Lab: the final models

Add a gwesp term to the faux.mesa.high model

And conduct model assessments

We will compare three models

Model	Network Statistics g(y)
Edges + nodal attributes + mixing by attributes + degree(0)	# of edges # edges for each grade and race group # edges that are within-race & within-grade (DH) # Isolates
Edges + Attributes + GWESP(0.25)	# of edges # edges for each grade and race group # edges that are within-race & within-grade (DH) weighted shared partners, with decay set to 0.25
Edges + Attributes + GWESP(0.5)	# of edges # edges for each grade and race group # edges that are within-race & within-grade (DH) weighted shared partners, with decay set to 0.5

These fits can take a while

- So we won't do this interactively now
  - We'll just show the results

- But you can implement these on your own when you have some time

Model Comparison

Current Model Summary   Current Model Fit Report   Model Comparison

```
                          Model1      Model2      Model3
edges                     -8.186***   -8.522***   -8.551***
nodefactor.Grade.8         1.633*      1.396*      1.428*
nodefactor.Grade.9         2.670***    2.178***    2.213***
nodefactor.Grade.10        3.053***    2.475***    2.516***
nodefactor.Grade.11        2.712***    2.247***    2.285***
nodefactor.Grade.12        3.501***    2.871***    2.910***
nodefactor.Race.Hisp      -1.396***   -1.114***   -1.092***
nodefactor.Race.NatAm     -1.336***   -1.087***   -1.093***
nodefactor.Race.Other     -2.165*     -2.074*     -2.134*
nodefactor.Race.White     -0.725*     -0.588*     -0.604*
nodematch.Grade.7          7.469***    5.975***    6.006***
nodematch.Grade.8          4.292***    3.237***    3.215***
nodematch.Grade.9          2.060***    1.613***    1.626***
nodematch.Grade.10         1.281*      1.052*      1.059*
nodematch.Grade.11         2.495***    1.831***    1.789***
nodematch.Grade.12         1.361.      0.971       0.911
nodematch.Race.Black        -Inf***     -Inf***     -Inf***
nodematch.Race.Hisp        0.678.      0.569.      0.555.
nodematch.Race.NatAm       1.272***    1.053***    1.058***
nodematch.Race.Other        -Inf***     -Inf***     -Inf***
nodematch.Race.White       0.340       0.315       0.342
degree0                    1.305***   NA          NA
gwesp.fixed.0.25          NA           1.398***   NA
gwesp.fixed.0.5           NA          NA           1.257***
AIC                        1806        1664        1659
BIC                        1965        1823        1818
```
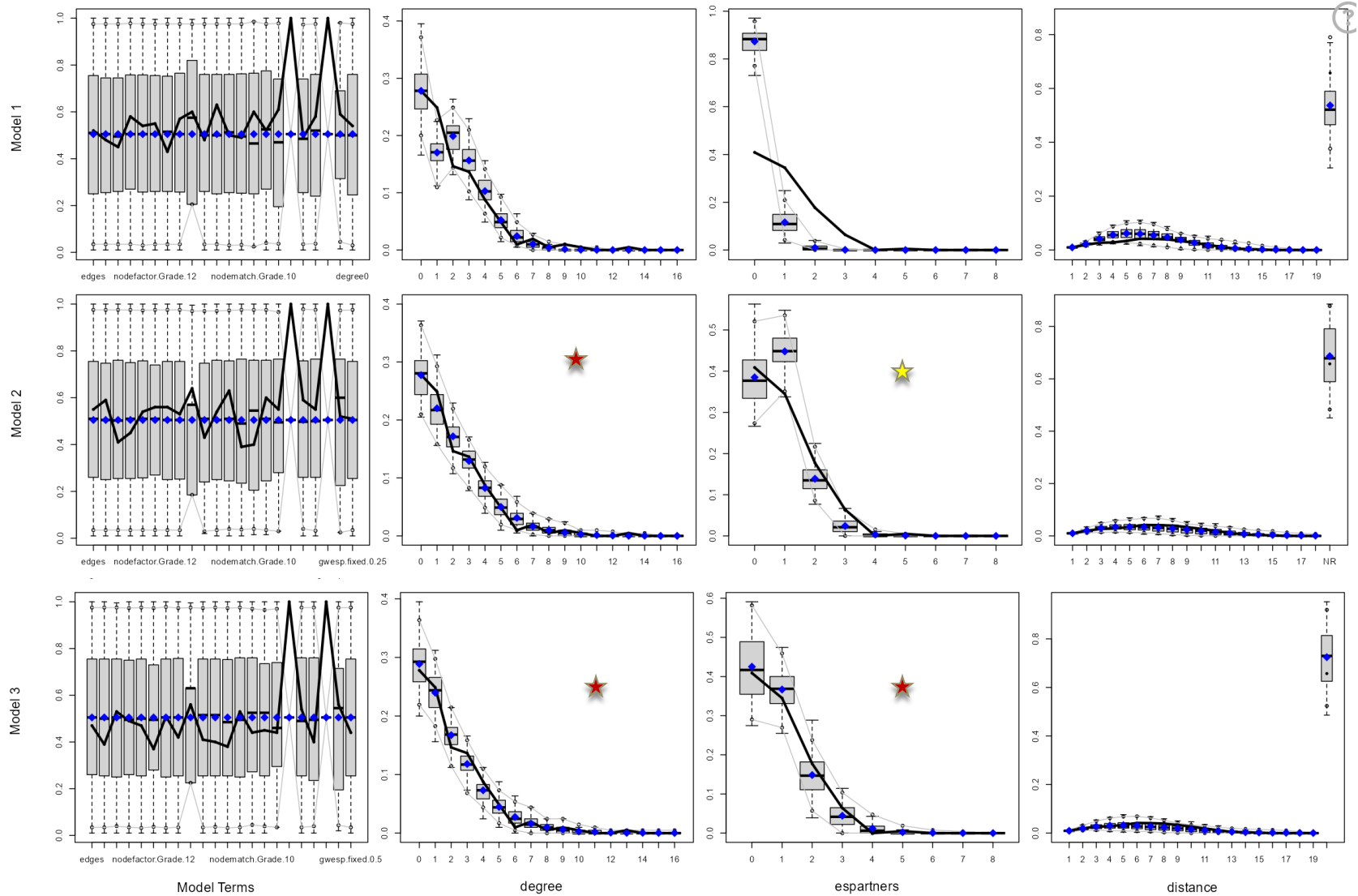
- The gwesp term is clearly significant.

- And note how the homophily coefs changes from model 1 after the gwesp is added to models 2 & 3
  - About 10-20% smaller
  - That's the impact of controlling for triadic closure effects

- Some weak evidence here that the 0.5 decay is a better fit for the gwesp term

GOF comparison for all 3 models:

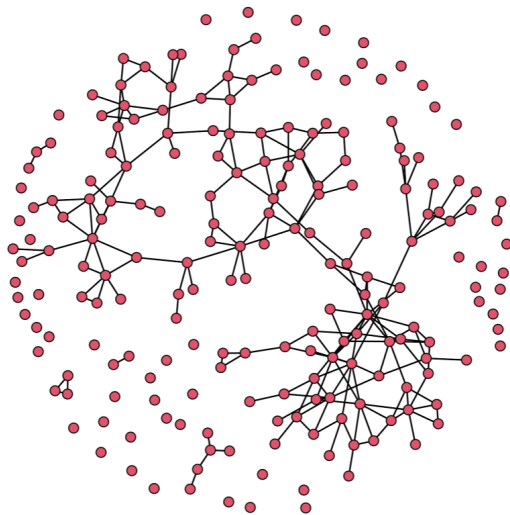Goodness-of-fit diagnostics

# Simulating networks from the model

- A fitted model describes a probability distribution across all networks of this size
  - The model assigns a probability to every possible network
  - The model terms and the estimated coefficients make some networks more likely than others

- You can simulate networks from this distribution
  - Using the same MCMC algorithm that was used for estimation and GOF

- And the simulated networks will be centered on the network statistics in the original observed network
  - This is why these models are really useful for network epidemiology
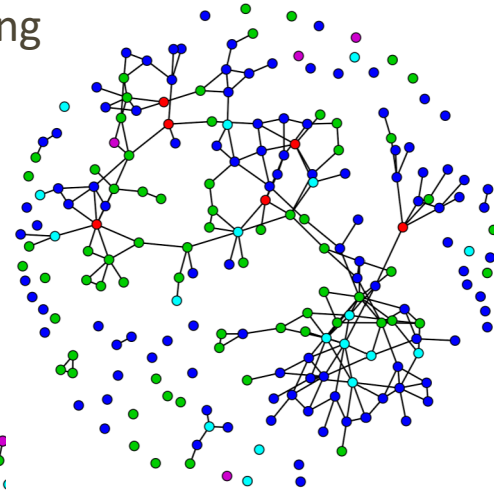
Simulation (finally!) from Model 3

So does mixing
by race

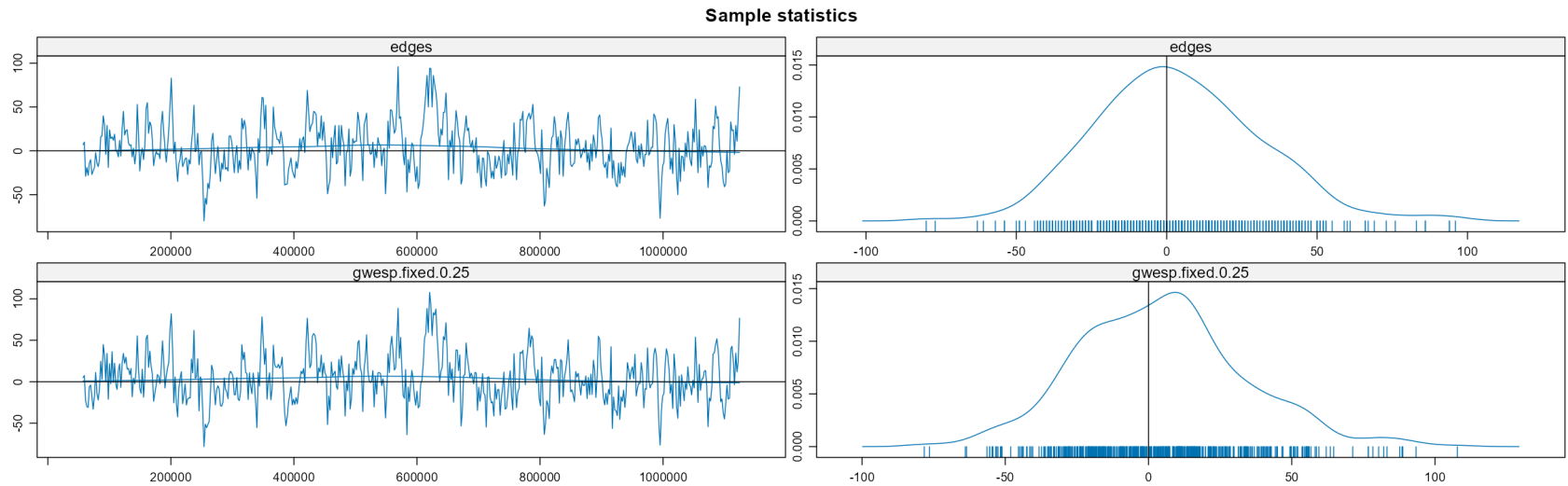Grade mixing
looks good

Overall structure
looks good
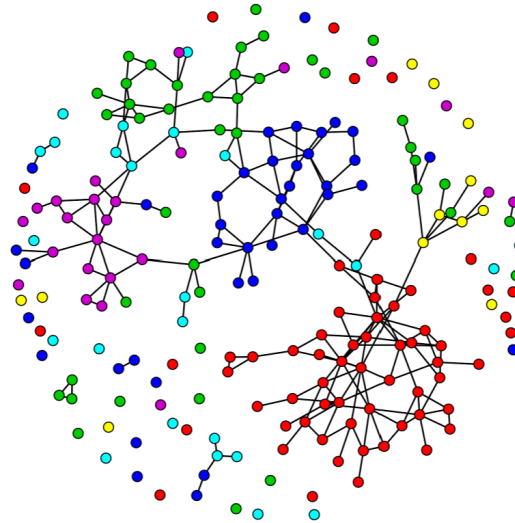
Race
- Black
- Hisp
- NatAm
- White
- Other

Grade
- 7
- 8
- 9
- 10
- 11
- 12

Now I'm curious

- What happens if we fit just edges + gwesp?

- So I tried it
  - gwesp(0.5) triggered a degeneracy stop
  - so did gwesp(0)
  - gwesp(0.25) did return a fit …
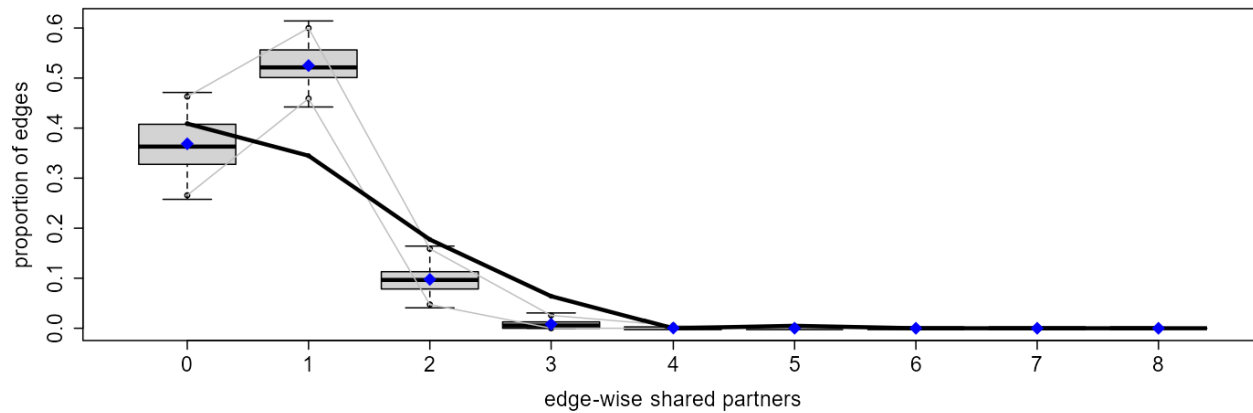
MCMC dx for edges + gwesp(0.25)

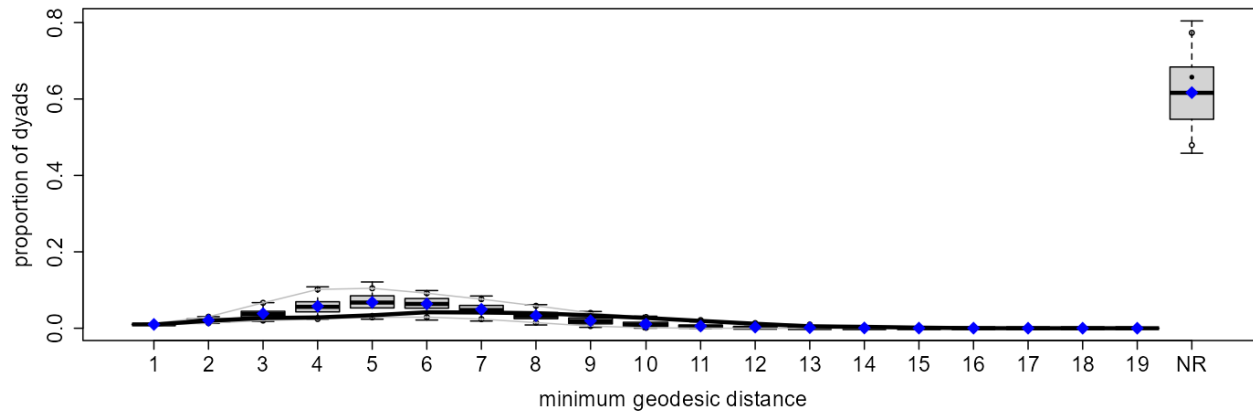The traceplots show moderate autocorrelation

The distributions look ok tho

Could try upping the MCMC.interval control parameter

GOF for edges + gwesp(0.25)

Better than Model 1, about the same as Model 2

But the fit to the geodesics is poor, especially near the mode

Summary

- Now we can say something about this network

- Both transitivity and homophily clearly play a role in clustering these friendships
  - Homophily
    - Also reproduces the geodesic distribution
    - But not the degree distribution of the local shared partner clustering
  - Transitivity (Triadic closure)
    - Reproduces the degree distribution and captures the local clustering (ESP) well
    - But not the geodesic distribution

- The model with both does best
  - And simulations from this model look remarkably similar to the observed network

This is what makes EpiModel so powerful

Believable network simulations, based on:

- Robust, principled statistical methodology for estimation and inference with a fully general modeling framework (ergms/tergms)

- Simulations deeply rooted in empirical network data that reproduce observed network statistics (in and out of the model)

- And simple data collection requirements (egocentric samples)

All of this is also embedded in a fully general stochastic epidemic modeling package