Research Paper:
Econ 8250 - Fall 2025

The goal of this research paper is to give you practice with health economics research without requiring you to have data or an innovative research question. There will be eight assignments due at different times throughout the semester. Each time you turn in a new assignment, you will attach it to previous work you turned in. In the end you should have something that resembles a paper, though with simulated data. Since research is a process of constant editing, to encourage you to edit prior work, there will be a grade on the final project as a whole. However, to keep you on schedule, late or unfinished assignments will lose points even if they are completed for the final product.

1 Introductions - Due August 28

First, you should read Don Davis' "Ph.D. Thesis Research: Where do I Start?" and look through Amy Finkelstein's "Finkelstein IAP talk 2007.pdf" slides that are uploaded to eLC. Then, find 3 articles, related to health care, on different topics, in a health policy or medical journal i.e. Health Affairs, New England Journal of Medicine, Health Services Research, BMJ, JAMA, Nature Reviews Drug Discovery, etc. Many of the articles in these journals are medicine specific, you want to look more for "commentary" or opinion type articles that describe a feature of the health care system. You may also look at newspaper articles. Your task is to look at the topic as an economist. Write a 2-3 paragraph introduction to an *economics* paper for *each* article (This should be between 1-3 pages of writing total).

Each introduction should be written in the style of research papers we read. Please do *not* look at the economics literature on these topics beyond the current paper. Each introduction should answer the following questions: What is an economics research question this article raises? Why is it interesting *for an economist*? What is hard about answering this question? What is the policy relevance of this question? The introduction can be a little longer if there is policy background that someone unfamiliar with the question would need to know to understand your introduction.

In addition, to writing mock introductions, you should briefly sketch what you would do with the model, data, and messy data section for each introduction you write. This should be roughly one paragraph for each question. This is meant to make you think through which introduction will be most appropriate to work with going forward.

2 Theoretical Model Intuition - Due September 11

Read Hal Varian's "How to Build an Economic Model in Your Spare Time". Then choose one introduction out of the three you combined to write. Your goal for this week is to think through the intuition of a model for this question. No math is required at this point, but I encourage you to think about the math now. Answer the following questions (if applicable): Who are your agents? What is their objective function? Their constraints? What does demand look like? Supply? What equilibrium concepts are you using? What mechanisms will drive your results? What types of comparative statics or policy counterfactuals would you want your model to have? Are there other things you anticipate needing to fully specify the model? If you were to describe in words, to professor Yoder, the economic environment you are studying how would you describe it?

Note that this is asking about a theoretical model, not an empirical model. My goal is a model that might look like something in Intermediate Microeconomics (i.e. demand and supply curves, utility functions, profit maximization) rather than something you would see in econometrics. A regression equation is **not** what I am looking for in this assignment. Likewise, you do not need to worry about the identification strategy, what data you might use, or controls you plan to include.

Turn in a writeup of your model's intuition after the 1-2 page introduction for just that question. Please try to avoid the literature on this topic for now.

3 Theoretical Model - Due October 2

Write down the math for your model's intuition. Solve for comparative statics (or some other sort of analysis) in a way that captures the economic environment you are working in. Keep it as simple as you can. You may use any assumptions you need to get your results, within reason. For example, feel free to use linear demand, two agents, a static setting. You may assume away adverse selection or moral hazard if it is not central to your mechanism or question.

I anticipate this will be difficult. Start early. You are encouraged to change your model intuition section as you work on this. If you get stuck, some suggestions include: try to simulate your model in Matlab/R, work a very simple example, redefine what you are trying to show, add or subtract assumptions, or write down precisely what is not working in your current setup.

This should read like a model section from the papers we read in class. That is, use words to introduce and explain the equations. You do not need to write down all steps, but there should be enough derivations and written directions that a reader who has taken the first year sequence could rederive your results quickly. Also, this is a theory section. While in general it is good to consider data availability and empirical considerations, I would prefer if these things did not play a role in your model at this stage. Turn this in along with the previous sections, in chronological order of when they were due. Also, incorporate a discussion of your model into the introduction. Please try to avoid the literature on this topic as much as possible for now, except for foundational papers.

4 Theoretical Model Revisions - Due October 16

This never turns out perfectly the first time. We will meet and discuss your first theoretical model and discuss a path to revise.

5 Data - Due October 30

Write a mock data section introduction for your paper. Address the following:

1. Having thought about the theory, why do we need empirics? For example: Can the model generate opposite comparative statics in different settings which all exist? Is there heterogeneity in the effect that is important? Do we care about magnitudes?

2. Reframe the research question to something empirical. Update your introduction to incorporate an empirical component.

3. In your data section, describe the ideal data/experiment to answer that question.

   (a) Simulate this data using the ideal data generating process you described. Write a description of how you simulate this data. (Do not send code, describe it!)

   (b) Allow for some error term to create noise in the data (i.e. your model doesn't predict perfectly). Given what you are simulating, what do you think would be in that error term? What assumptions are you making about it for your identification strategy? (You are free to make unrealistic assumptions here.)

   (c) Estimate your model with this simulated data. This can be reduced form or structural.

   (d) Present a table of results.

   (e) Pretending this is real data, describe your results (magnitudes) and how it answers your research question.

4. Turn this in along with the previous sections, in chronological order of when they were due. Please try to avoid the literature on this topic as much as possible for now.

6 Messy Data - Due November 13

Look through the literature on your topic and find datasets others have used. Save the papers because your next task will be to write a literature review. However, as much as possible try to avoid these papers' methods sections. The goal is to think carefully about data issues on your own. In many cases the authors will do this for you, but you may find they ignore a problem or you have a better way to address their problems. These types of issues are harder to spot if you read their methods sections first.

Find three datasets that have been used in research papers (or you can provide evidence is available for research). It is okay if you cannot get this data, I just want you to get a sense of what data is available and realistic to get. For some topics, there may not be a lot of great

data. That is okay. The goal is to be creative in thinking about ways to answer questions, even if the answer is somewhat indirect or you need to modify your research question.

For each dataset:

1. Describe the data and datasource.
    - How/why was the data generated? What variables are important? How many observations? (An educated guess for the number of observations is fine.)
    - What are the problems with this data?
    - What do you need to fix these problems?
    - If you had these fixes, what specifications would you run?
    - How does this answer your research question? The answer may be indirect or only suggestive. If necessary, you may propose a modified research question the data can answer, so long as your model also covers that question.

Now, choose one dataset or think about some combination of feasible datasets. Think carefully about the actual data generating process for this data. Simulate this data including at least one problem in this data i.e. add endogeneity, add sample selection, omit variables, etc.

1. Fix your empirical strategy by adding new data. Simulate whatever you need to solve the problem. For example, if there is endogeneity, simulate instruments. If applicable, describe concretely what those instruments could be (i.e. natural disasters which close hospitals or a change in the rating areas exogenously changes the risk pool). If this is challenging in your setting, describe what properties they would have to have and why data with those properties may be challenging to find.
2. Writeup a data and methods section.
3. Present a table of results. Pretending this is real data, describe your results (magnitudes) and how it answers your research question.

Turn this in along with the previous sections, in chronological order of when they were due.

7 Literature Review - Due December 4th

1. Write a two page meta-literature review of *economics* papers on your topic. What do you like about each paper? What do you wish you had thought of? What do you think you do better?
2. Write a $\frac{1}{2}$ - 1 page literature review of economics papers as though it was going into a paper. Identify and discuss your contribution. It is okay if someone did exactly what you did. This will not affect your grade for this assignment. If this does happen, I still expect you to find a contribution, even if it seems marginal, or describe small changes you could make to differentiate your project. In practice, this happens often and you will need to do one of those two things.

3. Update your introduction to reframe the sales pitch and your contribution to the literature.

Turn this in along with the previous sections, putting the literature review directly after the introduction.

8 Presentations - Due November 18, 20, or December 4

Each student will present their own paper and discuss another student's paper. Let's plan for a 15-20 minute presentation of your own paper and a 5 minute discussion by another student. For the paper presentation, the first slide should list what type of feedback you would like (pretend you are continuing on with this project). The remainder of the talk should simply be a standard introduction for a talk.

1. Introduce and motivate your question.
2. Explain how you address your question. Describe the theory with no equations. Discuss your "messy" data and empirical your strategy (including your fixes).
3. Preview results and economic intuitions (main take aways).

The goal of being a discussant is to practice giving constructive feedback. What do you like about the paper? If they were to continue on this topic, what could they do better or what should they worry about? Are there other interesting avenues for this question?