

## Appendix 4. Additional Matching and Rebalancing Results

### A4.1. Control Site Selection

The selected control sites generally have similar levels of bias to the wider buffer regions from which they were sampled (Fig. A4.1). A notable exception, however, is in GM1, where it was not possible to find a Control site with comparable population density and elevation characteristics to the Treatment site (Fig. A4.2).

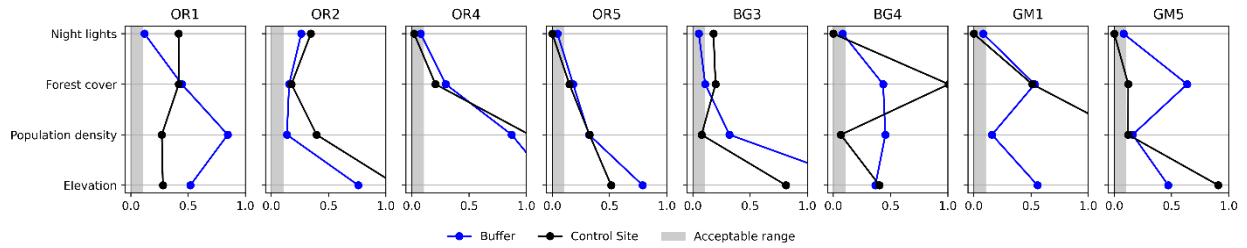


Fig. A4.1: Covariate bias for each covariate used in the first-stage Control site selection. The horizontal axes plot the absolute standardized mean difference (AMSD; Equation 1 in the main manuscript).

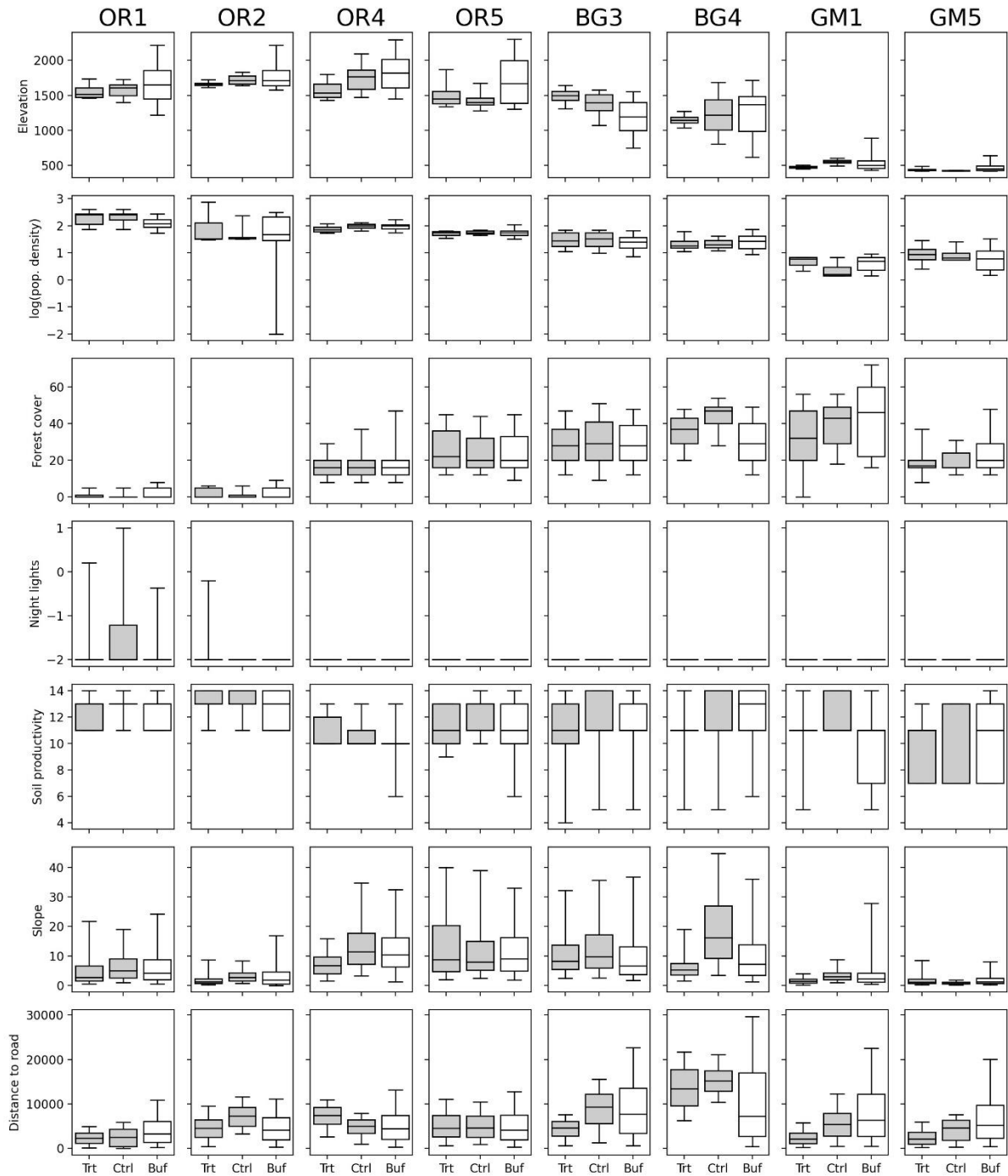


Fig. A4.2: Boxplots of the covariate distributions in each Treatment site, selected Control site, and 30km buffer from which the Control site was selected. The whiskers represent 5% and 95% percentiles of the data. Outliers are not shown. Note that the first stage only used four of these covariates (elevation, population, density, forest cover, and night lights) and the second stage used five (population density, forest cover, soil productivity, slope, and distance to road).

## A4.2. Point Sampling and Reweighting

### A4.2.1. Sample Size

The spatial correlation of the sampled points generally increased with larger sample sizes (Table A4.1). We chose to progress with sampling 75 points from each site; beyond this level, statistically significant spatial autocorrelation is observed. We believe this achieves an appropriate balance between statistical power and spatial independence.

Table A4.1: P-values from Moran's I tests of spatial correlation in the residuals of logistic regressions with dependent variables  $Y_1$  (smallholder expansion) and  $Y_2$  (smallholder abandonment) for different sample sizes. Shaded cells indicate instances in which the spatial correlation is significant at the 5% level.

Site	N=20		N=50		N=75		N=100	
	Y1	Y2	Y1	Y2	Y1	Y2	Y1	Y2
BG3	0.38	0.50	0.41	0.51	0.29	0.39	0.16	0.24
BG4 †	0.11	-	0.10	-	0.10*	-	0.10*	-
GM1 †	0.30	-	0.41	-	0.36	-	0.28	-
GM5 †	0.12	-	0.13	-	0.21	-	0.21	-
OR1	0.36	0.43	0.37	0.29	0.34	0.15	0.23	0.08*
OR2	0.52	0.41	0.28	0.48	0.16	0.47	0.03*	0.42
OR4	0.54	0.44	0.32	0.50	0.13	0.50	0.06*	0.47
OR5	0.51	0.44	0.36	0.46	0.06*	0.46	0.01*	0.42

\*p<0.1; \*\*p<0.01; \*\*\*p<0.001

† We did not assess this site's Y2 outcome because it initially contained less than 10% smallholder agriculture.

### A4.2.2 Covariate Balance

In most cases, entropy balancing successfully removed all bias in the weighted Control data (Fig. A4.3). For covariates containing high levels of bias between the Treatment and unbalanced Control site (e.g., BG4), the entropy balancing successfully reduced bias to within acceptable levels (Austin, 2009). The rebalanced data for GM5 did contain some residual bias, however the rebalancing provided a substantial improvement and the levels of residual bias are not drastically high for any covariate.

We note that the original presentation of the entropy balancing algorithm was designed to yield a weighted set of Control observations with zero bias (Hainmueller, 2012). Specifically, the optimization formulation specifies balance in first (and potentially higher order) moments as constraints. However, our data unfortunately contained limited overlap in some covariates (Fig. A4.2), meaning that the algorithm did not always converge with the original balance constraint specification. We therefore modified the source code for the *ebal* R package to exit the algorithm's iteration when instabilities were detected. With this modification, we were able to calibrate a set of weights for each site and outcome that substantially reduced the bias across all covariates (Fig. A4.3). However, for some sites (e.g., GM5 for  $Y_1$ ), these weights were highly skewed, i.e., most points received almost no weight with several points receiving large weights (Fig. A4.4). We note this issue but for the purposes of this article do not delve more deeply into modifying or developing methods for entropy balancing. Some recent research is beginning to generalize entropy balancing to achieve approximate balance (rather than exact) (Xu and Yang, 2021), but the R implementations are still in development.

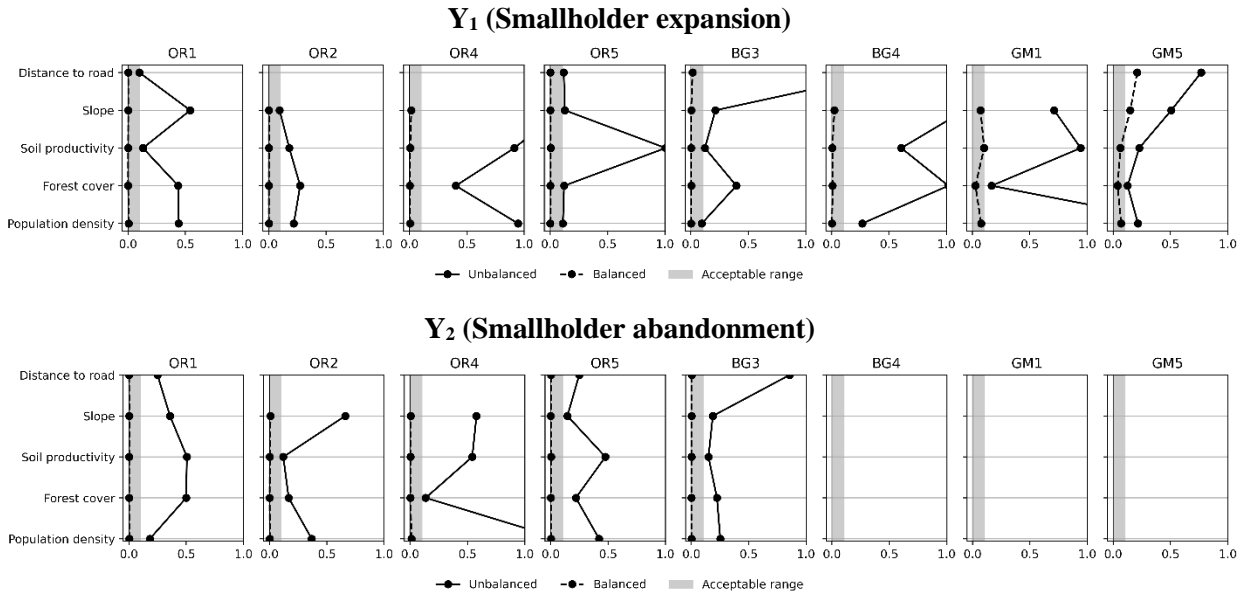


Fig. A4.3: Covariate bias for each covariate used in the second-stage rebalancing. The horizontal axes plot the absolute standardized mean difference (AMSD; Equation 1 in the main manuscript). BG4, GM1, and GM5 were excluded from the smallholder abandonment analysis ( $Y_2$ ). The distance to road covariate was excluded in four sites (BG4, GM1, OR2, and OR4) due to limited overlap in these data between the Treatment and Control sites (Fig. A4.2).

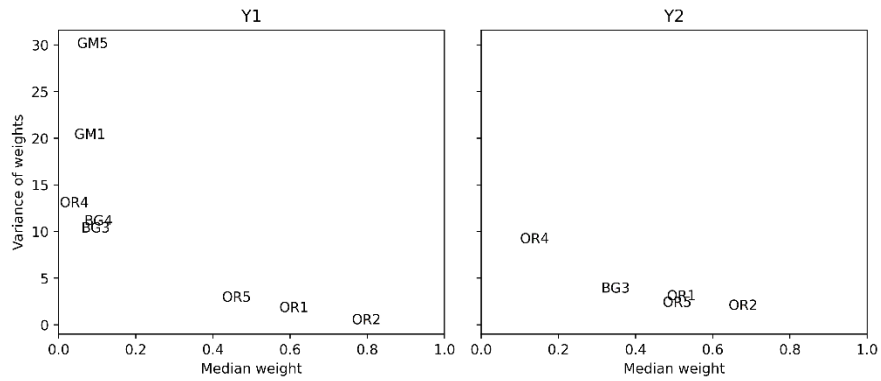


Fig. A4.4: Median and variance of resulting sample weights.

### A4.3. Robustness of Indirect LULCC Estimates

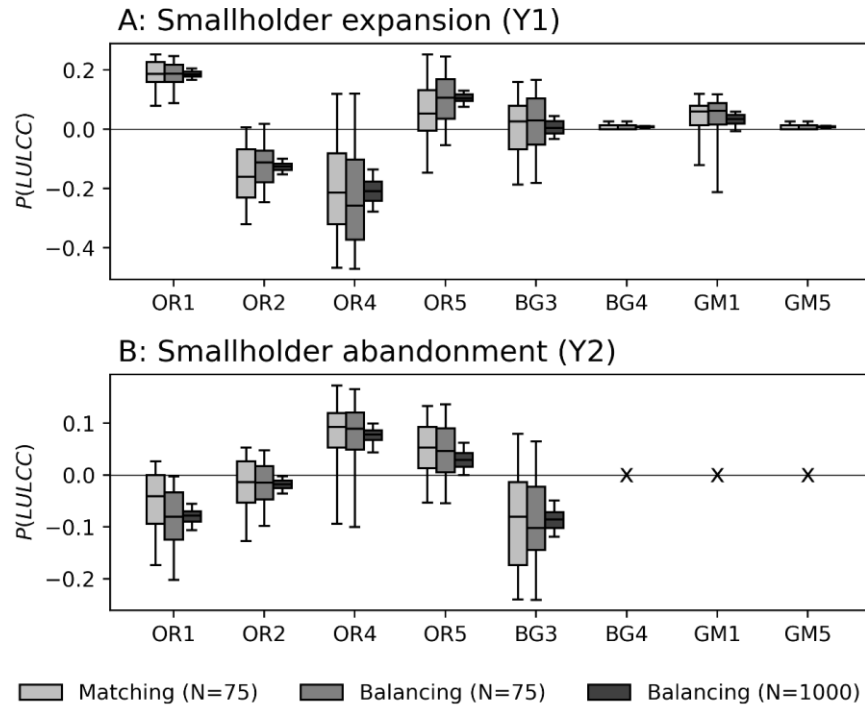


Fig. A4.5: Indirect LULCC estimates under different second-stage approaches. *Matching* ( $N=75$ ) used genetic matching with 75 points sampled from each Treatment and Control Site. *Balancing* ( $N=75$ ) used entropy balancing with 75 points and is the approach presented in the main manuscript. *Balancing* ( $N=1000$ ) used entropy balancing with 1000 points, which violate assumptions of spatial independence and consequently lead to lower variability in the estimated outcome.

### References

- Austin, P. C. 2009. “Balance Diagnostics for Comparing the Distribution of Baseline Covariates between Treatment Groups in Propensity-score Matched Samples.” *Statistics in Medicine*. <https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.3697>.
- Hainmueller, J. 2012. “Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies.” *Political Analysis: An Annual Publication of the Methodology Section of the American Political Science Association* 20 (1): 25–46.
- Xu, Y., and E. Yang. 2021. “Hierarchically Regularized Entropy Balancing.” *Available at SSRN*. <https://doi.org/10.2139/ssrn.3807620>.