Unsolicited Advice for Reading Empirical Papers

Josh Clinton

Dept. of Political Science, Vanderbilt University

Fall 2009

There are two types of reading that you will encounter in graduate school (and, indeed, the rest of your academic career). The first type involves reading to get a general sense of the argument being made by the author and to provide a few summaries/tidbits that can be regurgitated upon request. This type of reading can be done in front of the television or while lying on the beach. Passages can be skimmed and pages turned quickly. This type of reading is inappropriate for POL 358.

The second type of reading requires a notepad at the ready as well as intensive concentration. The reading is time-consuming and sometimes exhausting – requiring you to actively interact and question the material being consumed. It is this second form of reading that you should engage in for this class. Because of the required intensity, I have intentionally limited the material to be discussed to a single article. The goal is not to survey the current state of a literature, but rather to critically assess the contribution(s) of a single article.

Actively reading an empirical paper in political science requires you to answer the following questions:[1]

1. **What is the research question?**

   - What are the objectives of the paper?
   - How does the research question relate to existing theoretical and empirical literature?
   - Why is it worth asking?

2. **What is are the data being used here?**

   - How was it collected (and what are the possible sources of error?)
   - What are the important variables and how do they relate to theoretical quantities?
   - How are they defined and why (and what are the possible sources of error?)
   - What is the unit of observation?

3. **What is the empirical strategy for answering this research question?**

   - What are the methodological challenges to answering the research question?
   - If you had an ideal data set, what would it look like? What empirical strategy would you use on it?

---

[1]Similar lists by Philip Haile and John Asker were consulted in the construction of the following list.

- How is the data set in this paper different from that ideal data set?

- How does identification work in this paper? (i.e., what variation in the data identifies the relationship in the data?)

- What are the sources of this exogenous variation?

4. **What econometric/statistical techniques are being used in this paper?**

- Are they appropriate?

- What is the central estimating equation (or equations)?

- What is in the unobservable component (i.e., error)? What is left out of the specification that might matter?

- Are the variables plausibly exogeneous? If not, what are the instruments being used? Do you think they are valid?

- How well does the assumed model fit the data? Is there evidence of over/under-fit? Are there influential outliers that are driving the results?

- How does the econometric model relate to the theoretical framework?

5. **What are the main results of the paper?**

- What are the political science implications of the results?

- What is the statistical and substantive magnitude of the effects?

6. **What do we learn from this paper?**

7. **What questions does this paper leave unanswered? How might you answer them?**