# Biltong/GRASS Power Simulation — Practical Guide

This guide explains what the simulator does, what to edit, and how to read its outputs. It’s concise and geared for day-to-day use.

## What the program does

It estimates statistical power for GRASS/Meat Naturally RCTs by: - Randomizing communities to treatment arms (with stratification), - Simulating compliance (take-up), - Generating outcomes over time with clustering and AR(1) correlation, - Estimating ITT (OLS) and TOT (IV) effects with robust/cluster-robust SEs, - Repeating many times and computing power = share of significant results.

Outputs are saved to: - Tables: output/tabs - Figures: output/figs

## Files

* Biltong\_power\_simulations\_master.r — sets the scenario and runs sweeps.
* biltong\_power\_simulations\_engine.r — simulates data, runs estimators, writes outputs.
* biltong\_power\_simulations\_utils.r — helpers (AR(1), Poisson/NegBin draws, robust tests).

## What you edit in the master script

* Sample and design
  + n\_communities, avg\_ind\_obs\_per\_comm, sd\_indiv\_per\_comm
  + arms\_\* and alloc\_ratios\_\* (single vs multi-arm presets)
  + take\_up\_\* (per-arm compliance probabilities)
  + experiment\_type = "community\_level" or "individual\_within\_community"
  + Stratification vectors: year\_in\_program, ngo\_id, tribe\_id
* Outcomes and noise
  + Socio-economic counts: soc\_outcome\_dist (“negbin” | “poisson” | “none”), soc\_outcome\_theta, soc\_outcome\_ICC, soc\_outcome\_AR1\_rho, soc\_outcome\_AR1\_var
  + Environmental continuous: env\_outcome\_base\_mean, env\_outcome\_base\_sd, env\_outcome\_ICC, env\_outcome\_AR1\_rho, env\_outcome\_AR1\_var
  + Time grids: soc\_outcome\_T, soc\_outcome\_T\_months; env\_outcome\_T, env\_outcome\_T\_months
* Effects by arm (omit “control”)
  + Soc (multiplicative “rate ratio”): soc\_outcome\_ate\_pct\_\* (e.g., c(T1 = 1.20))
  + Env (additive): env\_outcome\_ate\_\* (e.g., c(T1 = 10))
* Inference and repetitions
  + sims, alpha, cluster\_se, hc\_type, cluster\_fe\_yn

## How one dataset is generated

### Randomization (with stratification)

Communities are assigned to arms using alloc\_ratios, balanced within strata (year\_in\_program, ngo\_id, tribe\_id).

### Compliance (take-up)

Realized treatment is a Bernoulli draw per observation with probability equal to the arm’s take\_up. ITT uses assignment. TOT instruments realized treatment with assignment.

### Outcome models

Socio-economic (counts; Negative Binomial or Poisson)

* Log-mean model for community c at time t:
* is the cluster random intercept from ICC; is AR(1): .
* is the treatment arm’s multiplicative rate ratio; adds stratum-specific TE when treated.
* Draws:
  + Negative binomial if soc\_outcome\_dist == "negbin":
  + Poisson if "poisson":
  + Deterministic if "none":

Environmental (continuous)

* Same ICC/AR(1) structure. is the additive arm effect.

ICC and AR(1)

* Cluster intercept variance:
* AR(1): with innovation variance \*\_AR1\_var.

## Estimation per dataset

* ITT (OLS):
* TOT (IV, per arm using only control + that arm): first stage , second stage .
* SEs: cluster-robust when cluster\_se=TRUE, else HC-type (hc\_type).

## Power and sweeps

Repeat the above sims times. For each arm and outcome:

* Power(ITT) = share of -values < alpha in the ITT model.
* Power(TOT) = share of -values < alpha in the IV model.

Typical sweeps vary effect size (rate ratio or additive effect) and plot power vs. effect; you can also sweep ICC, AR(1), T, or sample size.

## Outputs you’ll see

* Tables (wide and long) in output/tabs, including the last simulated dataset for a quick realism check.
* Figures in output/figs with captions listing arms, allocations, number/timing of observations, take-up, distributions, and inference settings.

## Quick checks

* Negative binomial in use: set soc\_outcome\_dist = "negbin" and confirm in the caption; the last simulated socio-economic data should have variance > mean.
* Avoid community FE in ITT for pure community-level RCTs (assignment is time-invariant within community).
* Watch cores when parallelizing.

## Minimal math summary

* AR(1):
* Cluster intercept variance:
* Socio-economic mean:
* Environmental level:
* Power: