POPULATION FIELD EXAM

2015 – 2016

Instructions

- You have 4 hours to complete this exam.
- This is a closed book examination. No written materials are allowed.
- You can use a calculator.
- There are 2 parts in this exam: Part I covers Lleras-Muney's course and Part II covers Edward Kung's course.
- Please answer Parts I and II in separate booklets.

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

Part I: Population Economics (Fall 2015)

Professor Adriana Lleras-Muney

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

A. <u>(40 points) Review. Please answer the following questions:</u>

1. (10 points) Education and health
    a. What is the evidence on the relationship between education and health?
    b. What are the policy implications of this debate?
    c. Why is it important to document mechanisms if the relationship exists?

2. (10 points) Review Manski's concepts on social effects.
    a. What are endogenous social effects? Contextual effects? Correlated effects? Give an example of each.
    b. Why we are interested in differentiating between them?
    c. What is the relation between networks and peer effects?

3. (10 points) Genes and social outcomes
    a. Why are old estimates of genetic effects likely to be biased?
    b. How is the effect of specific genes on outcomes identified today?
    c. What are the policy implications of finding genetic effects on social outcomes?

4. (10 points) Addiction
    a. What are the three main theories of addiction in economics?
    b. What are their policy implications?
    c. Why do economists care about addiction?

B. <u>(80 points) Value of a Statistical Life</u>

Using state-level data data on car accidents, time traveled and speed limits, Ashenfelter & Greenstone (2004) estimate the value of a statistical life.

a-(10 points) The authors argue that changes in the speed limit provide a natural setting to estimate the value of a statistical life. Why is this approach preferable than older approaches to estimating the value of life that regress wages on occupation-specific mortality rates?

b-(20 points) Ashenfelter and Greenstone write "We compare changes in fatality rates and speeds on rural interstates across states that did and did not adopt the 65-mph limit."

- (5 points) Based on this description, write down the estimating equation and the identifying assumptions.
- (5 points) What evidence can one present in support of the identifying assumptions?

- (5 points) What kind of omitted variable would ruin this DD estimation of this type? Given an example.
- (5 points) Should the standard errors be clustered and why?

c-(20 points) Table 2 reporting results from the DD estimation, is reproduced below.

- (5 points) The authors present results for fatality rates and for speed. Why?
- (5 points) What issue are the authors trying to address with columns 2 and 3?
- (5 points) Why are panels B and C presented? What do these estimates address?
- (5 points) how would you estimate a DDD model using unaffected roads? Write down the estimating equation. What is the identifying assumption now?

TABLE 2
DIFFERENCE IN DIFFERENCES (DD) ESTIMATES OF 65-MPH SPEED LIMIT ON FATALITY RATES AND SPEEDS

	DD of Levels (1)	DD of Levels Normalized by Preperiod Level in Adopting States (%) (2)	DD of Natural Logarithms (3)
	A. Rural Interstates (Affected Road Type)		
Fatality rate	.185	.130	.311
Speed	2.8	.047	.045
	B. Urban Interstates (Unaffected Road Type)		
Fatality rate	−.052	−.059	−.063
Speed	−.5	−.009	−.009
	C. Rural Arterials (Unaffected Road Type)		
Fatality rate	−.123	−.032	.005
Speed	.5	.009	.008

d-(20 points) For the main analysis the authors estimate two equations

$$\text{Ln(fatalities)}_{sry} = c_1 + a_1 * \text{I(=1 if 65 limit in force)}_{sry} + b_1 * \ln(\text{VMT})_{sry} + e_{sry}$$

$$\text{Ln(hours traveled)}_{sry} = c_2 + a_2 * \text{I(=1 if 65 limit in force)}_{sry} + b_2 * \ln(\text{VMT})_{sry} + u_{sry}$$

where ln(fatalities) is the natural log of fatalities in a given state *s*, road-type *r* and year *y*.

- (5 points) Why do we need to control for miles of travel?
- (5 points) The equations they estimate include many fixed effects. Which fixed effects need to be included to make this a proper DDD specification? Explain.
- (5 points) How can you use the estimated coefficients a_1 and a_2 to compute the value of life? What additional information/data do you require for this computation?
- (5 points) Results are presented in Tables 4 and 5 reproduced below. Use these numbers, make assumptions about additional data, and compute the value of life

e-(10 points). In class we studied the paper by Becker et al which uses data from GDP and life tables to estimate the value of life years. What are the main differences with the approach here?

TABLE 3

PROPORTIONATE (LOG) EFFECT OF THE ADOPTION OF THE 65-MPH SPEED LIMIT ON FATALITIES, CONTROLLING FOR THE OBSERVED MILEAGE, BY ROAD TYPE

	AFFECTED ROAD TYPE	UNAFFECTED ROAD TYPES	
SAMPLE	Rural Interstates (1)	Urban Interstates (2)	Rural Arterials (3)
	A. Annual Effects		
1982–86, 1987	−.098	−.203	−.062
	(.195)	(.174)	(.119)
	[165]	[162]	[162]
1982–86, 1988	.351*	−.223*	−.073
	(.165)	(.111)	(.050)
	[167]	[163]	[162]
1982–86, 1989	.473	−.062	.021
	(.259)	(.142)	(.071)
	[167]	[162]	[162]
1982–86, 1990	.268	.073	.181*
	(.163)	(.161)	(.090)
	[166]	[163]	[162]
1982–86, 1991	.202	−.097	.238**
	(.123)	(.135)	(.073)
	[166]	[163]	[162]
1982–86, 1992	.399**	−.012	.140
	(.162)	(.190)	(.087)
	[164]	[162]	[162]
1982–86, 1993	.493**	−.059	.113
	(.179)	(.154)	(.077)
	[165]	[162]	[162]
	B. Average Effect		
1982–93	.360**	−.056	.082*
	(.091)	(.073)	(.040)
	[326]	[327]	[324]

TABLE 4

PROPORTIONATE (LOG) EFFECT OF THE ADOPTION OF THE 65-MPH SPEED LIMIT ON HOURS REQUIRED TO TRAVEL THE OBSERVED MILEAGE, BY ROAD TYPE

	AFFECTED ROAD TYPE	UNAFFECTED ROAD TYPES	
SAMPLE	Rural Interstates (1)	Urban Interstates (2)	Rural Arterials (3)
	A. Annual Effects		
1982–86, 1987	−.039*	−.014	−.047
	(.018)	(.018)	(.025)
1982–86, 1988	−.041**	−.002	−.006
	(.009)	(.011)	(.007)
1982–86, 1989	−.038*	.004	.007
	(.018)	(.014)	(.013)
1982–86, 1990	−.025	−.011	.002
	(.017)	(.016)	(.013)
1982–86, 1991	−.043**	−.012	.005
	(.017)	(.015)	(.010)
1982–86, 1992	−.057**	−.021	.003
	(.017)	(.019)	(.014)
1982–86, 1993	−.054**	−.024	.002
	(.015)	(.014)	(.016)
	B. Average Effect		
1982–93	−.041**	−.009	−.000
	(.007)	(.007)	(.007)

Part II: Housing Markets and the Evaluation of Spatial Amenities (Winter 2016)

1. (40 pts) Answer the following questions in a few paragraphs. Be sure to cite relevant papers where appropriate.

   (a) Economists are often interested in estimating the value of amenities that are not explicitly traded in a marketplace, such as the value of neighborhood safety or of public education. Briefly describe how economists have attempted to do this. (10 pts)

   (b) When constructing quality of life rankings, Los Angeles often comes out near the top. Describe the intuition behind quality of life rankings and explain in a few short sentences why Los Angeles is often highly ranked. (10 pts)

   (c) Set up and describe Rosen's 1974 hedonic model (focus only on the consumer side). What can we learn from the equilibrium conditions of the model? What is Rosen's "two-step" method and what are some of its identification challenges? (10 pts)

   (d) Describe the user-cost model of house prices. What does the equilibrium relationship between house prices and rents depend on? What is missing from the user cost model? (10 pts)

2. (40 pts) Suppose you have a cross-section of data on housing transactions. The data allows you to see the street address of the home, the sale price, and the characteristics of the home, and neighborhood characteristics. You also have data on public schools, including each school's test score performance and attendance boundaries. You are interested in using this data to estimate the willingness to pay for school quality. (40 pts)

   (a) Suppose you regressed sale price on house characteristics and test scores for the school that house is assigned to. What would you expect to find? Describe some problems with this regression. (10 pts)

   (b) Come up with an alternative strategy for estimating the average willingness to pay for school quality. Make sure to write down the estimating equation. If you are deriving your identification strategy from another paper, cite that paper. How do you expect the new strategy to affect your estimates from part (a)? (10 pts)

(c) Describe some of the assumptions necessary for your answer in part (b) to be valid. How might you test these assumptions with the data that you already have? Feel free to draw the hypothetical pictures/tables you would use to test these assumptions. (10 pts)

(d) Are there remaining concerns that are not addressed in your answer to part (c)? What additional data would you need? (10 pts)

3. (40 pts) Consider a housing market with consumers $i = 1, \ldots, N$ and houses $j = 1, \ldots, N$. Each consumer chooses one house and each house gets chosen by one consumer. The utility that consumer i gets from choosing house j is:

$$V_{ij} = \alpha_i x_j - \beta_i p_j + \xi_j + \epsilon_{ij}$$

where x_j is an observed (to the econometrician) house characteristic, p_j is the price of house j, ξ_j is unobserved house quality, and ϵ_{ij} is a preference shock that is iid type 1 extreme value. α_i and β_i are preference parameters given by:

$$\begin{aligned} \alpha_i &= \alpha_0 + \sum_{k=1}^{K} \alpha_k z_{i,k} \\ \beta_i &= \beta_0 + \sum_{k=1}^{K} \beta_k z_{i,k} \end{aligned}$$

where z_{ik} are demographic characteristics of consumer i. As the econometrician, you observe x_j, p_j, z_{ik}, and d_{ij}—an indicator for whether consumer i chose house j.

(a) What is the probability that person i chooses house j? (10 pts)

(b) Let us write $\delta_j = \alpha_0 x_j - \beta_0 p_j + \xi_j$. Describe a computationally tractable strategy for estimating δ_j for each house. (10 pts)

(c) Describe a two-step procedure for estimating all the preference parameters, $\alpha_{0:K}$, $\beta_{0:K}$. (10 pts)

(d) Suppose one were to regress p_j on x_j. Interpret the coefficient on this regression. Under what conditions will it correctly estimate the average willingness to pay for x? When will it not? (10 pts)