Last Name:

First Name:

Requirements: You have 110 minutes. The last three pages are marked scrap: you must detach them. There are no aids allowed. Keep these papers closed on your desk until the start of the test is announced. Only test pages 5 – 8 and not scrap pages will be marked. Write your final answer on pages 5 and 6: use page 7 only if absolutely necessary and page 8 in an emergency. Remember: conciseness affects your mark. (Conciseness does *not* mean writing a lot but in small handwriting to fit in two pages.) No additional paper is provided beyond page 8: your answer must fit. Write legibly.

Background: These sources – provided ahead of time – prepare you for this test: **(1)** The first five weeks of course material including lectures, homework and required readings with a special emphasis on week 4 and 5 materials (scatter plots, association, correlation, causality, and regression analysis); **(2)** David Leonhardt, July 22, 2013, *The New York Times* "In climbing income ladder, location matters" http://www.nytimes.com/2013/07/22/business/in-climbing-income-ladder-location-matters.html; **(3)** Executive Summary, January 2014 "Where is the land of opportunity? The geography of intergenerational mobility in the U.S." Raj Chetty, Nathaniel Hendren, Patrick Kline and Emmanuel Saez http://obs.rc.fas.harvard.edu/chetty/website/v2/Geography%20Executive%20Summary%20and%20Memo%20January%202014.pdf; **(4)** "Frequently Asked Questions" on "The Equality of Opportunity Project" website: http://www.equality-of-opportunity.org/index.php/faq-s; **(5)** Abstract, Introduction (pp. 1 – 5), Conclusion (pp. 42 – 43) and Figure VIII (p. 91 of pdf) of Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014) "Where is the land of opportunity? The geography of intergenerational mobility in the United States" *NBER Working Paper*, http://www.nber.org/papers/w19843

Structure: This test gives quantitative results that you interpret. A single main question is followed by supporting questions that help guide you towards a full and organized answer to the main question.

Guide to answering effectively:

- Directly and fully address the question. A complete answer addresses each supporting question.
- To answer the questions, *apply* relevant course concepts to the quantitative results given.
- Critically evaluate the results: assess any violations of underlying assumptions and limitations of the data.
- Write your answer using *complete sentences and well-organized paragraphs*.
  - A typical paragraph is 4 to 6 sentences. A short paragraph, which is at times suggested, is 2 to 4 sentences.
  - Answer with multiple well-structured paragraphs that coherently answer the main question. Each paragraphs should have a clear topic sentence followed by supporting points.
  - Your response should be about 2 pages (handwritten with average handwriting).
  - Spend about 55 minutes planning/outlining your arguments and about 55 minutes writing and revising.
  - Use the scrap pages for a draft/outline.
  - Revise and edit your draft to achieve a clear, coherent, and concise writing style.

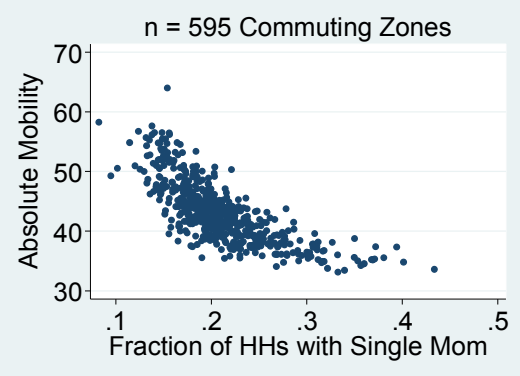
Student #:

Maximum Points	Raw Total (points earned: max 47)	Your Mark (out of 100%)
47 **(Note: 50 points possible, but marked out of 47)**		

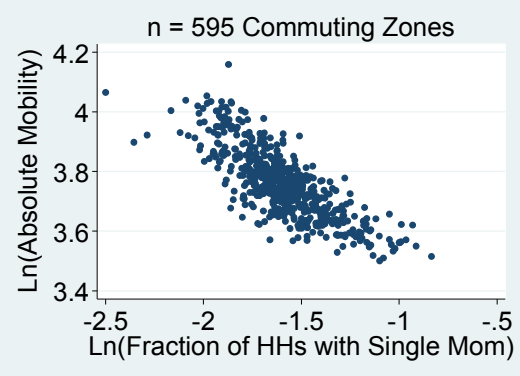
The Equality of Opportunity Project, led by a team of top academic economists – Raj Chetty, Nathaniel Hendren, Patrick Kline and Emmanuel Saez – provides data and analysis exploring intergenerational income mobility in the U.S. The data include many variables for each of 741 "Commuting Zones" (CZ). The analysis below considers a small subset of variables. The unit of observation in these cross-sectional data is a CZ, which is a collection of geographically adjacent counties. The analysis considers only CZs with a population of at least 25,000 people. Further, some variables are missing for some CZs and *some* regressions focus on seven states in the Southeast of the U.S. (Mississippi (MS), Alabama (AL), Georgia (GA), South Carolina (SC), Tennessee (TN), North Carolina (NC), and Florida (FL)), which is why the number of observations varies across regressions. When used, the Southeast subsample is abbreviated "(MS,AL,GA,SC,TN,NC,FL)."

Variable name	Description
absolute_mobility	The average percentile in the national income distribution of a child who is born to parents at the 25th percentile in the national income distribution: the higher it is, the higher mobility is
relative_mobility	The average percentile in the national income distribution of a child who is born to the richest parents (top 1 percent of the national income distribution) minus the average percentile of a child born to poorest parents (bottom 1 percent): the higher it is, the *lower* mobility is
frac_single_mom	Fraction of households (HHs) with children that are head by a single mother: number of single female households with children divided by total number of households with children
stud_teach_ratio	Average student-to-teacher ratio in public schools (i.e. how many students for each teacher): a measure of educational quality and investment

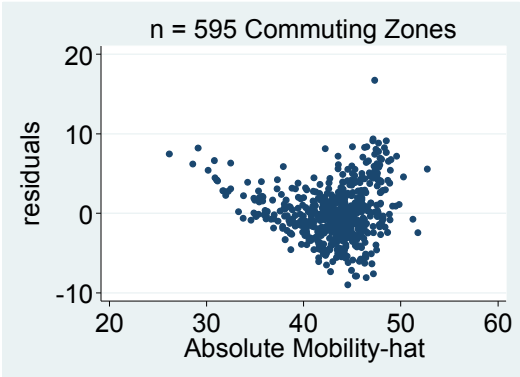
Graph 1:



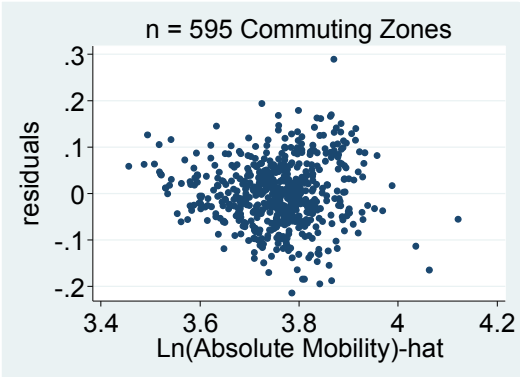
Graph 2:



Graph 1A:



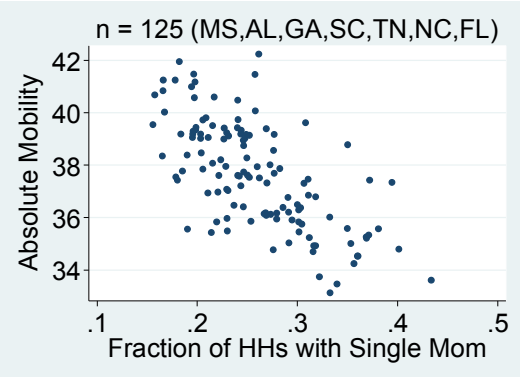
Graph 2A:



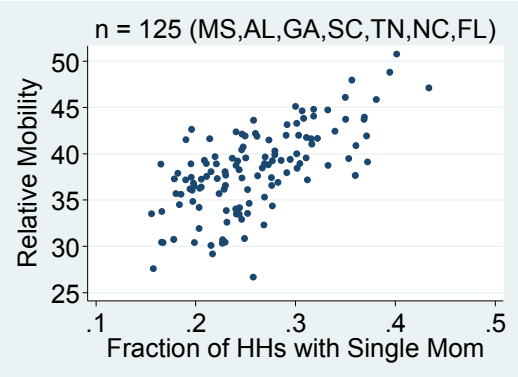
Regression 1: absolute_mobility-hat = 58.92 – 75.56*frac_single_mom, n = 595, R^2 = 0.56

Regression 2: ln(absolute_mobility)-hat = 3.12 – 0.40*ln(frac_single_mom), n = 595, R^2 = 0.63

Graph 3:



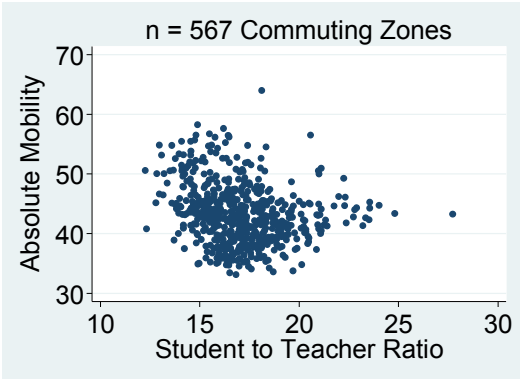
Graph 4:



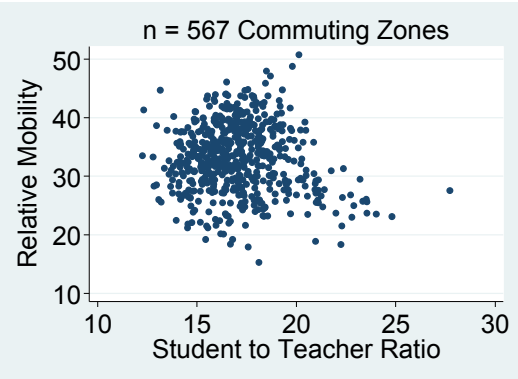
Regression 3: absolute_mobility-hat = 43.58 – 22.98*frac_single_mom, n = 125, R^2 = 0.45

Regression 4: relative_mobility-hat = 25.20 + 50.44*frac_single_mom, n = 125, R^2 = 0.44

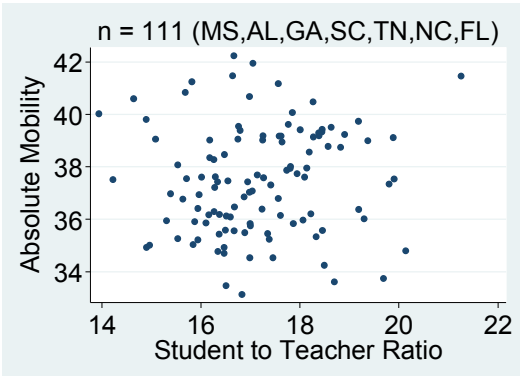
Graph 5:



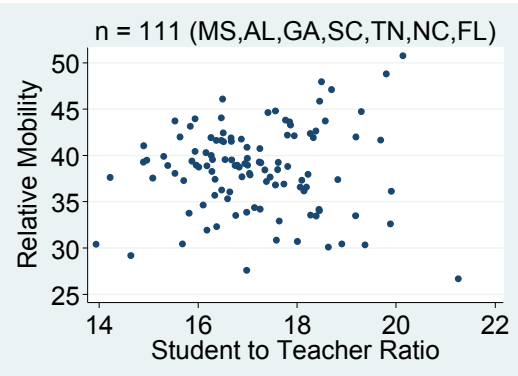
Graph 6:



Graph 7:



Graph 8:



Regression 5: absolute_mobility-hat = 52.50 – 0.54*stud_teach_ratio, n = 567, R^2 = 0.05

Regression 6: relative_mobility-hat = 37.96 – 0.28*stud_teach_ratio, n = 567, R^2 = 0.01

Regression 7: absolute_mobility-hat = 35.29 + 0.13*stud_teach_ratio, n = 111, R^2 = 0.007

Regression 8: relative_mobility-hat = 36.53 + 0.12*stud_teach_ratio, n = 111, R^2 = 0.001

Your audience: Presume that your readers have also just read the previous two pages. However, they need your help to understand what the numbers and graphs mean and what conclusions should be drawn. Last year, your readers took a course like ours but struggled with it. Strive to write a suggested solution that explains the results and conclusions using a professional tone.

Main question: What do the eight regressions mean?

Supporting questions: (These help guide you towards a full and organized answer to the main question.)

- Discussion of the research questions and data: (one short paragraph recommended)
  - Which research questions do the regressions raise?
  - Are the data observational or experimental? Explain.
- Comparison of Regression 1 and 2: (one paragraph recommended)
  - Making specific reference to the relevant graphs (including 1A and 2A), overall, how do Regressions 1 and 2 compare?
  - What is the *full interpretation* of the slope coefficient in Regression 2? Give a specific numeric example to illustrate the interpretation of the slope coefficient.
- Comparison of Regression 1 and 3: (one paragraph recommended)
  - What are the *full interpretations* of the specific numbers – intercept, slope, and R^2 – for Regression 3?
  - Making specific reference to the relevant graphs and regression results – including the slope and the R^2 – how do Regression 1 and 3 compare?
- Comparison of Regression 3 and 4: (one short paragraph recommended)
  - What are the *full interpretations* of the specific numbers – intercept, slope, and R^2 – for Regression 4?
  - Making specific reference to the relevant graphs and regressions, how do Regression 3 and 4 compare?
- Discussion of Regressions 5 – 8: (one paragraph recommended)
  - Overall, what do Regressions 5 – 8 and the associated graphs show?
  - What conclusions should be drawn from Regressions 5 – 8?
- Critical assessment: (one paragraph recommended)
  - The original report extensively uses the coefficient of correlation as a descriptive statistic, including for the cases in the previous two pages. Given the analyses presented in this test, what are TWO independent reasons to be concerned about using the coefficient of correlation as a *descriptive summary statistic* of the relationship between intergenerational income mobility and features of communities (CZ's) across the U.S.?
  - Overall, how informative are the presented analyses for the implied research questions? Explain.

Guide to answering effectively: (these are exactly the same as the first page, repeated here for emphasis)

- Directly and fully address the question. A complete answer addresses each supporting question.
- To answer the questions, *apply* relevant course concepts to the quantitative results given.
- Critically evaluate the results: assess any violations of underlying assumptions and limitations of the data.
- Write your answer using *complete sentences and well-organized paragraphs*.
  - A typical paragraph is 4 to 6 sentences. A short paragraph, which is at times suggested, is 2 to 4 sentences.
  - Answer with multiple well-structured paragraphs that coherently answer the main question. Each paragraphs should have a clear topic sentence followed by supporting points.
  - Your response should be about 2 pages (handwritten with average handwriting).
  - Spend about 55 minutes planning/outlining your arguments and about 55 minutes writing and revising.
  - Use the scrap pages for a draft/outline.
  - Revise and edit your draft to achieve a clear, coherent, and concise writing style.

Write your final answer on these pages to be graded. Use scrap pages for a draft/outline that is not graded.