Problem Set #1
MACSS 30100, Winter 2018
Nora Nickels

1. Classify a model from a journal.
(a) I am a psychology and biology PhD student in the department of Comparative Human Development, and my work is mainly experimental biopsychology. As a psychology student, I chose to search for a model from the most recent issue of *Psychological Science*. I work with social hormones, such as cortisol and testosterone, so I've selected a paper that models with baseline hormonal levels. The statistical model I've selected comes from testing the main effects of urinary oxytocin on the subjective experience of loving feelings, with oxytocin as the primary predictor of interest, and condition (experimental condition of expressing gratitude or not) and gender as control variables. It comes from a paper titled: Oxytocin and Social Bonds: The Role of Oxytocin in Perceptions of Romantic Partners' Bonding Behavior.

(b) Citation: Algoe, S. B., Kurtz, L. E., & Grewen, K. (2017). Oxytocin and Social Bonds: The Role of Oxytocin in Perceptions of Romantic Partners' Bonding Behavior. *Psychological science*, *28*(12), 1763-1772.

(c)

$$\Upsilon_{love} = \alpha + \beta_{gender} + \beta_{condition} + \beta_{oxytocin} + \varepsilon$$

with Υ_{love} being the dependent outcome variable of interest (the subjective ranking of experience of love from partner by the participant);
α = *y*-intercept, constant
β_{gender} = participant gender
$\beta_{condition}$ = experimental condition, with control participants not going through the subjective gratitude condition with their partner (assigned before experimental session)
$\beta_{oxytocin}$ = predictor variable of baseline hormonal oxytocin
ε = error term

(d) Υ_{love}: endogenous variable, or output of the model
β_{gender} = exogenous variable, taken as given from outside of the model
$\beta_{condition}$ = exogenous variable, taken as given from outside of the model
$\beta_{oxytocin}$ = exogenous variable, taken as given from outside of the model

(e) The model is static, as it does not take time into account, and is not differential in nature (e.g., this model does not look at oxytocin at various time points, but just takes oxytocin at baseline). The model is linear, as all of the parameter and predictor combinations are linear in nature. Finally, the model is stochastic, as the error term is included to define inherent randomness within the model.

(f) I believe that observing the reactivity of oxytocin should be included within the model. Using solely baseline oxytocin does not account for how participants' oxytocin levels react to subject behavioral stimuli of expressed and perceived gratitude, and perhaps the biological reactivity of each participant would be a further predictive variable or mediating variable within the model.

2. Make your own model.
(a)

$$Y_{mar} = \beta_1 rel + \beta_2 leng + \beta_3 age + \beta_4 religion + \beta_5 agediff + \beta_6 divorce + \beta_7 income + \beta_8 partinc + \beta_9 children + \beta_{10} big5 + \beta_{11} gender + \beta_{12} gender * age +$$

Where Y_{mar} is the endogenous variable of whether or not the individual gets married (1 = get married and 0 = not get married); rel is relationship status currently (0 or 1 variable); length is relationship length in months of current relationship; age is the age of the individual (in years); religion is a categorical religious affiliation of the individual; divorce is number of times individual has been previously married; income is the income of the individual; partinc is the partner's income; children is the desire to have children (on a subjective survey Likert scale of 1 – 10, with 1 being no desire and 10 being intent on having children); big5 is a personality score based on the big five inventory that measures aspects of personality such as extroversion, neuroticism, etc.; gender is the gender of the individual (0 or 1 variable); and finally, gender*age is an interaction term looking at gender and age of the individual.

(d) I believe that the key factors in my model are relationship length, religion, big5, and gender*age interaction.

(e) I believe that relationship length is a key factor because it encompasses whether or not someone is in a current relationship while also capturing how serious the relationship is in terms of length. Marriage is certainly a more serious consideration when an individual has been in a relationship for a long time, but I do believe that the other key factors I have chosen help to explain the variation in individuals who either have a shorter relationship length and get married or individuals who have a longer relationship length and do not get married. For example, religion as a categorical factor would certainly play a role in this. Because of cultural and social beliefs, religion often plays into why younger individuals would get married more quickly. As time progresses and traditional family dynamics and faiths phase out, individuals who are less religious may not choose to get married, which is why I believe religion is such a key factor. Big5, or personality, is also a key factor I've selected. I believe this is especially important based on the factors the Big 5 personality questionnaire covers – introversion/extroversion, agreeableness, neuroticism, openness and conscientiousness. Finally, I believe the gender*age interaction is important, as statistically it has been shown that women tend to marry younger than men do.

(f) I would test this longitudinally with a study. I would collect data from a random sampling on Mechanical Turk in three separate waves. The first wave would require individuals to not be married, the second wave would be two years later and would be open to the same individuals whether or not they had married, and the third would be two years after that and would also be open to the same individuals whether or not they had married. I would test the data collected on all these factors in multiple ways. I would run a model using just the four key factors I listed above, then run a model with all of the factors in my full model. I would compare the R squared values and F values from both these models, paying attention to any factors from the full model that I did not list as key factors that significantly predicted marriage status. This would be my preliminary test using real life data.