sealed envelope™

Sealed Envelope Ltd, Clerkenwell Workshops, 27-31 Clerkenwell Close, London EC1R 0AT, UK
https://www.sealedenvelope.com, email: contact@sealedenvelope.com, tel: +44 (0)20 3176 4242

Comparison of minimisation methods for unbalanced allocation ratios

Jon Ramsey and Tony Brady, Sealed Envelope Ltd. December 2018

Background

Research published in the last decade has highlighted some issues when using the standard minimisation method (Pocock and Simon 1975) in trials with unbalanced allocation ratios. It has been shown (Han, Enas, and McEntegart 2009) that the unequal ratio must be taken into account when calculating the allocation probabilities, otherwise the method can lead to deviations from the target allocation ratio, especially when the allocation probability given to the preferred treatment is relatively low. In addition, the randomisation (permutation) test (Ludbrook and Dudley 1998) is affected by minimisation with unequal allocation ratios (Proschan, Brittain, and Kammerman 2011). The randomisation distribution does not necessarily have a mean of 0, especially under strict minimisation (allocation probability of 1). This can lead to a substantial loss of power for the randomisation test. It has been demonstrated (Kuznetsova and Tymofyeyev 2011) that this problem occurs with unequal randomisation methods that do not preserve the unconditional allocation ratio at every allocation step.

One method proposed to tackle these problems (Han, Enas, and McEntegart 2009) results in an allocation ratio at the end of the study that is very close to the targeted one. However, the unconditional allocation ratio still varies from allocation to allocation. In theory this could lead to selection or evaluation bias, and reduces the power of the randomisation test.

We compare the behaviour of two minimisation methods, SBM (Madurasinghe 2017) and ARP (Kuznetsova and Tymofyeyev 2011), that are designed to preserve the unconditional allocation ratio at every allocation step for trials with unbalanced ratios. In a covariate-adaptive procedure this means preserving the unconditional allocation ratio at every allocation step for any given sequence of covariates.

Methods

We carried out simulations to compare our implementations of the sequence balance minimisation method (SBM) (Madurasinghe 2017) with the allocation ratio preserving biased coin method (ARP) (Kuznetsova and Tymofyeyev 2011). We also compare the published data for the SBM method with our own implementation.

Both SBM and ARP are minimisation methods that preserve the allocation ratio at every randomisation even when the ratio is unequal. SBM restricts its calculation of imbalance scores to the current allocation block. An allocation block, S, is defined as the number of allocations equal to the sum of desired allocation ratios (e.g. 3 for 2:1 ratio). It also calculates the respective assigning probabilities for each treatment on the basis of its imbalance score. ARP is a simple expansion of the standard minimisation algorithm (Pocock and Simon 1975) to execute an equal allocation to S 'fake' treatment arms, chosen to represent the treatments with the same weighting as the allocation ratio. For instance, 2:1 allocation of treatments A:B would be carried out by randomisation of treatments F_1, F_2 and F_3, where F_1 and F_2 represent A and F_3 represents B.

The published results data from the SBM paper were scraped from the tables in the HTML version.

For each of our implementations we carried out 1000 simulations of 630 trials parameterised to match those detailed in the SBM paper, with parameters

- Sample size $N = 30, 60, 120$.
- The number of equally weighted binary prognostic factors $F = 1, 2, \ldots, 10$.
- An allocation probability $p = 0.5, 0.6, 0.7, 0.8, 0.85, 0.9, 0.95$.
- A weighting assigned to total treatment group counts used as a balancing factor, $W = 0, 1, F$.

Every trial had a $1:2$ allocation ratio of treatment $T1$ vs $T2$.

The levels $l \in \{ 1, 2 \}$ for each prognostic factor were assigned randomly with equal probability before each simulated randomisation.

Note that the allocation probability parameter p is not directly comparable between the randomisation methods. For SBM the parameter expresses the high probability of assignment to the preferred treatment (Han, Enas, and McEntegart 2009), but this parameter is used only when the algorithm would otherwise certainly allocate a specific treatment. For ARP the parameter always describes the probability of assignment to the preferred treatment group as determined by the minimisation algorithm.

The SBM results are denoted SBM1 for the simulation results presented in (Madurasinghe 2017), and SBM2 for our own simulation results.

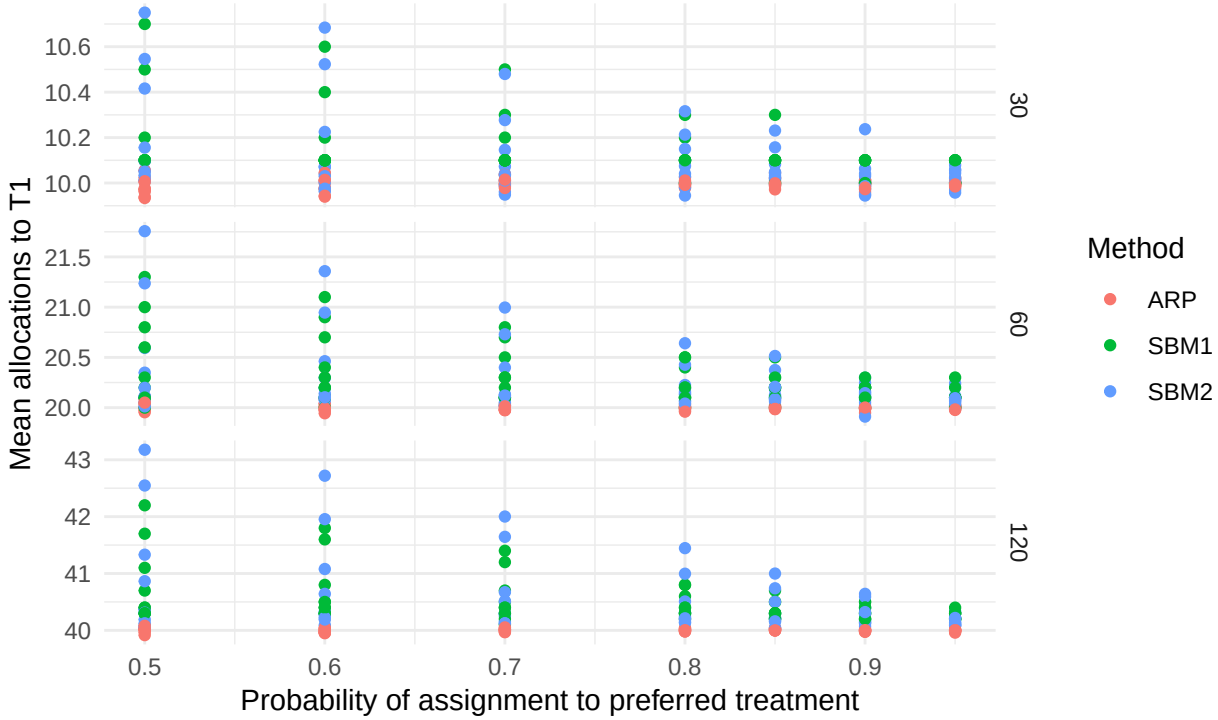
Results

Treatment group totals not included as a balancing factor

First we look at simulations with weighting for treatment totals as a balancing factor $W = 0$.

Figure 1. Mean allocations to group T1 for N = 30, 60, 120

Simulation results with 1 to 10 balancing factors for each sample size and probability



We expect mean allocations to group T1 of 10, 20, and 40 for $N = 30,\ 60,\ 120$ respectively (Figure 1). The results for ARP are noticeably closer to the desired mean for all probabilities across all sample sizes. The means for SBM skew high but get closer to the desired mean as p increases. The upwards bias in the mean for SBM appears to scale with sample size.

Figure 2. Standard error of allocations to group T1 for N = 30, 60, 120

Simulation results with 1 to 10 balancing factors for each sample size and probability

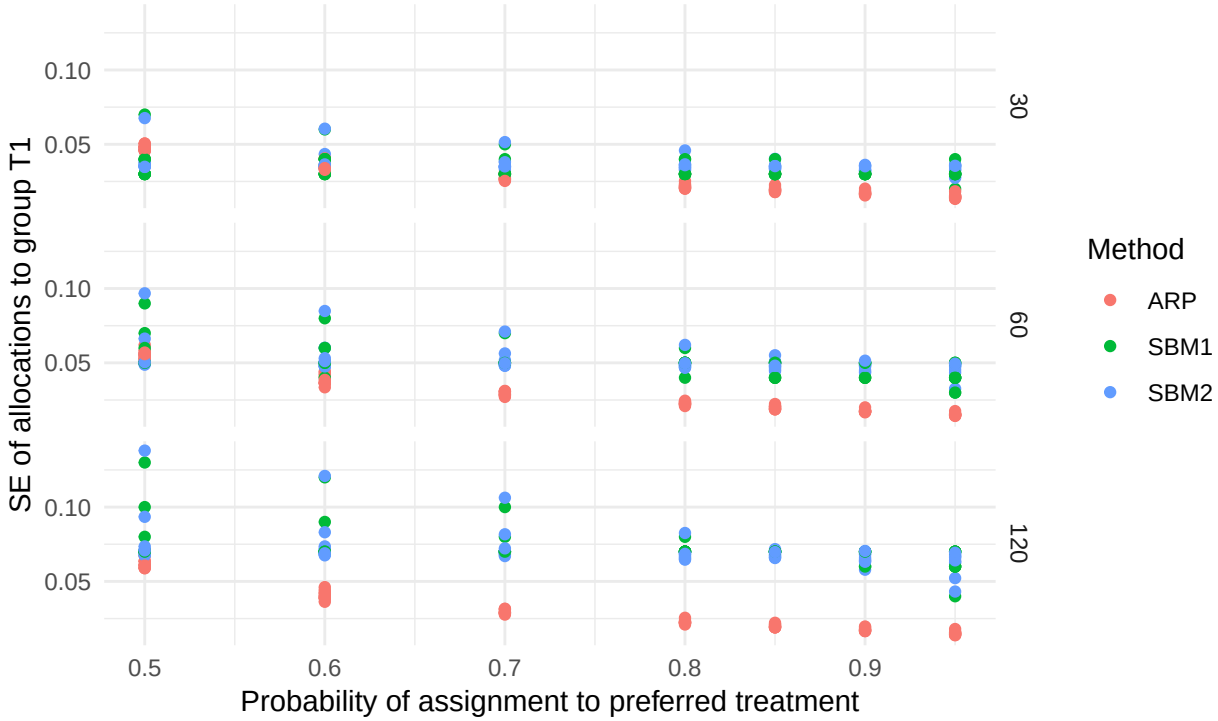


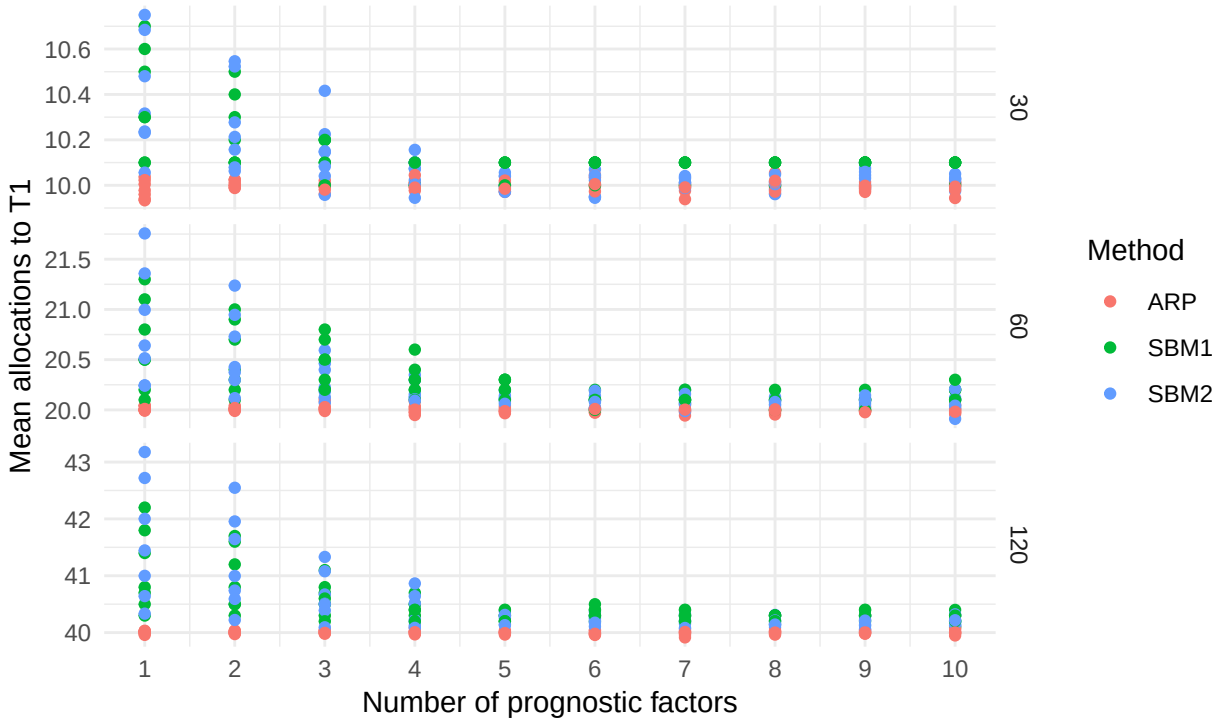
Table 1: Table 1. Standard error of mean allocations to group $T1$ (ARP method with 2 factors)

p	Sample size	SE
0.50	30	0.0504973
0.50	60	0.0583210
0.50	120	0.0671098
0.60	30	0.0405419
0.60	60	0.0432741
0.60	120	0.0461610
0.70	30	0.0311902
0.70	60	0.0311173
0.70	120	0.0315264
0.80	30	0.0234291
0.80	60	0.0236956
0.80	120	0.0236126
0.85	30	0.0209609
0.85	60	0.0207444
0.85	120	0.0210803
0.90	30	0.0186340
0.90	60	0.0180644
0.90	120	0.0192274
0.95	30	0.0153700
0.95	60	0.0153908
0.95	120	0.0161321

The standard error of the mean appears to be smaller for the ARP method (Figure 2, Table 1), indicating less variation in the number of allocations to $T1$ across simulations.

Figure 3. Mean allocations to group T1 for N = 30, 60, 120

Simulation results with varying p for each sample size and factor count



We expect mean allocations to group T1 of 10, 20, and 40 for $N = 30,\ 60,\ 120$ respectively (Figure 3). This shows increasing the number of covariates used in the balancing algorithm brings the simulation results for SBM closer to the desired mean.

Treatment group totals weighted 1

For simulations with $W = 1$ we find broadly similar results (Figure 4, Table 2, Figure 6), but the SE is noticeably smaller by inclusion of the treatment totals (Figure 5).

Figure 4. Mean allocations to group T1 for N = 30, 60, 120

Simulation results with 1 to 10 balancing factors for each sample size and probability

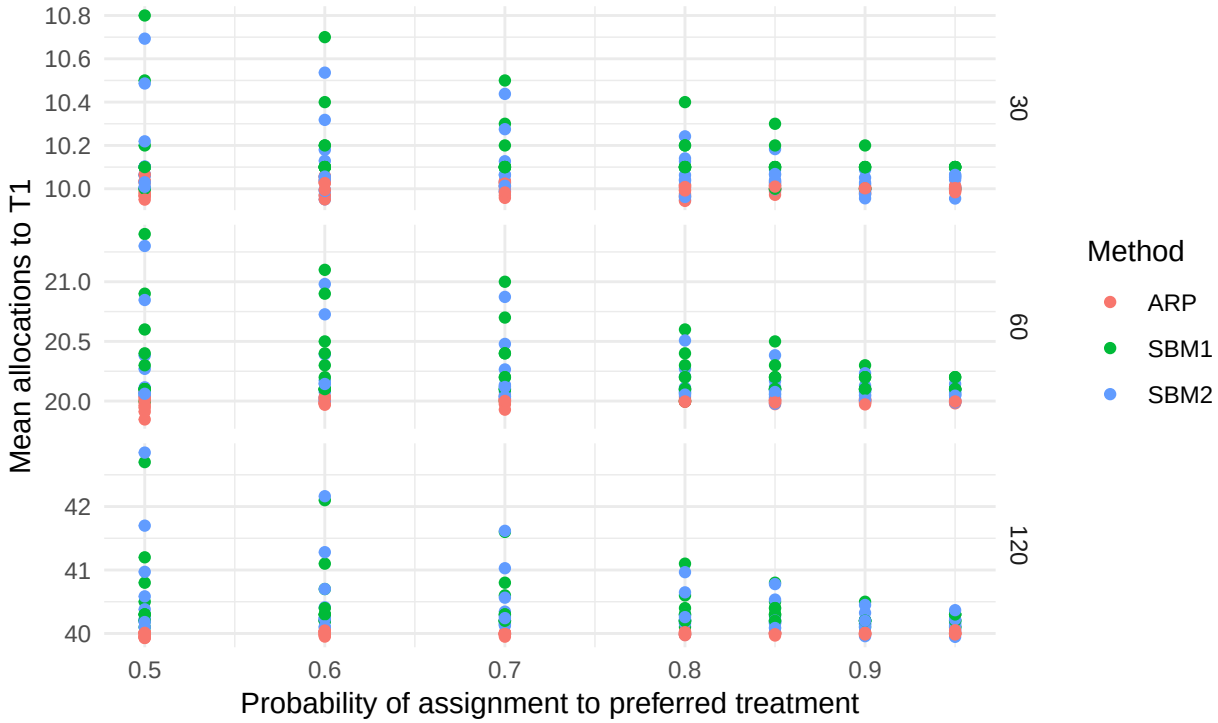


Figure 5. Standard error of allocations to group T1 for N = 30, 60, 120

Simulation results with 1 to 10 balancing factors for each sample size and probability

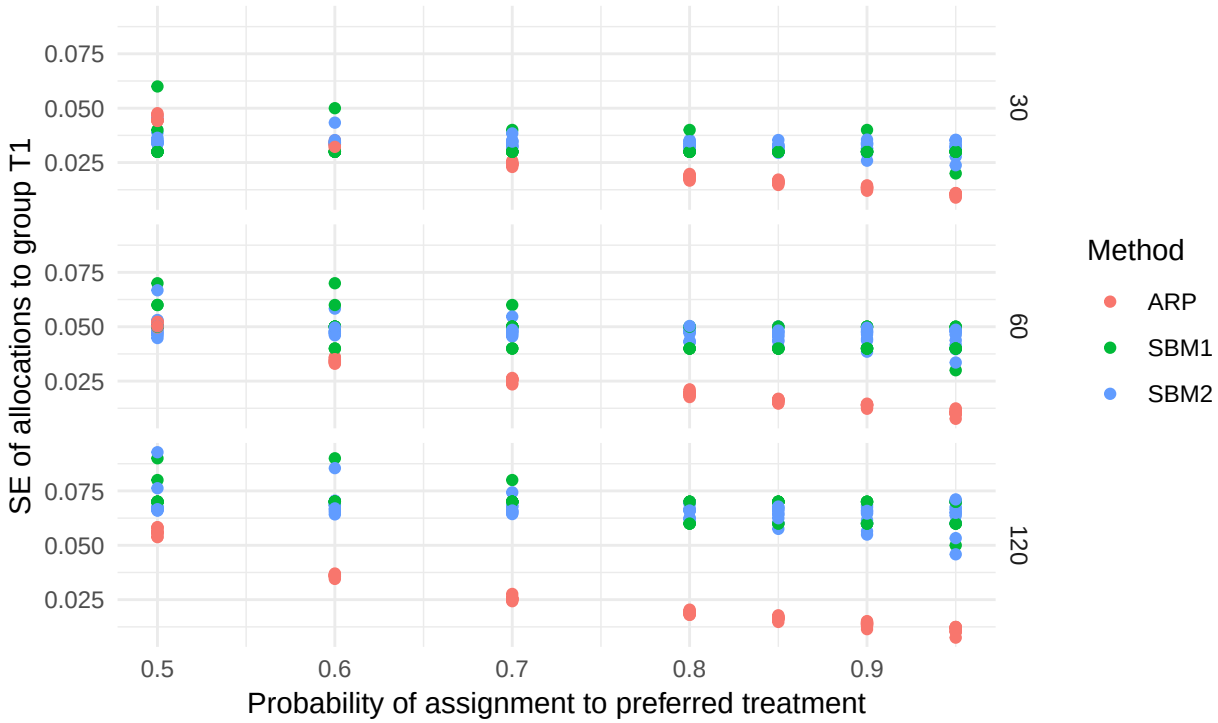
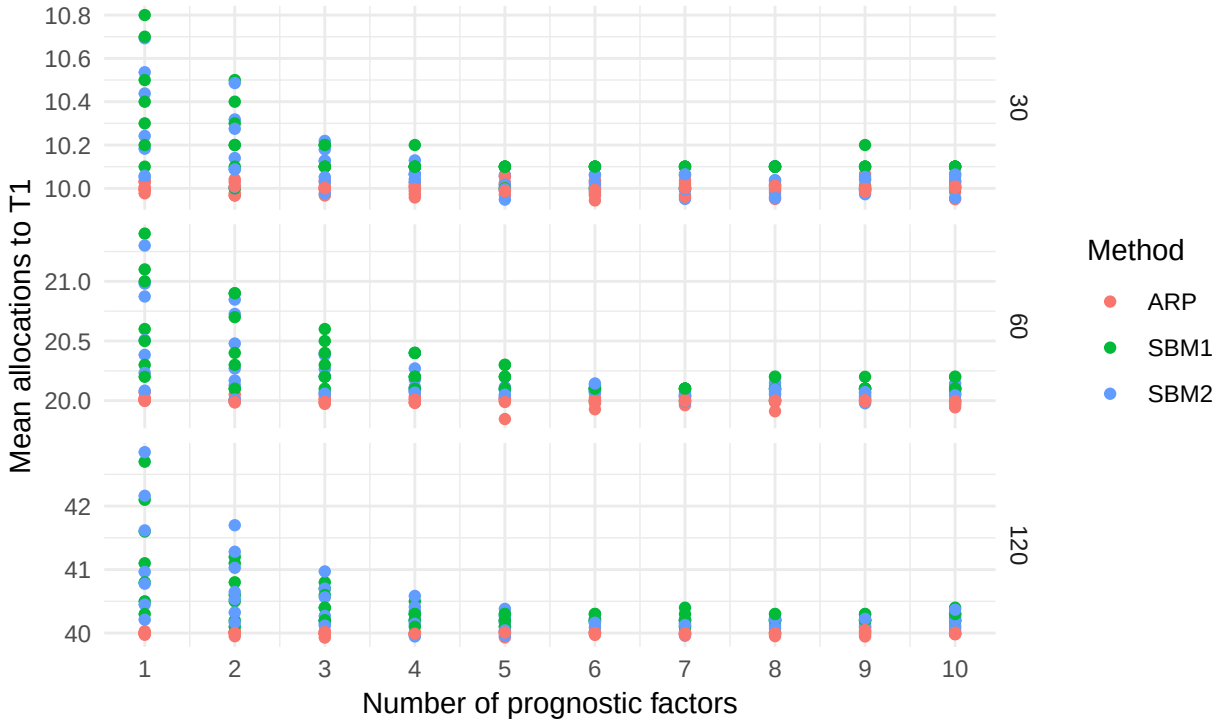


Table 2: Table 2. Standard error of allocations to group $T1$ (ARP method with 2 factors)

p	Sample size	SE
0.50	30	0.0444735
0.50	60	0.0508661
0.50	120	0.0581748
0.60	30	0.0345171
0.60	60	0.0347879
0.60	120	0.0355061
0.70	30	0.0242641
0.70	60	0.0259995
0.70	120	0.0254194
0.80	30	0.0173266
0.80	60	0.0190882
0.80	120	0.0185250
0.85	30	0.0152719
0.85	60	0.0149398
0.85	120	0.0163172
0.90	30	0.0129088
0.90	60	0.0127704
0.90	120	0.0127732
0.95	30	0.0100509
0.95	60	0.0098476
0.95	120	0.0103488

Figure 6. Mean allocations to group T1 for N = 30, 60, 120

Simulation results with varying p for each sample size and factor count



Treatment group totals weighted as prognostic factor count

The results for simulations with W=F (Figure 7, Figure 8, Table 3, Figure 9) are almost identical to those with W=1.

Figure 7. Mean allocations to group T1 for N = 30, 60, 120

Simulation results with 1 to 10 balancing factors for each sample size and probability

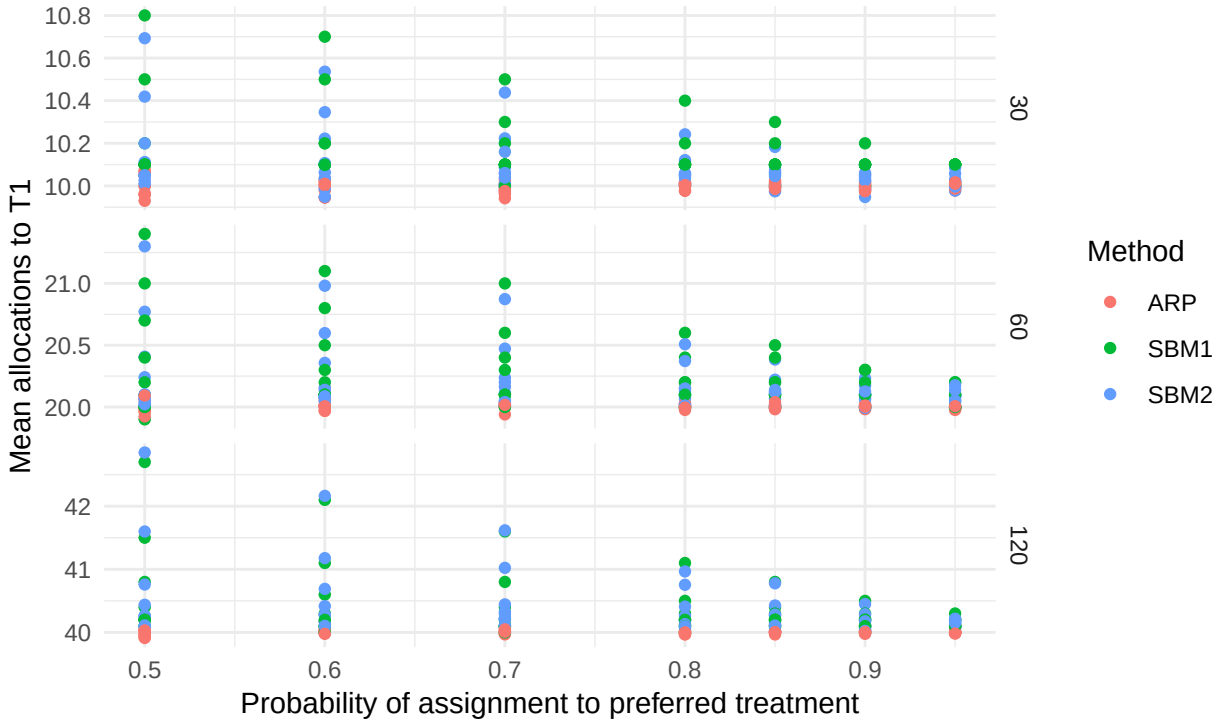


Figure 8. Standard error of allocations to group T1 for N = 30, 60, 120

Simulation results with 1 to 10 balancing factors for each sample size and probability

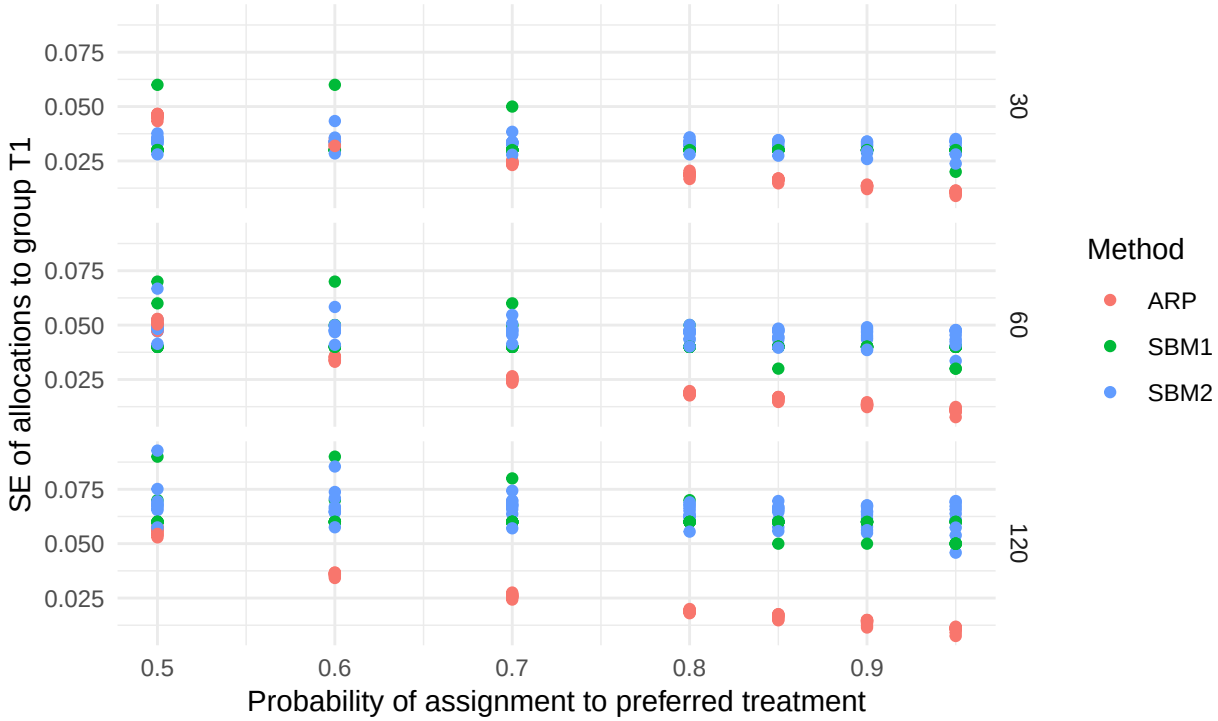
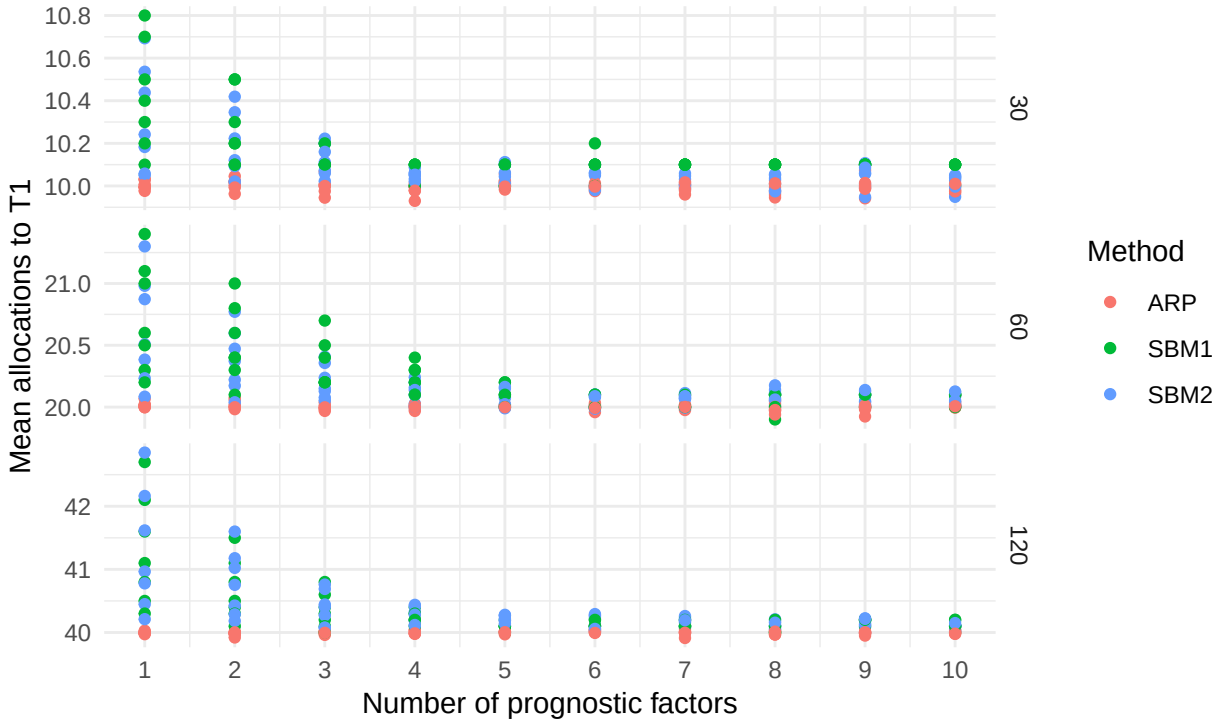


Table 3: Table 3. Standard error of allocations to group $T1$ (ARP method with 2 factors)

p	Sample size	SE
0.50	30	0.0465623
0.50	60	0.0527202
0.50	120	0.0569386
0.60	30	0.0337878
0.60	60	0.0344174
0.60	120	0.0342195
0.70	30	0.0247435
0.70	60	0.0259282
0.70	120	0.0261672
0.80	30	0.0186906
0.80	60	0.0179532
0.80	120	0.0188722
0.85	30	0.0163729
0.85	60	0.0149066
0.85	120	0.0150700
0.90	30	0.0128376
0.90	60	0.0132244
0.90	120	0.0122529
0.95	30	0.0099524
0.95	60	0.0106354
0.95	120	0.0107755

Figure 9. Mean allocations to group T1 for N = 30, 60, 120

Simulation results with varying p for each sample size and factor count



Position of group allocations within blocks

Within blocks of three sequential randomisations (since we have a $1:2$ allocation ratio) we expect allocation counts for treatment group $T1$ in each of the three possible block positions to be equal. To test this we assigned a block position number $1, 2, 3$ to each randomisation in sequence and counted the number of allocations to $T1$ at each position.

Treatment group totals weight $W = 0$

We first look at results obtained when treatment totals are weighted zero, sample size is 30, and there is one prognostic factor. With $N = 30$, $F = 1$, the level of the prognostic factor 1, and $W = 0$, we expect the count allocated to group $T1$ at each block position to equal $30 \times 1000 \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{2} = 1667$.

SBM2

For each probability p we checked that group $T1$ was allocated equally between the three block positions using a chi-squared test.

Table 4: Table 4. Chi-squared test results comparing allocation counts to $T1$ by block position, for different p

p	Chi-squared test statistic	P-value
0.50	0.6927644	0.7072421
0.60	1.4685682	0.4798489
0.70	1.8454493	0.3974347
0.80	2.8533178	0.2401098
0.85	0.6187304	0.7339127
0.90	1.9726242	0.3729496

p	Chi-squared test statistic	P-value
0.95	4.8658657	0.0877790

The P-values (Table 4) show that there is no evidence that allocation to group $T1$ is correlated with block position. The count data is given below (Table 5). This demonstrates that the SBM method preserves the unconditional allocation ratio at every allocation step.

Table 5: Table 5. Count of allocations to group $T1$ in each block position by p

p	block position	count
0.50	1	1825
0.50	2	1787
0.50	3	1778
0.60	1	1742
0.60	2	1773
0.60	3	1814
0.70	1	1775
0.70	2	1765
0.70	3	1701
0.80	1	1677
0.80	2	1773
0.80	3	1704
0.85	1	1675
0.85	2	1712
0.85	3	1717
0.90	1	1752
0.90	2	1680
0.90	3	1682
0.95	1	1759
0.95	2	1631
0.95	3	1687

The excess allocation to group $T1$ for the SBM method is apparent at lower values of p. This is consistent with the biased mean discussed earlier.

ARP

We repeat the chi-squared test for equal allocation of $T1$ to each of the three block positions.

Table 6: Table 6. Chi-squared test results comparing allocation counts to $T1$ by block position, for different p

p	Chi-squared test statistic	P-value
0.50	0.1767889	0.9153997
0.60	0.0867480	0.9575532
0.70	3.1921005	0.2026955
0.80	1.7057407	0.4261899
0.85	5.2154762	0.0737011
0.90	0.7570056	0.6848861
0.95	2.2089136	0.3313908

Again the P-values (Table 6) show that there is no evidence that allocation to group $T1$ is correlated with block position. The count data is given below (Table 7). This demonstrates that the ARP method preserves the unconditional allocation ratio at every allocation step.

Table 7: Table 7. Count of allocations to group $T1$ in each block position by p

p	block position	count
0.50	1	1677
0.50	2	1656
0.50	3	1656
0.60	1	1674
0.60	2	1658
0.60	3	1671
0.70	1	1730
0.70	2	1634
0.70	3	1649
0.80	1	1631
0.80	2	1703
0.80	3	1648
0.85	1	1620
0.85	2	1669
0.85	3	1751
0.90	1	1659
0.90	2	1644
0.90	3	1693
0.95	1	1629
0.95	2	1683
0.95	3	1714

Treatment group totals weight $W = F$

With treatment totals weighted as the number of prognostic factors, sample size 120, 10 prognostic factors, and the level of all prognostic factors limited to 1, we have $N = 120$, $F = 10$, and $W = F$, and we expect the count allocated to group $T1$ at each block position to equal $120 \times 1000 \times \frac{1}{3} \times \frac{1}{3} \times \left(\frac{1}{2}\right)^{10} = 13$.

SBM2

For each probability p we checked that group $T1$ was allocated equally between the three block positions using a chi-squared test.

Table 8: Table 8. Chi-squared test results comparing allocation counts to $T1$ per block position, by p

p	Chi-squared test statistic	P-value
0.50	1.3571429	0.5073412
0.60	4.1276596	0.1269668
0.70	1.2000000	0.5488116
0.80	0.0645161	0.9682567
0.85	0.3500000	0.8394570
0.90	2.9230769	0.2318793
0.95	3.6470588	0.1614549

The P-values (Table 8) show that there is no evidence that group $T1$ is not evenly allocated across block positions. The count data is given below (Table 9).

Table 9: Table 9. Count of allocations to group $T1$ in each block position by p

p	block position	count
0.50	1	12
0.50	2	9
0.50	3	7
0.60	1	11
0.60	2	22
0.60	3	14
0.70	1	12
0.70	2	15
0.70	3	18
0.80	1	10
0.80	2	10
0.80	3	11
0.85	1	12
0.85	2	15
0.85	3	13
0.90	1	11
0.90	2	10
0.90	3	18
0.95	1	16
0.95	2	12
0.95	3	23

ARP

We repeat the chi-squared test for equal allocation of $T1$ to each of the three block positions.

Table 10: Table 10. Chi-squared test results comparing allocation counts to $T1$ per block position, by p

p	Chi-squared test statistic	P-value
0.50	2.0000000	0.3678794
0.60	0.4905660	0.7824830
0.70	1.6818182	0.4313182
0.80	1.2258065	0.5417757
0.85	0.1666667	0.9200444
0.90	0.0487805	0.9759048
0.95	0.8387097	0.6574709

Again the P-values (Table 10) show that there is no evidence that group $T1$ is not evenly allocated across block positions. The count data is given below (Table 11).

Table 11: Table 11. Count of allocations to group $T1$ in each block position by p

p	block position	count
0.50	1	12
0.50	2	16
0.50	3	20
0.60	1	20
0.60	2	16
0.60	3	17
0.70	1	11
0.70	2	18
0.70	3	15
0.80	1	10
0.80	2	13
0.80	3	8
0.85	1	12
0.85	2	11
0.85	3	13
0.90	1	14
0.90	2	14
0.90	3	13
0.95	1	12
0.95	2	8
0.95	3	11

Conclusions

The results indicate that ARP is likely a better method than SBM for minimisation in trials with an unbalanced allocation ratio of $1:2$.

Both the original published SBM results and those from our own simulations demonstrated a bias towards allocating to group $T1$ at lower values of p, where ARP was consistently close to the desired mean at all tested values of p. Choice for the value of p may be important for preventing allocation prediction. The lower standard errors for ARP are also an advantage, as it shows that the method gets closer to the desired $1:2$ ratio more consistently.

Additionally there appears to be some discrepancy in the results for our implementation of SBM as compared with the published results and we have been unable to determine the cause (see Appendix).

As a simple extension of the minimisation method (Pocock and Simon 1975) ARP has the benefit of being easier to understand than SBM. It takes account of all randomisations to date, which allows randomisations marked in error to be retrospectively excluded from the calculation. This is generally not possible with SBM as it only considers randomisations in the current block.

(Kuznetsova and Tymofyeyev 2011) suggest that the ARP method may be unsuitable for smaller studies where the allocation ratio leads to a large block size. Further work would be useful to establish whether the SBM method has an advantage in this situation, as claimed by the author.

References

Han, Baoguang, Nathan H. Enas, and Damian McEntegart. 2009. "Randomization by Minimization for Unbalanced Treatment Allocation." *Statistics in Medicine* 28 (27): 3329–46. https://doi.org/10.1002/sim.3710.

Kuznetsova, Olga M., and Yevgen Tymofyeyev. 2011. "Preserving the Allocation Ratio at Every Allocation with Biased Coin Randomization and Minimization in Studies with Unequal Allocation." *Statistics in Medicine* 31 (8): 701–23. https://doi.org/10.1002/sim.4447.

Ludbrook, John, and Hugh Dudley. 1998. "Why Permutation Tests Are Superior to T and F Tests in Biomedical Research." *The American Statistician* 52 (2): 127–32. http://www.jstor.org/stable/2685470.

Madurasinghe, Vichithranie W. 2017. "Sequence Balance Minimisation: Minimising with Unequal Treatment Allocations." *Trials* 18 (1): 207. https://doi.org/10.1186/s13063-017-1942-3.

Pocock, Stuart J., and Richard Simon. 1975. "Sequential Treatment Assignment with Balancing for Prognostic Factors in the Controlled Clinical Trial." *Biometrics* 31 (1): 103–15. http://www.jstor.org/stable/2529712.

Proschan, Michael, Erica Brittain, and Lisa Kammerman. 2011. "Minimize the Use of Minimization with Unequal Allocation." *Biometrics* 67 (3): 1135–41. https://doi.org/10.1111/j.1541-0420.2010.01545.x.

Appendix

Comparison of mean allocations to group $T1$ for SBM1 vs SBM2

Comparing the mean allocations to group $T1$ for SBM1 and SBM2 reveals a systematic difference in results, with our implementation SBM2 tending to higher means than SBM1 at lower probabilities and to lower means at higher probabilities. This may indicate a difference in the implementation of the algorithm but we have been unable to determine this difference in correspondence with the author.

The difference is most apparent when treatment totals are not included as a balancing factor (Figures 10, 11, 12).

Figure 10. Mean allocations to group T1 for SBM1 vs SBM2, N = 30

Simulation results with varying p and factor counts

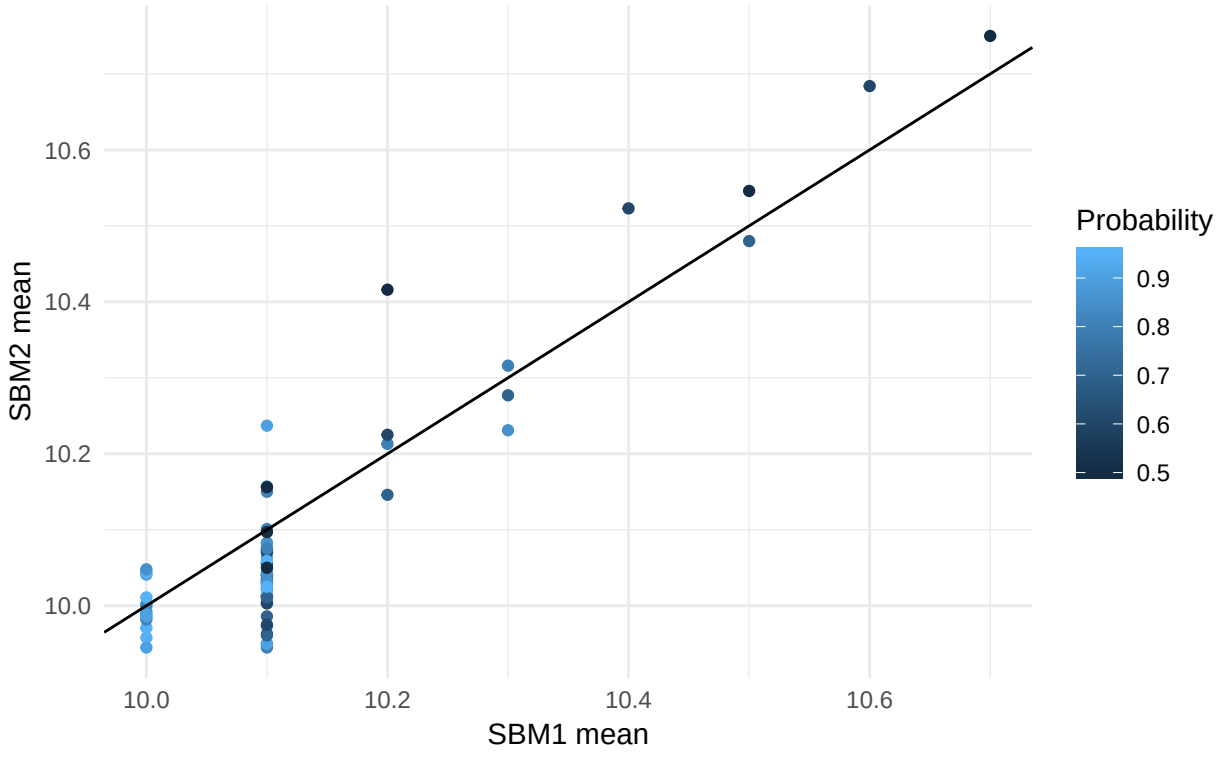


Figure 11. Mean allocations to group T1 for SBM1 vs SBM2, N = 60

Simulation results with varying p and factor counts

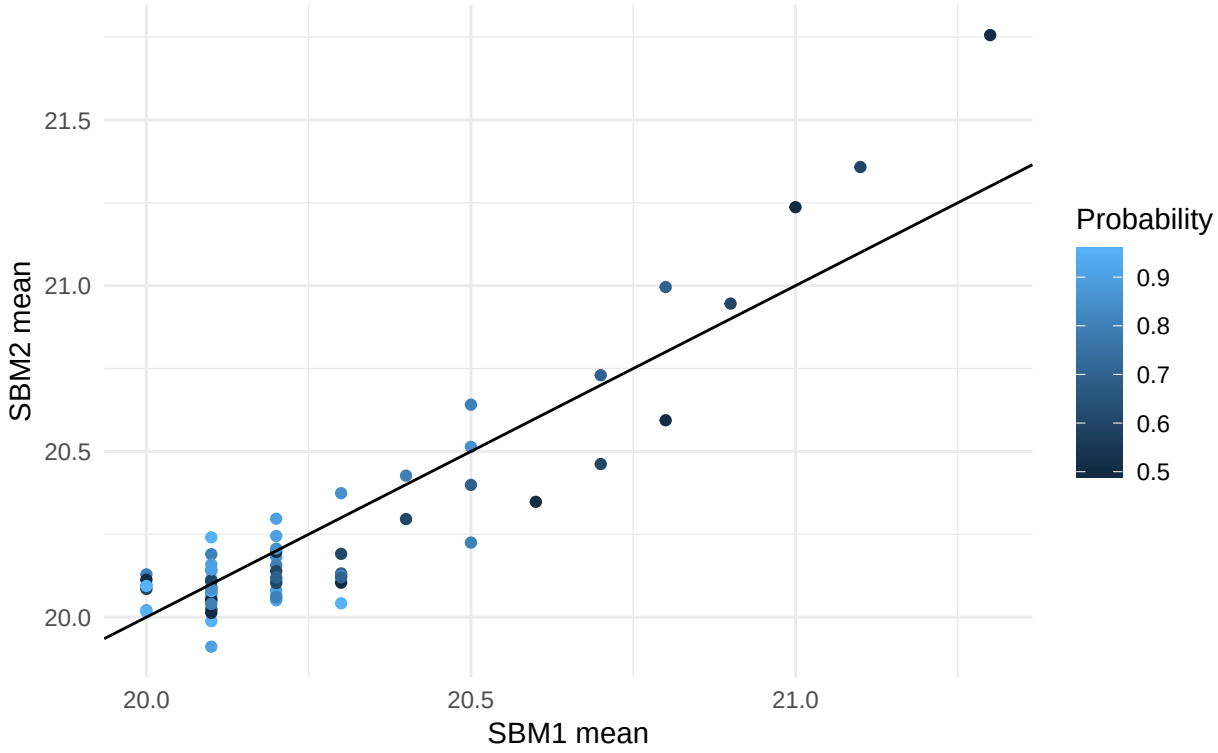
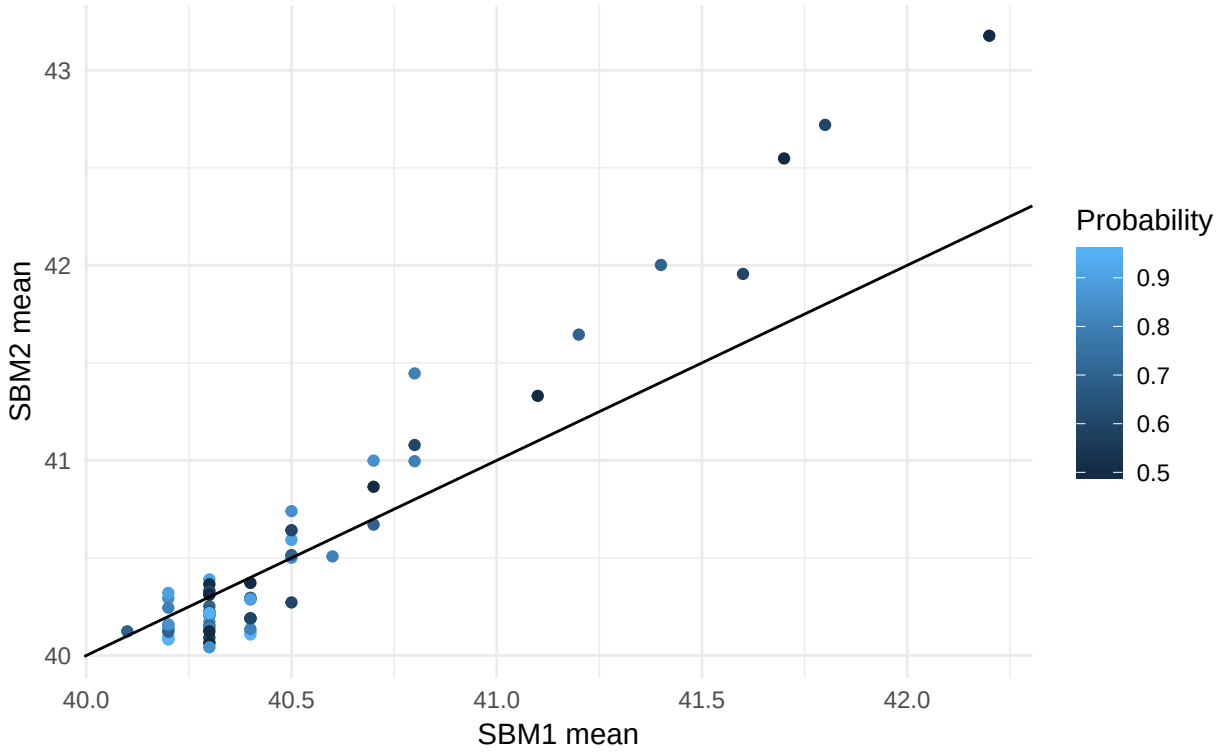


Figure 12. Mean allocations to group T1 for SBM1 vs SBM2, N = 120

Simulation results with varying p and factor counts



The difference is less obvious as treatment group totals are included (Figures 13, 14).

Figure 13. Mean allocations to group T1 for SBM1 vs SBM2, N = 120, W = 1

Simulation results with varying p and factor counts

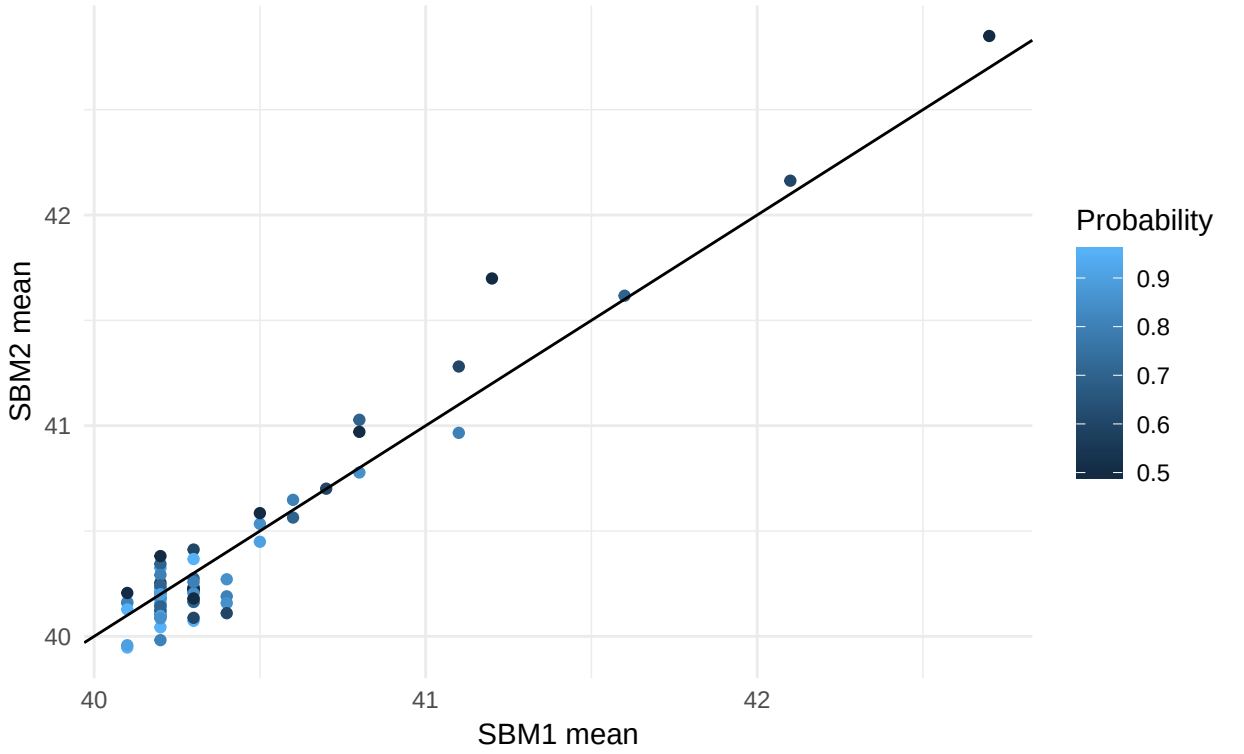


Figure 14. Mean allocations to group T1 for SBM1 vs SBM2, N = 120, W = F

Simulation results with varying p and factor counts