Optional Step 5: Estimating multivariate GWAS with userGWAS

userGWAS takes 3 required arguments:

1. **covstruc**: The output from LDSC.

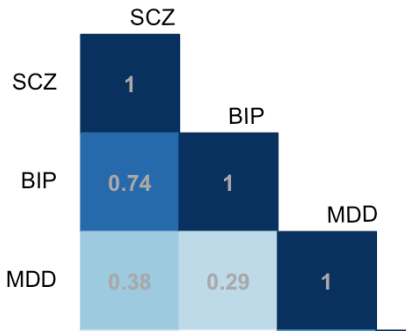
2. **SNPs**: The output from sumstats.

3. **model**: The model that is being estimated (written in lavaan syntax)

Behind the scenes

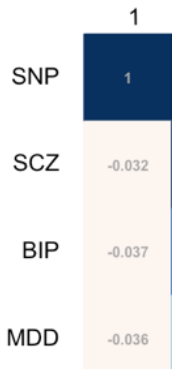
- *userGWAS* combines output from *ldsc* and *sumstats* to be able to specify a model with SNP effects
- Creates as many covariance matrices as there are SNPs across traits

Run ***ldsc***



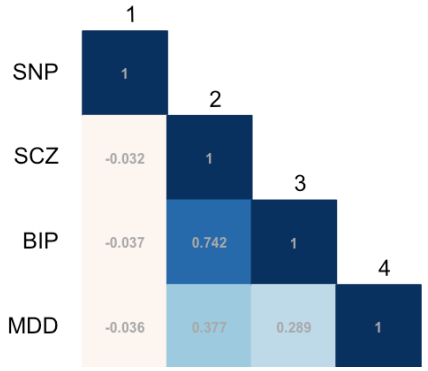
+

Run ***sumstats***



=

userGWAS combines the two



Frequently used optional arguments

- **sub**: Recommend argument that only specific components of the model output should be saved.

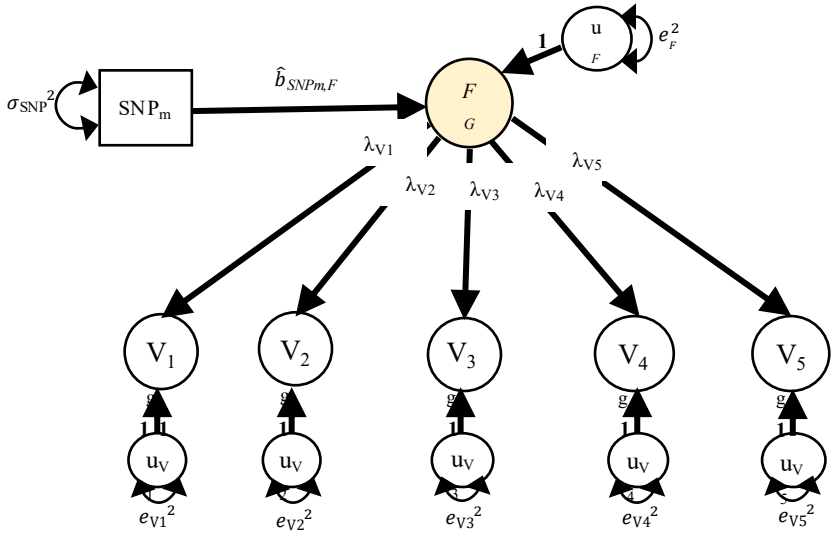
- **std.lv**: Whether you want the latent variables to have variances of 1.

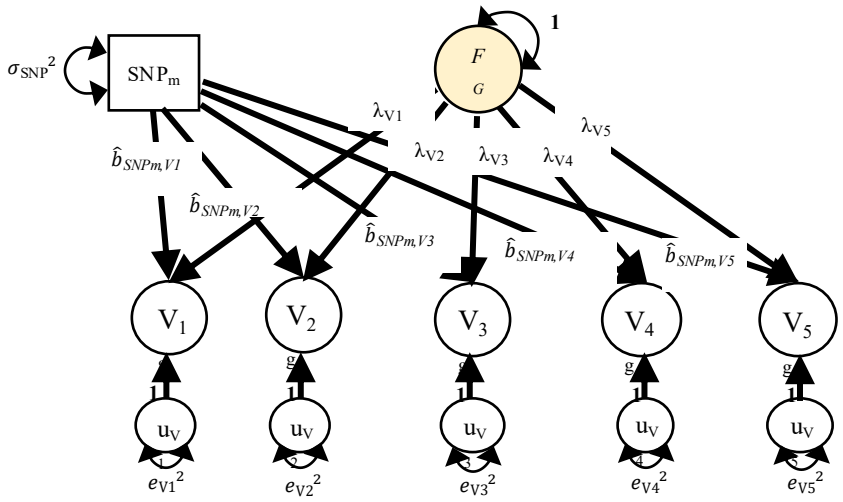
- **parallel**: Recommended optional argument specifying whether you want the function to be run in parallel, or to be run serially.
  - **cores**: How many cores to use (default = 1 less than available)
  - **Note each run is independent of the other, so as with standard univariate GWAS, the sumstats can be arbitrarily split to run across multiple jobs on a computing cluster*

- **Q_SNP**: Whether the function should automatically calculate Q_SNP for each latent factor.

Estimates of SNP level heterogeneity (Q_{SNP})

- Asks to what extent the effect of the SNP operates through the common factor
- χ^2 distributed test statistic, indexing fit of the common pathways model against independent pathways model



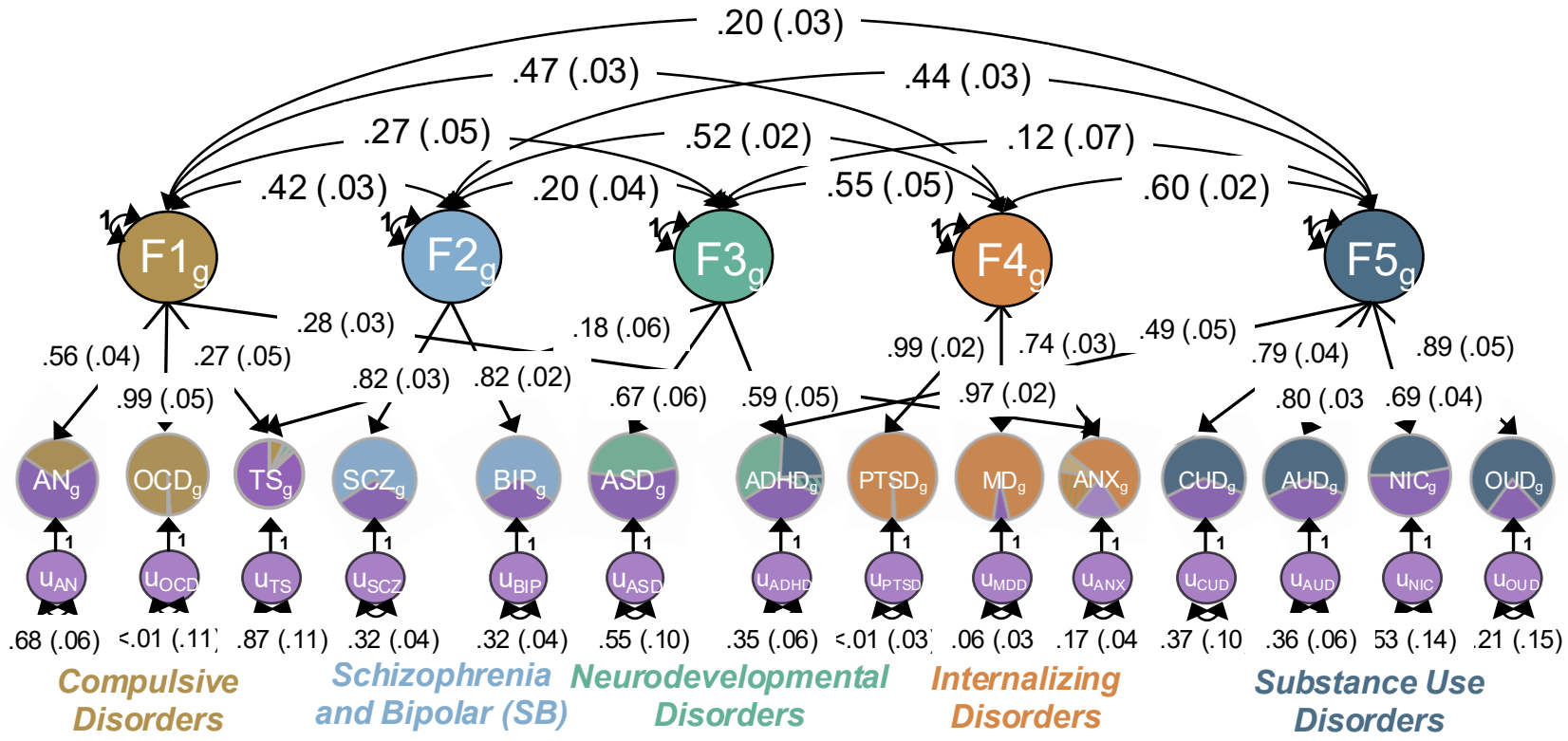


Q_{SNP} : QC Metric and a Result

- For some projects, you might just use Q_{SNP} to prune your factor GWAS results for SNPs that are unlikely to operate via the factor

- In other instances, your traits might be so highly correlated, but currently considered distinct phenotypes, such that your research question includes investigating what SNPs differentiate your traits

If you have multiple factors, you can get a Q_{SNP} specific metric for each factor. For example, you could have a SNP that fits one factor well, but is very disorder-specific for another factor



Grotzinger et al., 2025

Q_{SNP} : Better, Faster, Stronger

- It used to be that you had to estimate a separate independent pathways follow-up model.
  - This was slow, error prone, not computationally efficient

- Q_{SNP} is now an argument for *userGWAS* that will automatically calculate this metric for every factor that is predicted by a SNP without estimating a follow-up model