POPULATION FIELD EXAM

2020– 2021

Instructions

- You have 4 hours to complete this exam. The exam ends at 1pm PT.
- This is an open-book exam. You may consult your notes from class or source from the internet.
- You can also use a calculator.
- There are 2 parts in this exam: Part I covers Professor Lleras-Muney's course, Part II covers Professor Goncalves' course.
- Answer each part in a separate document. Upload each part **separately** to the website of each of the courses. There is a link on each course page where you can do this. The website will time stamp your submissions.
- You can answer this exam by hand and just take pictures of your answers.
- Please submit a SINGLE document for part I and another document for part 2. Make sure that your name and ID are clearly displayed. If you experience any problems submitting the exam you can always email it to Chiara or to us directly.
- If you have any technical issues and need to reach us directly please contact us at 310-359-5529 (Adriana) or 205-413-9193 (Felipe).

YOU MUST ANSWER BOTH PARTS AND OBTAIN AT LEAST 75% IN *EACH* PART TO PASS THE FIELD EXAM.

Part I: Population Economics (Fall 2020)

Professor Adriana Lleras-Muney

TOTAL: 100 POINTS

Answer both questions. Partial credit will be given whenever possible.

A. **(50 points) Apply findings from the course to the current COVID crises**

1. (10 points) Historically how were infectious diseases defeated? Are there any lessons from the eradication of infectious diseases (historically or in developing countries today) that apply to the COVID crises?
2. (10 points) Given what we know about education and health, what would you expect to be true about the relationship between education and COVID mortality? Between education and COVID-protective behaviors? What mechanisms might be at play today that mediate this relationship?
3. (10 points) Use estimates of the value of a statistical life to estimate an upper bound on the optimal level of government expenditures to fight COVID.
4. (10 points) Based on the existing literature on the relationship between disease and development, what do we predict the effects of COVID would be on economic output and individual incomes? Are there differences in the short and long term predictions?
5. (10 points) RCTS and COVID. Discuss the ways in which RCTs for COVID vaccines have advantages and limitations.

B. (50 points) Empirical methods

Aizer et al. (2020) use a difference-in-difference approach to estimate if WWII contracts, given to companies to produce war-related goods (such as bombers and ships) from 1941 to 1945, affected the labor market outcomes of Black and White workers from 1940 to 1950. These contracts were given on the condition that the companies not discriminate in hiring on the basis of race. Using data from 143 cities (m) and 2 census years (t = 1940 and 1950), Aizer et al. estimate the following regression, separately by race:

$$Y_{mt} = \beta_0 + \beta_1 WW\$_m * Post_t + Metro\ FE_m + Post_t + \gamma X_{mt} * post + \epsilon_{mt}$$

where Y_{mt} is an outcome of interest (such as the employment rate among 25–54-year-old population, the share employed in skilled occupations or average wages) for a given city and time period, $WW\$_m$ is the total war expenditure per capita for metro *m, Post* is an indicator for 1950 and X_{mt} is a vector of controls from 1940 Census, such as % employed in manufacturing. Table 2 below presents the main estimates.

1. (5 points) Column 1 shows the estimates without any controls. What kind of estimation of this? What is the main identifying assumption?

2. (5 points) Column 2 shows the estimates with controls for Black workers. What is the main identifying assumption now? Why do we have X_m*post instead of X_{mt}?
3. (5 points) How do you interpret these coefficients? Use the results of wages and skilled occupations to illustrate.
4. (5 points) The racial gap in the share of workers semi-skilled decreased by 8.1 percentage points between 1940 and 1950. Can the DD results presented here be used to estimate how much of this gap is explained by WWII contracts? Explain how you can use the DD estimates to do this and also why this back of the envelope computation might be wrong.
5. (5 points) How can the authors verify identification assumption with additional data?
6. (5 points) There was a large amount of migration as a result of WWII$ (the authors document this). What concerns does migration raise? How can the authors address these concerns?
7. (5 points) Suppose you have data that is further disaggregated by industry. So that now an observation is a city, year and industry mean, where the industry is just broken down into defense and non-defense industry. Defense industries are defined as those that receive WWII monies. Write down the DDD specification you could estimate with these data.
8. (5 points) What would the identifying assumption be in these new DDD approach? List one advantage and one disadvantage of employing this strategy.
9. (5 points) Suppose instead that you wanted to use a matching approach to identify or construct counterfactuals. How would you implement a propensity score DD approach? Write down the estimating equations.
10. (5 points) What advantages and disadvantages would this approach have?

Table 2: Effect of war expenditures (1940-1950)

	(1)	(2)	(3)	(4)
	Black Men		White Men	
	Basic	Controls	Basic	Controls
Panel A: Share skilled				
War exp per capita * Post	0.012**	0.015***	-0.000	0.001
	(0.005)	(0.006)	(0.002)	(0.002)
Mean Y - 1940	0.33	0.33	0.77	0.77
Mean Y - 1950	0.48	0.48	0.83	0.83
Panel B: ln(Average yearly wage)				
War exp per capita * Post	0.035**	0.031**	0.006	0.007
	(0.015)	(0.015)	(0.004)	(0.004)
Mean Y - 1940	6.55	6.55	7.26	7.26
Mean Y - 1950	7.06	7.06	7.56	7.56
Panel C: Prime-age employment rate				
War exp per capita * Post	-0.005	-0.006	-0.004**	-0.003*
	(0.005)	(0.006)	(0.002)	(0.002)
Mean Y - 1940	0.80	0.80	0.88	0.88
Mean Y - 1950	0.84	0.84	0.92	0.92
Metro areas	147	147	147	147
Mean war exp per capita	1.46	1.46	1.46	1.46
Metro FE	X	X	X	X
Division-Year FE	X	X	X	X
Baseline controls	-	X	-	X
Draft control	-	X	-	X

Note: Sample is 147 metro areas. See equation 1 for the basic specification. War expenditure is $1000s per capita. Share skilled is the share of employed men who are not farmers, laborers, or service workers. Wages are total wage earnings in the previous year for men who are currently employees. Prime-age employment is the share of men ages 25-54 who are employed. Baseline controls are 1940 variables interacted with a post indicator: average years of education, share employed in manufacturing, share employed in agriculture, and share Black. Draft control is predicted draft rate based on 1940 demographics. Primary data sources are 1940 (100%; 5% sub-sample for whites) and 1950 (1%) Census samples. Metro area definitions are based on 1940 and 1950 Census Bureau definitions. All values are in 1940 dollars. Regressions are weighted by relevant population. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Population Economics Exam, Part II

Felipe Goncalves

Winter 2021

I have just opened a karate dojo in Los Angeles. The purpose of the dojo is to help teenagers improve their self-defense abilities. Some critics say that the training is not effective, so I have decided to conduct a study to settle this debate.

Define the following variables: $T_i \in \{0,1\}$ is an indicator for whether an individual participates in the dojo. $W_i^* \in \{0,1\}$ is a latent variable for whether the individual would win a fight if they encountered that situation, and $F_i \in \{0,1\}$ is an indicator for whether the individual does enter a fight. The product of these two variables is an indicator for whether the individual has won a fight, $W_i = F_i \times W_i^*$.

The variables F_i, W_i^*, and W_i are all potentially a function of whether they join the dojo, $F_i(t)$, $W_i^*(t)$, $W_i(t)$.

I have run an experiment where I randomly send notices to half the teenagers in Los Angeles encouraging them to join, $Z \in \{0,1\}$. I have also collected administrative data on all LA teenagers, whether they got into a fight in the last year, and whether they won.

The participation rates are the following:

$$\begin{aligned} E(T_i|Z_i=0) &= 1/4 \\ E(T_i|Z_i=1) &= 1/2 \end{aligned}$$

 # 1 (10 points) Using the language of Angrist, Imbens, and Rubin (1996), what share of teenagers are always takers? Compliers? Never Takers?

 # 2 (10 points) Now suppose we observe the fighting rates F_i for teenagers separately by dojo

participation and whether they received a notice:

$$
\begin{aligned}
E(F_i|T_i{=}0,Z_i{=}0) &= 1/2 \\
E(F_i|T_i{=}1,Z_i{=}0) &= 7/8 \\
E(F_i|T_i{=}0,Z_i{=}1) &= 1/2 \\
E(F_i|T_i{=}1,Z_i{=}1) &= 3/4
\end{aligned}
$$

What is the IV estimate of the impact of dojo participation on fighting propensity?

 # 3 (10 points) What are some possible violations of the IV assumptions that would lead to the IV estimate not reflecting a LATE for compliers? Explain in words for this specific context.

 # 4 (10 points) Describe in words how you would test for the validity of the IV design in this setting.

 # 5 (10 points) Who has higher fight propensity, F_i, with karate training, the always takers or the compliers?

 # 6 (10 points) Who has higher fight propensity, F_i, *without* karate training, the compliers or never takers?

 # 7 (10 points) I also want to know the impact the dojo has on the fighting success of its students. We estimate the following values:

$$
\begin{aligned}
E(W_i|T_i{=}0,Z_i{=}0) &= 0 \\
E(W_i|T_i{=}1,Z_i{=}0) &= 7/16 \\
E(W_i|T_i{=}0,Z_i{=}1) &= 0 \\
E(W_i|T_i{=}1,Z_i{=}1) &= 3/8
\end{aligned}
$$

What is the IV estimate for the impact of T_i on W_i?

 # 8 (10 points) The estimate of the dojo's impact on winning a fight, W_i, includes both the propensity to fight F_i and the student's underlying winnability, W_i^*, where $W_i = F_i \times W_i^*$.

Construct bounds on the impact of T_i on W_i^* for compliers who would fight regardless of T_i. *Hint: Calculate* $E(F_i(0)|compliers)$, $E(F_i(1)|compliers)$, $E(W_i(0)|compliers)$, $E(W_i(1)|compliers)$, *and use the Lee (2009) approach among compliers.*

 # 9 (10 points) I am hoping to eventually expand my operations to the whole city and increase the number of students who join. Should we expect that the effect of the dojo's training on fight propensity, F_i, will be greater or smaller among the new students from an expansion than the students induced to join by the randomized notice? *Hint: Use the Brinch, Mogstad, and Wiswall (2017) approach.*

 # 10 (10 points) Suppose I want to separately know the impacts of joining the karate dojo relative to no karate training *and* the impacts relative to joining a competing karate dojo. In words, explain how we could answer this question. Suppose we can conduct a new experiment, collect new data, etc. Anything necessary to answer the question.