Instructions:

• IDENTIFY YOURSELF BY CODE LETTER, not name, on all pages.

• Start each question at the top of a new page.

• Answer FOUR (4) out of the following SIX (6) questions. If you answer more than four questions, only the first four will be graded.

• Note that some questions continue onto a second page. Be sure to answer all parts of the question.

• You have four hours to complete the examination.

• This is a closed book exam. No notes, articles, books or other sources may be used.
**Question 1**

a) Compare, contrast and evaluate the use of audits, correspondence studies and field experiments vs. Blinder-Oaxaca type analysis of outcome data in measuring and detecting discrimination. Distinguish, in your answer, between attempts to measure market discrimination (e.g. rental housing, employment, or lending) vs nonmarket discrimination (child maltreatment reporting, traffic stops, or school suspensions).

b) Provide a formal representation of the pathways to racial/ethnic disparities in traffic stops. Distinguish between disparities in outcomes and disproportionalities in outcomes and discuss how or whether audit methods for measuring and detecting disparities and disproportionalities differ along the pathways. Detail some of the biases that arise when measuring disparities vs disproportionalities along these pathways.

c) Consider the following field experiment:

Six disabled and six non-disabled testers each visit six auto repair shops. Each disabled driver is matched with a non-disabled driver. The disabled driver’s body-damaged automobile is driven to the repair shop to get an estimate of repairs by the disabled driver. The non-disabled driver drives the same automobile to another repair shop.

The designers of the experiment identified 12 accessible body shops in a specific city. An accessible body shop is one that a person with a physical disability is able to enter and exit without difficulty.

The body shops are randomly split between the disabled and the non-disabled tester in each pair; both testers in any given pair approached body shops with the identical car. Two testers are sent to each body shop, using a different car.

The non-disabled (disabled) tester in pair i visited body shops that were previously visited by the disabled (non-disabled) tester of pair j, where i ≠ j.

The testers obtained a written estimate (without bargaining or negotiation) for the repair.

i. How many repair estimates are the result of the experiment (assuming no rejections or refusals?)

ii. Explain how unobserved heterogeneity among repair shops is accounted for in this experiment.

iii. Discuss wherein you agree or disagree with the experimental design that requires testers to seek a repair estimate without bargaining or negotiation.

iv. The summary results of the main test are the following:
where the information in the last column denotes the mean estimate and the standard error of the estimates across body shops and testers.

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>$1212</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Disabled</td>
<td></td>
<td>se=$214</td>
</tr>
<tr>
<td>Disabled</td>
<td></td>
<td>$1425</td>
</tr>
<tr>
<td></td>
<td></td>
<td>se=$212</td>
</tr>
</tbody>
</table>

v. Discuss how one might determine whether there is discrimination against disabled price seekers in this experiment.

vi. How does this type of field experiment differ from correspondence studies and visual vignette methods?
Question 2

The following is an excerpt from a recent article in *Bloomberg BusinessWeek* about a proposal to reduce obesity with a new state tax on soda:

“…Legislation before the state's General Assembly [legislature] would add one penny per ounce tax to the cost of regular soda or other drinks sweetened with sugar, or $1.44 to the cost of a 12-pack of 12-ounce sodas. Diet sodas or other non-sugared drinks would not be subject to the tax.

The projected $45 million in revenue from the tax would go toward public health efforts to reduce obesity…”

Imagine that you are an analyst for the state legislature, and you have been asked to evaluate the merits of this proposal. In your evaluation, answer the questions below. Include any models or literature you find relevant. Your evaluation should demonstrate an advanced graduate level of economics knowledge.

a) Can economic principles justify government intervention to reduce obesity? Under what circumstances?

b) Legislators want to know whether the proposed tax is likely to be effective in reducing obesity. What information would you need to gather and what approach would you use to make that determination?

c) Legislators also want to know whether the proposal represents sound tax policy.

   c.1 What does it mean to say that a tax is economically efficient?
   c.2 Why is economic efficiency a goal of tax policy?
   c.3 Do you think the proposed tax would be efficient?
   c.4 Why is the distributional burden of a tax a consideration in tax policy?
   c.5 What information would you need to gather and what approach would you take to determine how this tax is likely be distributed among taxpayers?
   c.6 What tax compliance issues (if any) are likely to be raised by the proposal?
   c.7 What information would you need to gather and what approach would you take to predict the likely rate of compliance with this tax?

d) What else would you consider in evaluating whether the proposal represents good state policy?

---

Question 3

The earliest version of the Blinder-Oaxaca decomposition of, say, ln-earnings disparities between males and females, can be given by the following set of equations:

\[
\ln y^F = \sum \beta_i^F x_i^F + \varepsilon^F \\
\ln y^M = \sum \beta_i^M x_i^M + \varepsilon^M
\]

Where superscripts \(F\) and \(M\) denote females and males, where the \(x\)’s are human capital (and other) predictors of ln-earnings (\(\ln y\)) and the \(\beta\)’s are coefficients to be estimated.

Ordinary Least Squares estimates produce the following results:

\[
\ln \tilde{y}^F = \sum \hat{\beta}_i^F \tilde{x}_i^F \\
\ln \tilde{y}^M = \sum \hat{\beta}_i^M \tilde{x}_i^M
\]

The difference or gap in ln-earnings can thereby be decomposed into an explained portion and an unexplained portion:

\[
\ln \tilde{y}^M - \ln \tilde{y}^F = \sum \hat{\beta}_i^M \tilde{x}_i^M - \sum \hat{\beta}_i^F \tilde{x}_i^F
\]

Or:

\[
\ln \tilde{y}^M - \ln \tilde{y}^F = \sum \hat{\beta}_i^M \tilde{x}_i^M - \sum \hat{\beta}_i^F \tilde{x}_i^F + \sum \hat{\beta}_i^M \tilde{x}_i^F - \sum \hat{\beta}_i^M \tilde{x}_i^F \\
= \sum (\hat{\beta}_i^M - \hat{\beta}_i^F + \hat{\beta}_i^M - \hat{\beta}_i^F) \tilde{x}_i^F
\]

a) What portion of the gap is explained?
b) What portion of the gap is unexplained?
c) What assumptions must be made in order for the resulting decomposition above to be valid?
d) The unexplained portion in the above illustration is often termed “discrimination.” Under what conditions will the unexplained portion overestimate market discrimination? Under what conditions will the unexplained portion underestimate discrimination?
e) Provide a detailed expansion of the above decomposition to account for what is known as the “Cotton Correction” and provide a rationale for this alternative method of measuring the explained and unexplained portions of the gap.
f) Discuss the implications for the measurement of discrimination using the model detailed above when instead of two groups, there are multiple groups (e.g. whites, blacks, Hispanics, American Indians and Asians.)

g) Detail how one can compute the portion of the explained gap that is explained by, say, $x_k$, the kth predictor of ln-earnings.

h) Suppose that $x_k$ is a dichotomous variable (e.g., disabled). Does the portion of the explained gap explained by $x_k$ depend on which value is the excluded value (e.g., disabled vs non-disabled)? Why or why not?

i) Suppose that the dependent variable is dichotomous (e.g. labor force participation). How is the measurement of the portion of the explained gap explained by the dichotomous variable, $x_k$ (e.g., disabled) affected?
Question 4

a) What are the theoretical underpinnings of Benefit-Cost Analysis and what are the challenges in implementing a theoretically sound BCA? What are some of the methods used to measure non-market and/or intangible benefits? Discuss the assumptions required to justify including these benefits in the calculation of Benefit-Cost Ratios.

b) Provide an illustration from the literature on the benefits of early childhood education and compare estimates that are limited to improved labor outcomes vs those that incorporate social benefits from reductions in crime and welfare dependency.

c) The following quotation comes from Kenneth Arrow, “The Social Discount Rate,” in Cost-Benefit Analysis of Manpower Policies (1969)

“In the social evaluation of a competitive economy, it is customary and proper to start by using market prices as an estimate of social costs. [but]...the future is risky, and existing risk-bearing markets are not, in principle, sufficiently complex and differentiated for optimal allocation” (p.56)

Discuss the implications of this comment for the use of the market interest rates for discounting future wages and/or reductions in crime/welfare expenditures arising from early childhood education programs.
Question 5

The appropriate role for Government in providing health care is an item of considerable debate. Some believe that the free market model with no government intervention is appropriate. Others argue that a substantial government presence is necessary. And, of course, there is disagreement over the form that intervention should take and how it should be paid for.

a) What is the economic argument for the public sector intervention in the provision of medical services? What, if any, is the market failure? Compare the public sector’s role if there is a market failure to that if health care is considered to be either a merit good or a right.

b) The Affordable Care Act requires that all individuals not covered by an employer sponsored insurance plan, Medicare or Medicaid purchase a private insurance policy or pay a penalty. Is mandatory universal coverage necessary for economic efficiency? Explain.

c) A portion of the additional revenue required to pay for the Affordable Care Act will come from a 2.3 percent excise tax on medical devices. Minnesota Congressman Paulsen claims that the tax, when added to the existing corporate income tax, will cause our medical device industry to face one of the highest tax rates of any industry in the world and will cause sales of devices to fall substantially. Examine the medical device tax from an economic efficiency point of view comparing its dead-weight loss to that of an increase in the payroll tax which raises the same revenue.

d) The Affordable Care Act will also receive funding from a 40 percent excise tax on so called “Cadillac” employer based health insurance coverage whose premiums exceed $27,500 per year for a family of four. An alternative to that excise tax on premiums would be to include employer paid health insurance premiums as part of gross income, then allow an itemized deduction capped at $27,500 for health insurance premiums paid. Discuss the difference in the distributional impact of the two alternatives. Would the alternative make the tax system more progressive than the excise tax?

e) Another portion of the revenue needed to pay for the Affordable Care Act will come from a 10 percent federal excise tax on indoor tanning services. Evaluate this tax from an economic efficiency point of view.
**Question 6**

a) Numerous innovative strategies have been offered for reducing racial gaps in test scores. Some of these innovations or interventions are informed by results such as those found in Fryer and Levitt, “Understanding the Black-White Test Score Gap in the First Two Years of School,” *The Review of Economics and Statistics*, May 2004, 86(2): 447–464. One innovation is subsidies to low-income families to increase the number of books that children have in their homes.

Explain the theoretical rationale for providing subsidies to low-income families to purchase books vs direct income supplements.

b) Table 2 (attached) from the Fryer-Levitt paper reports results of the estimation of an equation where the dependent variable is a normalized score on mathematics and reading examinations. The coefficients on the race variables denote the racial gaps in test scores relative to non-Hispanic whites. Various specifications are considered for both the estimation of the mathematics score and the estimation of the reading scores. Columns 4 and 9 control for such factors as gender, age, birth weight, mother’s age at time of birth, and WIC participation in addition to socio-economic status and numbers of books in the home (which the authors state measures home environment.) Columns 5 and 10 present still another specification. The key differences between columns 4 and 5 for mathematics and columns 9 and 10 for reading arise from the inclusion of more than 98 other covariates in the specifications reported in columns 5 and 10 as compared to columns 4 and 9. These additional covariates capture: city size, neighborhood characteristics, region of the country, parental education, parental income, parental occupational status, family size and structure, whether the mother worked, type of preschool program participation, whether English is spoken at home, and the extent of parental involvement in a child’s life and school.2

i. Carefully review Table 2 and discuss what can and what cannot be inferred about the prospects of improving mathematics and reading test scores by increasing the number of books available in the home.

ii. In particular, does inclusion of neighborhood characteristics, preschool program participation, parental involvement, parent’s occupation or other variables not included in Columns 4 and 9 alter conclusions about the availability of books?

---

2 The full text of the footnote to Table 2 reads: NOTES: The dependent variable is the math or reading test score in the fall of kindergarten. Test scores are IRT scores, normalized to have a mean of 0 and a standard deviation of 1 in the full, unweighted sample. Non-Hispanic whites are the omitted race category, so all of the race coefficients are gaps relative to that group. The unit of observation is a student. Standard errors in parentheses. Estimation is done using weighted least squares, using sample weights provided in the data set. In addition to the variables included in the table, indicator variables for students with missing values on each covariate are also included in the regressions. In addition, columns 5 and 10 report only a subset of the coefficients from regressions with 98 covariates included in the specification. The full results for columns 5 and 10 are reported in appendix table A1. Note that the specifications in columns 5 and 10 include age and age squared; that is why the coefficient on age changes so dramatically relative to other columns in the table.
iii. What, if anything, can be concluded about the impacts of availability of books on racial gaps in test scores?

iv. Detail the strengths and weaknesses of using nearest neighbor propensity score matching vs Heckman’s differences-in-differences matching to address the problem that the data are not drawn from a randomized experiment.

v. Discuss the nature of selection that might be involved with regards to the availability of books in the home. Detail how one might control for selection effects when measuring the impacts of book availability on test scores and distinguish between Instrumental Variables (IV) and Heckman Selection Correction methods for controlling for such selection.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Math</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Black</td>
<td>-0.638</td>
<td>-0.368</td>
</tr>
<tr>
<td>(0.022)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.722</td>
<td>-0.429</td>
</tr>
<tr>
<td>(0.022)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>0.150</td>
<td>0.070</td>
</tr>
<tr>
<td>(0.056)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other race</td>
<td>-0.503</td>
<td>-0.329</td>
</tr>
<tr>
<td>(0.041)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socioeconomic status composite measure</td>
<td>0.456</td>
<td>0.389</td>
</tr>
<tr>
<td>Number of children’s books</td>
<td>-</td>
<td>(0.014)</td>
</tr>
<tr>
<td>(Number of children’s books)² (×1000)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Age at kindergarten fall (in months)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Birth weight (ounces) (× 10)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Teenage mother at time of first birth</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mother at least 30 at time of first birth</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WIC participant</td>
<td>-</td>
<td>(0.019)</td>
</tr>
<tr>
<td></td>
<td>0.108</td>
<td>0.223</td>
</tr>
</tbody>
</table>

| Number of observations | 13,290 | 12,601 |
| Full set of covariates included in regression? | N       | N       | N       | N       | Y       | N       | N       | N       | Y       |

**NOTES:** The dependent variable is the math or reading test score in the fall of kindergarten. Test scores are CRT scores, normalized to have a mean of 0 and a standard deviation of 1 in the full, unweighted sample. Non-Hispanic whites are the omitted race category, so all of the race coefficients are gaps relative to that group. The unit of observation is a student. Standard errors in parentheses. Estimation is done using weighted least squares, using sample weights provided in the data set. In addition to the variables included in the table, indicator variables for students with missing values on each covariate are also included in the regressions. In addition, columns 5 and 10 report only a subset of the coefficients from regressions with 98 covariates included in the specification. The full results for columns 5 and 10 are reported in appendix table A1. Note that the specifications in column 5 and 10 include age and age squared; that is why the coefficient on age changes so dramatically relative to other columns in the table.