

POLICY ANALYSIS FIELD EXAM
9 JUNE 2023
O'NEILL SCHOOL OF PUBLIC AFFAIRS

The exam is in three parts. You should answer all questions. You may use a calculator and/or excel to support calculations in Part 3. Good Luck!

PART 1 – Optimal Social Programs

One strand of research in economics is concerned with the “optimal” generosity of social welfare programs.

- a) Explain the concept of the optimal generosity of social programs, and how it is related to microeconomic theories of behavior and normative theories of social welfare. What are the central theoretical relationships that are being balanced or traded off, according to research that operates from this broad point of view?
- b) The basic theory is somewhat abstract. To make things concrete, choose a specific government social program as an example.
 - Explain the purpose of this government program from the point of view of the optimal generosity literature. That is: what problem is the program supposed to address and how is it supposed to help people. How does the optimal generosity point of view compare with other rationales for this program?
 - Explain what the core theoretical/behavioral relationships mean in the case of this particular program.
 - Explain the ways that the program can be made more generous or less generous by altering certain program details or parameters. What would be the possible downsides of changing program parameters to increase or decrease generosity involve?
- c) How have empirical researchers used data to answer questions that are motivated by theories of optimal generosity of social programs?
 - What relationships does the theory suggest should be estimated?
 - What kinds of causal identification problems come up in empirical work. How are these problems addressed in applications. Define theoretical relationships and constructs carefully, and explain why they may be hard to recover from an econometric point of view.
 - To what extent has the empirical literature has changed the way economists think about the program and its design.
- d) Suppose that the program you are considering is “optimal” at the moment. What situational or contextual factors would need to change in order to justify a change in generosity? Specifically, under what conditions would it be socially optimal to pass a law altering program parameters to increase or decrease generosity?
- e) Suppose that convincing empirical work suggests that reducing the generosity of your program would improve welfare from the point of view of an optimal social programs point of view. That is, cutting generosity would raise social welfare overall. How would you explain this idea to an intelligent and well educated non-economist who thinks it doesn't make sense that reducing program generosity will make people better off overall.

PART 2 – Those Tax Preparers

Background

For most Americans, preparing income tax returns is pretty simple business. That is because their only source of income is “W-2”, or wage, income. For these cases, employers pay the state and federal government the expected tax bill based on wages. At the end of the calendar year, employees receive a “W-2” that indicates total income paid to the employee and total taxes paid to the government on their behalf. The individual taxpayer then needs to fill out their income tax returns that reconciles what their taxable income should actually be based on after various deductions, credits, etc. that may adjust these final numbers. Taxpayers with fairly simple cases (e.g. single earner only from wage sources, no children or dependents, not eligible for special credits, etc.) will probably have paid the exact correct amount over the year. Others, particularly low-income households, will likely be eligible for extra tax credits (e.g. the Earned Income Tax Credit) and therefore will likely have overpaid their taxes and receive a “refund.” Higher income people with lots of other sources of more complicated non-employer income (e.g. interest income, capital income, etc.) will likely have underpaid their taxes from wage income alone and therefore will owe extra payments at the end of the year.

The administrative burden of state and federal income taxes for low-income households is in principle, pretty light. However, research has shown that the prospect of filling out the tax forms seems daunting to them, so they often avoid it and never receive the “refunds” through income assistance programs for which they are eligible. To combat this, there are a variety of public and nonprofit taxpayer assistance programs (TAPs) which will assist low income households for free. There is also a robust private for-profit sector, such as H&R Block and Turbo Tax, that are paid tax preparers (PTPs). Anyone can be a paid tax preparer under federal law, so in addition to well-known corporate names, anyone can make a self-occupation of preparing other people’s taxes. Because low-income households are relatively simple cases for households that are likely to be eligible for a refund, many paid tax preparers attract customers by promising to give them their refund immediately less a percentage of the expected refund for preparing the returns. In this way, they kind of operate like high interest short-term loans: the taxpayer receives a fraction (about 90%) of their future refund a couple of weeks earlier than they otherwise would, and the paid preparer receives the actual refund from the government.

Critics of paid tax preparers are upset with “predatorily” high-interest rates for low-income households. Because low-income wage earners tend to be simple tax returns and there are many forms of free taxpayer assistance programs, it seems unnecessary for them to pay anyone to do the taxes at all. In this way, it is argued that paid tax preparers take advantage of consumer ignorance. A second complaint is that paid tax preparers do a poor job, with the exceptions being the major corporate brands like H&R Block or Turbo Tax. Despite representing a minority of the filings, studies by the Internal Revenue Service found that tax returns prepared by these small single-person paid tax preparers are the source of the majority of tax reporting errors. This suggests that paid tax preparers are poorly informed or too hasty in completing the work. Since anyone can be a paid tax preparer under federal law, this is likely true at least some of the time.

Unlike the federal government, states have the power of occupational licensing which can regulate paid tax preparation. Seven states do regulate paid tax preparers, requiring them to complete a variety of credentials if they prepare taxes for profit (the regulation does not extend to nonprofit taxpayer assistance programs or anyone else performing the task for free). A major advocate for these regulations are the largest corporate paid tax preparers (H&R Block, Turbo Tax, etc.). Critics of the regulation see this as a form of anti-competitive regulatory capture: Large firms can more easily absorb the cost of regulation that will drive their smaller competitors out of business, allowing them to charge higher prices and hold a larger market share, and higher prices are particularly hard on low-income households. Large firms counter that people unnecessarily avoid paid tax preparers like themselves because of the negative reputational harms the industry suffers from these low-quality paid tax preparers (i.e. the regulation solves an adverse selection model and raises product demand). These competing concerns raise questions about the impact of regulating paid tax preparers on their consumers.

Data

Seven states regulate paid tax preparation:

- California since 1997
- Maryland since 2010
- New York and Oregon since 2014
- Connecticut since 2017
- Illinois and Nevada since 2018

All the states do some mix of requiring coursework and training for licensure as a tax preparer, and levy fines and penalties for compliance failures.

Under the Statistics of Income program, the Internal Revenue Services provides public use data on the number of tax returns by state. **For each state and year since 2014**, the following data is available:

- The number of tax returns prepared by the taxpayer (self-prepared) by income group
- The number of tax returns prepared by a paid tax preparer (paid prepared) by income group
- The number of tax return prepared by a volunteer non-paid preparer (volunteer prepared) by income group

For example, in 2017, Ohio for households with less than \$30,000 in gross income, 140,000 returns were submitted with a paid preparer, 240,000 submitted with a volunteer, and another 1,480,000 submitted a self-return. For our purposes, you can assume that we have this data by income deciles.

Furthermore, the Bureau of Labor Statistics conducts an Occupational Survey of Employment and Wages. This survey takes a large nationally representative sample of persons, and collects information about their employment circumstances (employment status, time since changed jobs, wage, educational attainment, etc.). Since 2000, “tax preparer” is an occupation the survey respondent can list. The respondent’s state of residence is also an available field, so we can find information on how prevalent employment in tax preparation services are in a state and what they make in wages.

Questions

- (a) Develop an economic model to help understand how state regulation of paid tax preparers might affect some aspect of the market for these services. Your model does not need to account for all possible factors or explain every possible outcome. There are many ideas in the set-up above, and you should probably focus on a very limited set. Make sure to give precise explanations of the inputs to the model, what is being maximized by the agents in the model, and what tradeoffs are at play in the model.
- (b) Using the model you developed in part a, generate testable hypotheses about the effect of state regulation of tax preparation services from some data described above. For each hypothesis, describe how it follows from the intuition of your theoretical model. Be sure to express in the hypothesis statement whether the effect of the policy is positive, negative, or ambiguous. (Your model might suggest countervailing effects that results in ambiguous predictions, this is fine and you can simply point it out.) To the extent possible, you should describe indicators you could construct from the data provided that might help you test the hypothesis and comment on how well the data-based indicators connect with the theory.
- (c) Identify some hypotheses that are interesting but do not emerge directly from your model, or are not testable from the available data. Comment on what is missing/ignored in the theoretical model or data that would have to be changed in order to explore the new hypothesis. If possible, describe an existing literature where the new hypothesis, modeling framework, or type of data would be at home.
- (d) The Fancy Think Tank of Occupational Licensing wishes to hire a new young scholar whose research agenda will be attractive to state policy makers on how to regulate labor markets. They have a lot of money to fund an ambitious research agenda. Develop a high-level research plan around state regulation of paid

tax preparation that will result in publishable academic work. Your research plan can involve causal program evaluation, exploration of policy-relevant economic parameters, a predictive model for describing policy impacts, qualitative studies, etc. Whatever you want to do, you should clearly describe the key details of your overall plan: research question, unit of analysis, research design, statistical model, assumptions, threats to validity, and proposed sensitivity checks).

PART 3 – Wellness

Suppose that a collection of midwestern states have been experimenting with new school based mental health interventions which are supposed to promote a healthier mindset and reduce rates of depression among adolescents. The same intervention was introduced in middle schools at different times in Michigan (2000), Indiana (2002), and Ohio (2011). (This means – for example – that Indiana actually operated the program in 2002 and every year after that but it did not operate the program in 2001 or before.) Wisconsin has not adopted the program for its public schools but is starting to review evidence on the matter to decide if the program would be worthwhile.

Each of these states administers a survey based mental health assessment to all students each year. The data from the assessment can be used to form a summary *wellness score*. The normalized wellness score theoretically ranges 0 to 200. The average teenager scores about 100 on the normalized score. People experiencing feelings of depression and anxiety have lower scores than people who feel comfortable and engaged.

The attached table shows the average wellness score among 7th grade girls in each state from 2000 to 2010. The attached figure shows a plot of these numbers to make things easier to see. Throughout, you can assume that the averages in the table are averages across the entire population of 7th grade girls attending public schools in each state.

- a) Develop some mathematical notation to describe the causal effects of the mental health intervention on wellness outcomes among 7th grade students. Your notation should show the connection between observed outcomes and potential outcomes and allow the possibility of heterogeneous treatment effects.
- b) Andy says that the Midwest provides a great platform for studying the effects of the mental health program. He thinks that it would make good sense to analyze the data using a twoway fixed effects regression model. Write down the statistical model that Andy has in mind. Under what assumptions will this approach produce “good” estimates of the causal effects of the mental health program? Are the assumptions credible? Which assumptions do you think are most problematic.
- c) Michelle says that Andy’s model runs some risks of the so-called “negative weighting problem” discovered in papers by Goodman-Bacon (2021) and de Chaisemartin and D’Haultfortuile (2020). What does she mean? Give an example of such a problem using the numbers from in the table below. Do you think this type of bias appears to be an actual problem in this example? Or is it just an obscure theoretical possibility?
- d) Amy agrees that Andy’s idea has some problems and that she thinks it would be wiser to think in terms of group-time average treatment effects on the treated following ideas developed by Callaway and Sant’Anna (2021). Use the notation you developed above to define this idea of a group-time causal effect. Use the idea to describe (conceptually) the causal effect of the program among grade 7 students in Ohio in the years 2010, 2011, and 2012. Your answer should use a mix of words and notation to explain the idea of the effect. (No need to actually calculate anything at this stage. Just define the causal effect we are talking about as though you had “complete” information about potential outcomes.)
- e) Suppose that you want to estimate the group-time ATTs for Indiana at three points in time: 2002, 2003, and 2004. Propose simple 2 x 2 DID estimators for each of these three effects. Carefully state the assumptions under which your estimator will identify each group-time ATT.
- f) Calculate the estimates of these three ATTs using the method you laid out above and the information in the table. (You can use **excel** or a **calculator** to help with the math. But you should show your work in your answer.)
- g) Maria thinks the results are really only interesting after the program has been in place for five years. She would like to take a simple average of the Indiana and Ohio group-time ATTs at five years post adoption in each state. Compute this average of the two effects.
- h) The overall research design here rests on some important assumptions. What kinds of robustness checks and sensitivity analyses would you pursue if had more time and computational resources.

Year	Indiana Adopts in 2002	Ohio Adopts in 2011	Michigan Adopts in 2000	Wisconsin Never Adopts
2000	110	137	160	97
2001	106	133	156	93
2002	117	134	157	94
2003	113	129	152	89
2004	118	133	156	93
2005	120	134	157	94
2006	119	132	155	92
2007	119	131	154	91
2008	125	136	159	96
2009	119	130	153	90
2010	122	133	156	93
2011	122	143	156	93
2012	126	152	160	97
2013	127	158	161	98
2014	119	155	153	90
2015	122	163	156	93

