Paper components (and suggestion for 20-page breakdown):

[*Note: This is a GUIDE – you don't have to include every single component listed on this outline in your paper. You can also move things around as you see fit.*]

I. Introduction (1-2 pages)
- Overview of:
  - Question/why care about question?/why important
  - Identification Strategy
  - Results
  - Conclusion/policy implication

II. Background/ Literature review (2 pages)
- Give motivation: Why is this question important?
- Discuss key articles (5-6) only
- What are the shortcomings of these studies or what specific question is left unanswered (or not convincingly answered)? External validity problems?
- Why is it important to learn more about this (what's left to be done; what's contribution of your paper to topic)?

III. Setting (1 page)
- Why did you choose this setting?
- Describe relevant features of the setting key to interpretation
- What makes question appropriate to this population?

IV. Data description (1 page)
- What data set do you use (name), and why did you choose these data? (If it's the only choice, say so!)
- Basic features of data:
  - Cross-section or panel?
  - Years; ages, gender, etc of sample
  - Description of sampling technique
  - Overview of content
  - Obvious shortcomings (i.e. no income data; no men interviewed, only people attending school interviewed, etc.)

V. Identification Strategy (2 pages)
- What are biases of a "naïve" estimate (just regressing y on x, or using a full sample from any time or place)?

- Chosen naïve estimate depends on your identification strategy. Examples:
- Pooling different types; ignoring a key regressor; excluding the interaction term you're using to identify effect; using cross-section instead of changes; looking across rather than within families, looking at wrong outcome, excluding fixed effects, etc.)
- Describe comparison groups, give intuition/justification for this comparison
- Present findings from Table 1, discuss anything that stands out in comparison of means. You don't have to discuss every single mean in the table.

VI.   Econometric specification (1 page)
- Describe precise sub-sample used
- Describe outcome variable (including potential choices)
- Specify estimating equation in Table 2
- Give list of control variables (be precise about anything how anything is measured). Why is it necessary to include these?

VII.   Results: 2 page
- Interpret coefficients on key variables only presented in table 2
- Note both statistical significance & point estimates, interpret magnitude of
- estimated effect
- Note whether difference between naïve and alternative estimates bigger or smaller
- Are any other coefficients strange/large/unexpected?

VIII. Measurement issues: 2-3 pages
- Discuss any potential remaining biases/shortcomings of table 2 results
- Suggest ways to check these (i.e. describe columns in table 3)
- Present and interpret results from Table 3

IX.   Discussion: 2 pages
- What are the ambiguities of your results?
- What are different possible interpretations?
- What are strongest arguments for one or the other?
- What directions for future work suggested by your results?

X.   Conclusion/policy implications of findings: 1-2 pages
- "Policy brief" of your study:
  - What were main findings
  - Why important
  - What does this imply for policy

Tables/figures: 4 pp (this content mandatory: minimum # of tables and columns in tables)

Table 1 (3 columns): Summary Statistics
- Divide sample into rough comparison groups
- Comparison of means of all variables in your regression AND outcome variables
- T-test of mean differences in third column
- Note: If you want to (or if you accidentally did this on the last problem set), can also present mean differences between naïve comparison groups to show something about biases in naïve regression

Table 2: (4 columns): Estimate
- Naïve estimate
- Alternative estimate without controls
- Alternative estimate with control variables
- Alternative estimate with different outcome or different proxies for variable of interest

Table 3: (3 columns): Robustness check
- Control experiment: Is there a group among whom we wouldn't see this effect if it were true? Or a group among whom it should definitely be smaller?
- Robustness check 1: Is there a group among the effect should be concentrated? Or a group among whom it should definitely be smaller?
- Robustness check 2: Is the Table 2 estimate robust to at least one variation in specification that should be innocuous (i.e. adding more covariates, adding fixed effects, expanding sample, slight variation in outcome variable, etc.)

Figure:
Find one way to present your results graphically: bar chart, graph of time trend, etc.