}_{\text{est}}^{(b)}$ for est $\in \{\text{adj}, \text{wadj}, \text{IVW}\}$

11 **end**

12 Compute the sample mean, and sample variance of $\{\hat{\alpha}_{\text{est}}^{(b)}: b = 1, \ldots, B\}$ for est $\in \{\text{adj}, \text{wadj}, \text{IVW}\}$

Note that degeneracy of $\mathbf{W}^*$ will not occur under $\mathcal{B}_c$ because if $w_i^* = 0$ for some i and the weights are known, $w_i^* = 0$ makes it impossible for bootstrap estimate of $\hat{\alpha}_i$ to contribute to the bootstrap estimate of distribution of $\hat{\alpha}_{\text{wadj}}$.

There are some philosophical distinctions between $\mathcal{B}_u$ and $\mathcal{B}_c$. $\mathcal{B}_u$ treats the donor pool as realizations from some infinite super-population of potential donors. In contrast, $\mathcal{B}_c$ treats the donor pool as being fixed and known before the analysis is conducted, where the randomness comes from parameters and idiosyncratic error.

3   Simulations for fixed donor pool bootstrap procedure

In the first experiment, we consider the parameter combination of $n \in \{5, 10, 15, 25\}$ and $\sigma_\alpha \in \{5, 10, 25, 50, 100\}$ where we fix σ to 10. Note that $\mathrm{E}(\mathrm{E}(\alpha_1)) = 54$, where the last expectation is operated under the density of the covariates. In other words, data with $\sigma_\alpha \in \{5, 10, 25, 50, 100\}$ should well represent the situations when the signal of the covariates is strong and when it is nearly lost.

Table 1 shows the results, with the 3rd to 5th column representing the averaged $I(\hat{\Delta}(\hat{\alpha}) > 0)$ (i.e., guess of $\delta_{\hat{\alpha}}$), and the 6th to 9th columns representing the LOOCV consistency and best consistency following notations in Section **??**. First, assuming that $\bar{C}^{(k)}(\delta_{\hat{\alpha}})$ well estimates $\mathrm{E}(\mathcal{C}(\delta_{\hat{\alpha}}))$ and fixing n, we observe that $\delta_{\hat{\alpha}}$ is highly consistent for $\hat{\alpha} \in \mathcal{A}$ when σ_α is small from Table 1.

Table 1. 30 Monte Carlo simulations for $\mathcal{B}_c$ with varying n and σ_α (risk-reduction propositions $\delta_{\hat{\alpha}}$)

n	σ_α	$\hat{\alpha}_{\mathrm{adj}}$	$\hat{\alpha}_{\mathrm{wadj}}$	$\hat{\alpha}_{\mathrm{IVW}}$	$\bar{\mathcal{C}}^{(k)}(\delta_{\hat{\alpha}_{\mathrm{adj}}})$	$\bar{\mathcal{C}}^{(k)}(\delta_{\hat{\alpha}_{\mathrm{wadj}}})$	$\bar{\mathcal{C}}^{(k)}(\delta_{\hat{\alpha}_{\mathrm{IVW}}})$	$\bar{\mathcal{C}}^{(k)}(\mathcal{A})$
5*5	5	1 (0)	1 (0)	1 (0)	0.89 (0.03)	0.90 (0.02)	0.89 (0.03)	0.35 (0.04)
	10	1 (0)	1 (0)	1 (0)	0.89 (0.03)	0.89 (0.03)	0.89 (0.03)	0.37 (0.05)
	25	1 (0)	1 (0)	1 (0)	0.78 (0.03)	0.83 (0.03)	0.78 (0.03)	0.44 (0.05)
	50	0.90 (0.06)	0.97 (0.03)	0.93 (0.05)	0.66 (0.04)	0.65 (0.04)	0.65 (0.04)	0.45 (0.06)
	100	0.77 (0.08)	0.97 (0.03)	0.80 (0.07)	0.57 (0.04)	0.53 (0.04)	0.57 (0.04)	0.49 (0.05)
5*10	5	1 (0)	1 (0)	1 (0)	0.91 (0.02)	0.92 (0.02)	0.91 (0.02)	0.25 (0.03)
	10	1 (0)	1 (0)	1 (0)	0.89 (0.02)	0.89 (0.03)	0.89 (0.02)	0.29 (0.03)
	25	1 (0)	1 (0)	1 (0)	0.75 (0.03)	0.78 (0.03)	0.77 (0.04)	0.3 (0.04)
	50	0.83 (0.07)	1 (0)	0.80 (0.07)	0.59 (0.04)	0.63 (0.04)	0.59 (0.04)	0.37 (0.05)
	100	0.80 (0.07)	0.93 (0.05)	0.80 (0.07)	0.47 (0.04)	0.51 (0.04)	0.46 (0.04)	0.41 (0.05)
5*15	5	1 (0)	1 (0)	1 (0)	0.91 (0.02)	0.93 (0.02)	0.91 (0.02)	0.31 (0.05)
	10	1 (0)	1 (0)	1 (0)	0.87 (0.02)	0.89 (0.03)	0.87 (0.02)	0.31 (0.04)
	25	1 (0)	1 (0)	1 (0)	0.75 (0.03)	0.78 (0.03)	0.76 (0.03)	0.37 (0.04)
	50	0.933 (0.05)	1 (0)	0.9 (0.06)	0.61 (0.04)	0.69 (0.03)	0.64 (0.04)	0.41 (0.04)
	100	0.67 (0.09)	1 (0)	0.63 (0.09)	0.55 (0.04)	0.51 (0.04)	0.55 (0.04)	0.47 (0.04)
5*25	5	1 (0)	1 (0)	1 (0)	0.95 (0.02)	0.94 (0.02)	0.95 (0.02)	0.29 (0.04)
	10	1 (0)	1 (0)	1 (0)	0.93 (0.02)	0.91 (0.02)	0.93 (0.02)	0.30 (0.04)
	25	1 (0)	1 (0)	1 (0)	0.78 (0.04)	0.79 (0.04)	0.77 (0.04)	0.31 (0.04)
	50	0.90 (0.06)	1 (0)	0.9 (0.06)	0.57 (0.04)	0.60 (0.04)	0.58 (0.04)	0.34 (0.04)
	100	0.83 (0.07)	1 (0)	0.80 (0.07)	0.49 (0.04)	0.48 (0.04)	0.50 (0.04)	0.39 (0.03)

The reasons can be explained as follows. When σ_α is small, the signal of the covariates is strong so that $\hat{\alpha}_{\mathrm{wadj}}$ will be expected to capture the signal according to construction of $\hat{\alpha}_{\mathrm{wadj}}$ in Section **??**. Moreover, when σ_α is small, $\mathcal{M}_{22}$ approximates $\mathcal{M}_{21}$ such that estimation of $\mathrm{E}(\alpha_1)$ should be nearly unbiased according to Proposition **??**. However, when the signal of the covariates is poor (σ_α is big), the $\delta_{\hat{\alpha}}$ become more inconsistent for $\hat{\alpha} \in \mathcal{A}$. It is to be expected since the bootstrap estimates become more biased. However, users can be warned by $\bar{C}^{(k)}(\delta_{\hat{\alpha}})$ to have an idea of the effectiveness of $\delta_{\hat{\alpha}}$. Second, fixing σ_α, we can observe that $\delta_{\hat{\alpha}}$ becomes more consistent when n increases. It is due to the robustness gain in estimation when n increases.

Interestingly, we observe different patterns of $\bar{\mathcal{C}}^{(k)}(\mathcal{A})$ as opposed to $\bar{\mathcal{C}}^{(k)}(\delta_{\hat{\alpha}})$: (1) as n increases, $\bar{\mathcal{C}}^{(k)}(\mathcal{A})$ decreases; and (2) as σ_α increases, $\bar{\mathcal{C}}^{(k)}(\mathcal{A})$ increases. It is possible when n is small or σ_α is large, the differences among $\hat{\Delta}(\hat{\alpha})$ for $\hat{\alpha} \in \mathcal{A}$ are large so that the ranking gains more precision. It is because in these situations, the estimation of bootstrap estimates is less precise so that the variance of $\hat{\Delta}(\hat{\alpha})$ is larger.

Additionally, we observe that in most cases $\delta_{\hat{\alpha}_{\mathrm{wadj}}}$ reports $\hat{\alpha}_{\mathrm{wadj}}$ reduces the risk even when $\bar{\mathcal{C}}^{(k)}(\delta_{\hat{\alpha}_{\mathrm{wadj}}})$ starts to break down. Recall from Section **??** that $\hat{\Delta}(\hat{\alpha})$ contains the squared bias for estimating $\mathrm{E}(\alpha_1)$. But it is not present for $\hat{\Delta}(\hat{\alpha}_{\mathrm{wadj}})$ since we applied the fact $\hat{\alpha}_{\mathrm{wadj}}$ is unbiased for $\mathrm{E}(\alpha_1)$ from Proposition **??**. Therefore, when the signal from covariates is poorer, $\delta_{\hat{\alpha}_{\mathrm{wadj}}}$ becomes less conservative. Besides, the averaged $I(\hat{\Delta}(\hat{\alpha}) > 0)$ times $\bar{\mathcal{C}}^{(k)}(\delta_{\hat{\alpha}})$ can provide an approximation for the probability that $\hat{\alpha}$ actually reduces the risk assuming an symmetry of consistencies between the cases when $\hat{\Delta}(\hat{\alpha}) > 0$ and when $\hat{\Delta}(\hat{\alpha}) < 0$. For example, when $n = 5$ and $\sigma_\alpha = 50$, the

Table 2: 30 Monte Carlo simulations for $\mathcal{B}_c$ with varying σ and σ_α

σ	σ_α	$\hat{\alpha}_{\text{adj}}$	$\hat{\alpha}_{\text{wadj}}$	$\hat{\alpha}_{\text{IVW}}$	$\bar{\mathcal{C}}^{(k)}(\delta_{\hat{\alpha}_{\text{adj}}})$	$\bar{\mathcal{C}}^{(k)}(\delta_{\hat{\alpha}_{\text{wadj}}})$	$\bar{\mathcal{C}}^{(k)}(\delta_{\hat{\alpha}_{\text{IVW}}})$	$\bar{\mathcal{C}}^{(k)}(\mathcal{A})$
5*5	5	1 (0)	1 (0)	1 (0)	0.99 (0.01)	0.99 (0.01)	0.99 (0.01)	0.4 (0.04)
	10	1 (0)	1 (0)	1 (0)	0.91 (0.02)	0.95 (0.02)	0.92 (0.02)	0.46 (0.04)
	25	0.97 (0.03)	1 (0)	0.97 (0.03)	0.83 (0.03)	0.89 (0.02)	0.83 (0.03)	0.43 (0.04)
	50	0.87 (0.06)	1 (0)	0.87 (0.06)	0.67 (0.04)	0.69 (0.03)	0.67 (0.04)	0.38 (0.04)
	100	0.7 (0.09)	0.93 (0.05)	0.77 (0.08)	0.52 (0.04)	0.51 (0.04)	0.51 (0.05)	0.51 (0.05)
5*10	5	1 (0)	1 (0)	1 (0)	0.93 (0.02)	0.93 (0.02)	0.93 (0.02)	0.27 (0.03)
	10	1 (0)	1 (0)	1 (0)	0.86 (0.03)	0.89 (0.03)	0.87 (0.03)	0.37 (0.04)
	25	0.97 (0.03)	1 (0)	0.97 (0.03)	0.78 (0.03)	0.82 (0.03)	0.79 (0.03)	0.29 (0.04)
	50	0.93 (0.05)	1 (0)	0.9 (0.06)	0.66 (0.04)	0.67 (0.04)	0.66 (0.04)	0.32 (0.04)
	100	0.77 (0.08)	0.97 (0.03)	0.7 (0.09)	0.54 (0.05)	0.54 (0.04)	0.55 (0.04)	0.43 (0.05)
5*25	5	0.97 (0.03)	1 (0)	1 (0)	0.78 (0.03)	0.77 (0.03)	0.79 (0.03)	0.22 (0.03)
	10	1 (0)	1 (0)	1 (0)	0.81 (0.03)	0.79 (0.03)	0.81 (0.03)	0.21 (0.03)
	25	1 (0)	1 (0)	1 (0)	0.72 (0.03)	0.77 (0.03)	0.71 (0.03)	0.25 (0.04)
	50	0.87 (0.06)	0.97 (0.03)	0.87 (0.06)	0.55 (0.04)	0.57 (0.04)	0.52 (0.04)	0.35 (0.04)
	100	0.9 (0.06)	1 (0)	0.87 (0.06)	0.54 (0.04)	0.54 (0.04)	0.56 (0.04)	0.39 (0.05)
5*50	5	0.83 (0.07)	0.73 (0.08)	0.8 (0.07)	0.54 (0.04)	0.57 (0.04)	0.55 (0.04)	0.19 (0.03)
	10	0.77 (0.08)	0.8 (0.07)	0.77 (0.08)	0.53 (0.04)	0.55 (0.05)	0.53 (0.04)	0.23 (0.03)
	25	0.73 (0.08)	0.8 (0.07)	0.73 (0.08)	0.58 (0.04)	0.58 (0.04)	0.57 (0.04)	0.21 (0.03)
	50	0.8 (0.07)	0.83 (0.07)	0.8 (0.07)	0.59 (0.04)	0.53 (0.04)	0.6 (0.04)	0.29 (0.05)
	100	0.5 (0.09)	0.73 (0.08)	0.53 (0.09)	0.51 (0.05)	0.48 (0.05)	0.51 (0.05)	0.31 (0.04)
5*100	5	0.43 (0.09)	0.37 (0.09)	0.43 (0.09)	0.47 (0.04)	0.47 (0.04)	0.47 (0.04)	0.19 (0.03)
	10	0.63 (0.09)	0.6 (0.09)	0.67 (0.09)	0.51 (0.04)	0.49 (0.04)	0.51 (0.04)	0.23 (0.03)
	25	0.57 (0.09)	0.6 (0.09)	0.57 (0.09)	0.51 (0.04)	0.53 (0.04)	0.49 (0.04)	0.21 (0.04)
	50	0.53 (0.09)	0.43 (0.09)	0.5 (0.09)	0.47 (0.04)	0.49 (0.05)	0.49 (0.05)	0.19 (0.03)
	100	0.63 (0.09)	0.63 (0.09)	0.63 (0.09)	0.43 (0.04)	0.41 (0.04)	0.45 (0.04)	0.17 (0.03)

where we fix n to 10. Likewise, $\sigma, \sigma_\alpha \in \{5, 10, 25, 50, 100\}$ will produce situations when the signal of the covariates is strong and when it is nearly lost in the model of both $y_{i,t}$ and α_i.

From Table 2, we observe that as σ_α increases fixing σ, $\bar{\mathcal{C}}(\delta_{\hat{\alpha}})$ and $\bar{\mathcal{C}}(\mathcal{A})$ will decrease, which is a pattern similar to one shown in the first experiment. Furthermore, as σ increases fixing σ_α, $\bar{\mathcal{C}}(\delta_{\hat{\alpha}})$ and $\bar{\mathcal{C}}(\mathcal{A})$ decrease as well. Note that the consistency or best consistency hinges on the estimation of the parameters. Since $\hat{\alpha}_{\text{wadj}}$ is a linear combination of OLS estimates, as σ increases, $\text{Var}(\hat{\alpha}_{\text{wadj}})$ increases as well. Therefore, $\hat{\alpha}_{\text{wadj}}$ become more volatile and its estimation of $\text{E}(\alpha_1)$ can be less reliable. Those reasons can explain why an increase of σ_α contributes to a decrease of $\bar{\mathcal{C}}(\delta_{\hat{\alpha}})$ and $\bar{\mathcal{C}}(\mathcal{A})$.

With respect to averaged $I(\hat{\Delta}(\hat{\alpha} > 0))$ (i.e., the guess), it starts to decrease as σ increases. It is reasonable if we believe the bootstrap estimate $S_{\hat{\alpha}}^2$ provides a good approximation for $\text{Var}(\hat{\alpha})$ for $\hat{\alpha} \in \mathcal{A}$. The reasons can be outlined as below. Recall in Section **??**, the conditions of risk-reduction propositions involve $(\text{E}(\alpha_1))^2 > \text{Var}(\hat{\alpha}) + (\text{E}(\hat{\alpha}) - \text{E}(\alpha_1))^2$ for $\hat{\alpha} \in \mathcal{A}$, where we note that those parameters are *true* ones but not estimated ones. Notice that $\text{Var}(\hat{\alpha})$ is a function of σ since $\hat{\alpha}$ is estimated by OLS. Therefore, it explains the reason why the increase of σ would result in a decrease of averaged $I(\hat{\Delta}(\hat{\alpha} > 0))$ since the conditions are not likely to hold when $\text{Var}(\hat{\alpha})$ increases.

Table 3: 30 Monte Carlo simulation with varying n and σ_α (distance to y_{1,T_1^*+1})

n	σ_α	Original	$\hat{\alpha}_{\text{adj}}$	$\hat{\alpha}_{\text{wadj}}$	$\hat{\alpha}_{\text{IVW}}$
5*5	5	51.92 (4.04)	19.23 (2.55)	20.64 (2.76)	19.36 (2.52)
	10	52.58 (4.35)	21.565 (2.58)	23.38 (2.86)	21.72 (2.50)
	25	55.02 (5.84)	30.30 (3.59)	31.99 (4.36)	30.20 (3.50)
	50	64.42 (8.11)	50.00 (5.78)	52.56 (6.92)	49.55 (5.70)
	100	91.51 (13.20)	94.27 (10.41)	97.26 (12.57)	93.61 (10.27)
5*10	5	52.34 (4.00)	17.23 (2.96)	18.55 (2.84)	17.39 (2.95)
	10	52.59 (4.05)	19.02 (3.24)	20.95 (3.14)	19.23 (3.24)
	25	54.07 (4.97)	27.66 (4.55)	31.79 (4.47)	27.85 (4.61)
	50	60.32 (7.53)	47.70 (7.04)	52.97 (7.54)	47.78 (7.17)
	100	85.60 (12.99)	89.86 (12.82)	100.74 (13.61)	90.40 (12.91)
5*15	5	49.86 (4.01)	18.07 (2.88)	18.38 (2.71)	18.04 (2.88)
	10	48.73 (4.30)	19.45 (2.97)	19.35 (2.86)	19.32 (3.00)
	25	47.06 (5.13)	26.16 (3.31)	26.81 (3.13)	26.23 (3.33)
	50	48.75 (6.86)	40.27 (4.43)	42.09 (4.49)	40.77 (4.38)
	100	64.29 (11.11)	68.85 (8.27)	74.08 (8.72)	69.91 (8.14)
5*25	5	57.60 (6.94)	21.58 (5.90)	20.00 (5.86)	21.58 (5.90)
	10	56.80 (7.01)	22.20 (5.97)	20.47 (5.89)	22.22 (5.97)
	25	56.58 (7.49)	28.96 (6.49)	27.06 (6.19)	29.03 (6.48)
	50	64.33 (8.75)	47.16 (8.28)	46.30 (7.26)	47.35 (8.26)
	100	95.61 (13.02)	90.10 (13.29)	86.81 (11.75)	90.55 (13.23)

Simulation for $\mathcal{B}_u$ with the same parameter setup as that of $\mathcal{B}_c$ are implemented. See table **??** and table **??** for results. Comparing table 1 and table **??** yields that when n is small ($n = 5$ or $n = 10$) and σ_α is small ($\sigma_\alpha = 5$), $\mathcal{B}_u$ is better than $\mathcal{B}_c$ with statistical evidence. For other situations, $\mathcal{B}_u$ and $\mathcal{B}_c$ are rather similar. It is likely that the extra randomness from sampling with replacement from donor pool compensates for the possible noises from a small donor pool. Concerning table 2 and table **??**, it appears that when $n = 10$ and $\sigma_\alpha = 5$, $\mathcal{B}_u$ is better than $\mathcal{B}_c$ when σ increases. It might be the case that additional layer of bootstrap in the donor pool buffers the negative effects on $\bar{\mathcal{C}}(\delta_{\hat{\alpha}})$ introduced from increasing variation of $y_{i,t}$. However, when σ_α increases over 5 and $n = 10$, $\mathcal{B}_c$ and $\mathcal{B}_u$ are quite similar under situations of different σ and σ_α. In conclusion, $\mathcal{B}_u$ is better than $\mathcal{B}_c$ when the signal of the covariates is strong and n is small; otherwise, they are similar. See more discussions for differences between $\mathcal{B}_u$ and $\mathcal{B}_c$ in Section **??**.

References

John I Marden. Multivariate statistics: Old school. *University of Illinois*, 2015.

$\mathcal{B}_c$ when σ increases. It might be the case that additional layer of bootstrap in the donor pool buffers the negative effects on $\bar{\mathcal{C}}(\delta_{\hat{\alpha}})$ introduced from increasing variation of $y_{i,t}$. However, when σ_{α} increases over 5 and $n = 10$, $\mathcal{B}_c$ and $\mathcal{B}_u$ are quite similar under situations of different σ and σ_{α}. In conclusion, $\mathcal{B}_u$ is better than $\mathcal{B}_c$ when the signal of the covariates is strong and n is small; otherwise, they are similar. See more discussions for differences between $\mathcal{B}_u$ and $\mathcal{B}_c$ in Section **??**.

References

John I Marden. Multivariate statistics: Old school. *University of Illinois*, 2015.