Appendix H
Statistical Models for Data Analysis

Aim 2d. Generalize Linear Mixed Effect Model (GLMM)

To model the sustainability outcome of the Eban II as functions of characteristics of the CBOs, the GLMM can be written in general as: $Y_i = X_i\alpha + Z_i\gamma_i + \varepsilon_i$, for i=1,2,...,$n_i$,

where Y_i is the $n_i \times 1$ vector of log odds of probability of being effective with sustainability (e.g., continue to implement Eban II with fidelity or has integrated the program into existing services) for the i^{th} agency, n_i is the number of measurements for the i^{th} agency, X_i is $n_i \times p$ known predictors (e.g., number of implemented groups with fidelity and high retention and maintaining high facilitator competence). The fixed effects parameter vector of α is $p \times 1$. Z_i is $n_i \times k$ known predictors, which are associated with the random part of the model (e.g., change within an agency). γ_i is the random effect parameter vector which is $k \times 1$ and $r_i \sim N(0, D_i)$, ε_i is $n_i \times 1$ and $e_i \sim N(0, \sigma^2 I)$ which is independent of γ_i. Let $\Sigma_i(\theta) = Z_i D Z_i{}^T + \sigma^2 I$ with θ being the parameter of covariance vector. Integrating out γ_i in the above random effect model, we can write $Y_i \sim N(X_i\alpha, \Sigma_i(\theta))$.

Aim 3 Hierarchical Models

To model sustainability, the hierarchical model can be written as:

Step a: individual-couple-level models model: $y_{ijk} = \beta_{0jk} + \beta_{1jk} X_{ijk} + \varepsilon_{ijk}$

where $i = 1, 2..., n_i$, $j = 1, 2..., n_j$, $k = 1, 2..., n_k$. y_{ijk} is the outcomes (e.g., condoms use with intercourse) of i^{th} individual-couple, j^{th} couple, interviewed by k^{th} agency organization. n_i is the number of measurements for the i^{th} individual-couple (e.g., at baseline, post, and 3-month follow-up), X_{ijk} is $n_i \times p \times q$ known predictors (such as time), $\varepsilon_{ijk} \sim iid\ N(0, \sigma_p^2 I)$. The effective observations of modeling at this level will be 1800 (=180x10x3). This individual-couple-level model sets-up the foundation for the class of hierarchical models.

Step b: across couples (180 couples) level model to model the regression coefficients β_{0jk} and β_{1jk}

$$\beta_{0jk} = \gamma_{00k} + \gamma_{01k} Z_{0jk} + \omega_{0jk}$$
$$\beta_{1jk} = \gamma_{10k} + \gamma_{11k} Z_{1jk} + \omega_{1jk}$$

where Z_{0jk} and Z_{1jk} are known couples characteristics (such as demographics; health history; history of alcohol and recreational drug use; HIV/AIDS risk reduction knowledge; perceived couple sexual norms; intervention components and a binary variable indicating immediate active treatment or wait list control). $\omega_{0jk} \sim iid\ N(0, \sigma_{0ph}^2 I)$, $\omega_{0jk} \sim iid\ N(0, \sigma_{1ph}^2 I)$

Step c: across agency (10 agencies) level model to model the regression coefficients γ_{00k}, γ_{01k}, γ_{10k} and γ_{11k}

$\gamma_{00k} = \lambda_{000} + \lambda_{01k} P_{00k} + \rho_{00k}$ where $\rho_{00k} \sim iid\ N(0, \sigma_{00site}^2 I)$

$\gamma_{01k} = \lambda_{010} + \lambda_{01k} P_{01k} + \rho_{01k}$ where $\rho_{01k} \sim iid\ N(0, \sigma_{01site}^2 I)$

$\gamma_{10k} = \phi_{100} + \phi_{11k} Q_{10k} + \psi_{10k}$ where $\psi_{10k} \sim iid\ N(0, \psi_{10site}^2 I)$

$\gamma_{11k} = \phi_{110} + \phi_{11k} Q_{11k} + \psi_{11k}$ where $\psi_{11k} \sim iid\ N(0, \psi_{11site}^2 I)$

where $P_{00k}, P_{01k}, Q_{10k}$ and Q_{11k} are known characteristics of agency level organizations (such as organization size and budget; number of staff;). Combine (a), (b) and (c), we have the following integrated model, which encompasses individual, couple, and agency level characteristics and will be estimated through special procedures and software:[118,119]

$$y_{ijk} = [(\lambda_{000} + \lambda_{01k} P_{00k} + \rho_{00k}) + (\lambda_{010} + \lambda_{01k} P_{01k} + \rho_{01k}) Z_{0jk} + \omega_{0jk}]$$
$$+[(\phi_{100} + \phi_{11k} Q_{10k} + \psi_{10k}) + (\phi_{110} + \phi_{11k} Q_{11k} + \psi_{11k}) Z_{1jk} + \omega_{1jk}] X_{ijk} + \varepsilon_{ijk}$$