Labor Economics
Guide to Readings

In this class, you will be expected to read many academic papers. This is likely a new experience for most of you, but don't be intimidated! In the first week of the course, we will go through the key empirical tools that you need to get through the main technical parts of the papers. However, do not get too bogged down in the details. Using your economic intuition will help you go a long way in understanding a paper (and a well-written paper should make the intuition crystal clear).

The following are my suggested steps for how to read a paper. Every paper is different, of course, so you'll have to be flexible and adapt. Over time, you'll develop your own steps but hopefully these will help in the beginning.

Steps

1. **Introduction**. Skip the abstract. Go straight into the introduction. After reading the introduction, see if you can answer these questions:
    - What is the motivation? Why should we care about this paper's topic?
    - What question are they trying to answer? Why has no one else been able to answer it before?
    - What is their setting?
    - What is their identification strategy? (i.e. do they explicitly state a technique?)
    - What is the punchline of the results? (i.e. do they find an effect: yes or no? big or small?)
2. After reading the introduction, ask yourself if you could talk (confidently) about this paper for 5 minutes. In other words: do you have a good sense of what this paper is trying to do and what their main result is? If not, re-read the introduction.
    - At this point, you don't need to know anything about the nitty-gritty details (e.g. the exact specifications, their main assumptions, various robustness checks etc).
3. **Literature Review**. Sometimes this is a separate section and sometimes it is interwoven as part of the introduction. Skim this pretty quickly. Your main take-aways should be: a) what are they doing differently than other papers?, and b) why does it matter? Lit reviews often given you a sense of what they are *not* doing, which then makes what they *are* doing clearer.
    - If this is a paper you are reading for your policy memo/research, then take your time here because lit reviews are a gold mine of information. You'll want to learn about the current research on a topic and the authors have done the work for you by summarizing it concisely and logically. It also gives you a bunch of new papers to go and check out.
4. **Model/Theory.** I tried to pick papers that focus more on empirics rather than theoretical models. There'll be a few with a model. This will certainly be too technical for you, so don't worry. Read the beginning, where they should give an overview of what's going on (this might also be in the intro). Start reading the math until you get lost. Then skip through it only stopping to read when you find one of these key phrases: "intuitively" (a simple explanation for the complicated stuff they do, or

"testable implication" (the actual thing they will look for in the data). Read around those words and then just keep skipping.

5. **Background/Setting**. Papers will start out by talking about these grandiose general economic ideas. Then, you get to the data and realize they are actually looking at a very specific case. Make sure you understand the setting, especially if it's something unfamiliar to you (e.g. something in a foreign country). This is a good place to slow down and try to understand most of it. It is usually written in an accessible way and shouldn't involve any math/metrics. Plus, you often just get a cool history/general knowledge lesson.

6. **Data**. It's going to feel a little boring to read about data, but it's essential to know what they are working with. I think your key take-aways from this section should be:

   - What is the unit of observation? What does the data look like? You should be able to sketch out a little example of the dataset
   - What variables do they see? What is the level of aggregation? (e.g. monthly/yearly)
   - Are there any important variables they do not see? Do they have a work-around?
   - Are there any people *not* in their sample that should be? Do they make a case of why it is not a big deal?

   I find that data sections can get very nit-picky about measurement issues and other technical things. Just skim through or skip these parts all together. My take is that if it's important, the authors will re-iterate it when you get to the analysis. Also, these are all well-published papers by top notch researchers, so it's probably not *that* bad (and they're just being meticulous because they're pros).

7. **Summary Statistics**. This sometimes appears as the first part of the Results section, or it might be in the Data or in Setting sections. Either way, after knowing the background and the data, it's often a good time to look at the summary statistics. This is a good way for you to "meet" the data and get a sense of the sample. Take your time here. Don't just focus on what the authors tell you to look at – look at all the parts of the tables/graphs. This is the most transparent part of a paper, so make sure you look out for anything that might be weird and brushed under the rug.

   - I recommend you read with two copies of the paper. One where you read the written parts (physical or electronic) and one copy open on your computer/tablet that you use to look at the figures and tables. I find that printing out figures can sometimes make them much harder to read, so I prefer to keep it electronic.
   - The papers will often have appendices with more summary stats. I recommend you also take a look at those. The appendix will either be right at the end or (for some published papers) it will be a separate file on the journal website.

8. **Empirical Strategy**. This is an important section, but will probably be the hardest to read for you (other than models). Try to do your best, but if you are not getting it, move onto the next section and then come back after seeing the results (I find that it's easier to think about the abstract regressions once you actually see the output). Try to think about the following questions in this section:

   - What is the strategy that they use? (e.g. IV, RD, DID). What is the policy/institutional feature that allows them to use this?
   - Imagine conducting an RCT to study their research question. What would that look like? How does this study try to mimic the RCT?

- Look at their regression equations. There might be multiple. My rule of thumb is to spend time on the first one (if it's first, it's probably an important baseline to understand) and whichever one they call their "main" or "preferred" specification. Look at the subscripts - at what level do they vary? What does the equation look like when the dummy variables turn on ($= 1$) and off ($= 0$)?
- Do they make any assumptions? (hint: the answer is always yes). How believable is it? Try to come up with a counter-example that would violate their assumptions.
- What is the variable that we are really interested in? i.e. which coefficient will give us the causal effect that we are looking for? What doesn't appear in the regression (i.e. is in the error term) but could be correlated with this variable?

9. **Results**. It's easy to get overwhelmed with all the numbers in the tables and graphs. This is a section where it is important to let the authors guide you on where to look (at least on a first read). My few tips:
   - Always map the results back to the equations in the empirical strategy. Which coefficient estimate is a table or graph showing? What is the interpretation of that coefficient?
   - Which results are statistically significant? Does the sign go in the direction you expect to? What happens to the coefficient (and the precision of the estimate) as they add more controls/change the specification?
   - Graphs are often more informative than tables. Make sure you understand every aspect of a graph. The two most fundamental questions are: 1) what is on the x-axis? 2) what is on the y-axis? You cannot understand a graph if you can't even answer those two questions, so take your time in figuring it out (the notes at the bottom will usually help). If there are multiple plots on the same axis, what do they each represent?
10. **Robustness Checks**. Once the authors get through the main question they want to answer, then they get into lots of smaller details. Often at this point, you might be tired of the paper. Just take a quick skim. There are usually sub-headings for each check. If the robustness check sounds interesting to you, then read it. Otherwise, just skip it or read the first and last paragraph so you know what they did (and they'll probably tell you that their analysis is correct). If you have some lingering concerns about the empirical strategy/results see if they address it in the robustness checks. It's always fun when they have a great response for the exact thing you were worried about.
11. **Conclusion**. Read the conclusion. This is not to learn anything new. Instead, use it as a check that you got the main take-aways of the paper.
    - Similarly, I read the abstract at the end. If there is something in the abstract that I totally missed, then I know I've *definitely* done something wrong.
12. Re-read the introduction. You will learn so much more now that you've gone through the paper. Trust me, it's incredible how much information is packed into a (well-written) introduction.
    - Don't feel bad if you don't feel 100% confident about the paper at this point - that's why we will have discussions in class. If you have time, re-read the whole paper again more carefully. You won't fully understand a paper until you've read it two or three times (at least).