Problem Set #1
MACS 30000, Dr. Evans
Cooper Nederhood

Problem 1: classify a model from a journal

Part (a): *The Surprisingly Swift Decline of US Manufacturing*

Part (b): Pierce, Justin, and Peter Schott. 2016. "The Surprisingly Swift Decline of US Manufacturing." *American Economic Review* 106(7): 1632-1662

Part (c): The following statistical model evaluates the relationship between PNTR (permanent normal trade relations) status for China and US manufacturing employment using a generalized OLS difference-in-difference specification. Specifically, it tests whether industries with a higher NTR (normal trade relations) gap have a larger employment change after the imposition of PNTR

$$ln(Emp_{i,t}) = \theta PostPNR_t \times NTRGap_i + PostPNR_t \times \mathbf{X}_i \gamma + \mathbf{X}_{i,t}\lambda + \delta_t + \delta_i + \alpha + \epsilon_{i,t} \quad (1)$$

Part (d): The endogenous variable is $ln(Emp_{i,t})$, the natural log of the manufacturing employment in industry i at time t.

The exogenous variables include: $PostPNR_t$, a dummy variable indicating whether the year is past the policy change; $NTRGap_i$, the difference between the normal trade relations tariff price and the alternative; $\mathbf{X}_i$, time-invariant industry-specific control variables; and $\mathbf{X}_{i,t}$, time-variant industry-specific control variables like union membership.

The variables δ_i and δ_t are the industry and time fixed-effects, respectively. And $\epsilon_{i,t}$ is the error term.

Part (e): The model is dynamic rather than static because it uses panel data and, through the interaction terms with $PostPNR_t$, the model allows for relationships to change between the pre and post period.

The model is linear because each parameter and each exogenous variable is linear.

The model is stochastic because of the random error term $\epsilon_{i,t}$.

Part (f): Because the model uses a difference-in-difference framework it detects the *relative* differences in employment changes pre and post policy change but it does not provide estimates on the *absolute* change in employment due to the policy change.

Problem 2: Make your own model

Part (a): The following equation describes a statistical model of whether an individual gets married. Note, for categorical variables like *religion* that would require multiple flag variables, I simply list the variable rather than each individual flag, for clarity.

$$married_i = \alpha + \beta_1 height_i + \beta_2 weight_i + \beta_3 sexuality_i + \beta_4 trans_i + \beta_5 countryOrigin_i + \beta_6 religion_i + \beta_7 income_i + \beta_8 IQ_i + \beta_9 education_i + \beta_{10} psychDisorder_i + \beta_{11} race_i + \beta_{12} dateOfBirth_i + \epsilon_i \quad (2)$$

Part (b): Here the dependent variable, $married_i$, is a 1/0 flag denoting whether the individual is married

Part (c): The right-hand-side of the equation includes all exogenous variables. Assuming we have data on these demographic variables we could fully simulate the data generating process.

Part (d): One could view the marriage process from many angles, but here I am viewing marriage as a matching problem. An individual evaluates each possible mate with regards to compatibility and if a compatible mate is found, a match (a marriage) occurs. This evaluation process is a function of the individual's observable characteristics. For example, perhaps an exceptionaly intelligent person finds it difficult to locate individuals with similar intelligence levels and thus her prospective mate pool is smaller, resulting in a more difficult search. The exact relationship between the observable characteristics and the marriage decision is left up to the model to discern.

Part (e): A marriage decision is an infinitely complex problem. Many of the factors involved, namely love, are impossible to measure or to even arrive at some agreed-upon concrete definition. Thus, it could be informative to remove all notions of humanity and romance from the process and simply view it as finding two shapes that mutually fit together. Even a non-result, in which the exogenous variables have no bearing on mate selection, would be an interesting result as it would suggest that humans *are not* simply cold rational agents searching for mates.

Part (f): Because I have focused on concrete demographic variables we can easy test this using a cross-sectional population survey. We could run a logit regression to test for statistical and economics significance in the β estimates in (2). Also, we could split the data into training and testing data to test the predictive power of the statistical model. Because the model is a simple cross-section while the marriage decision is sustained over a lifetime perhaps the timing effect could confound the results, but it would be a good first step before a more complex model is developed.