Advanced Research Design for Causal Inference
Winter 2019

INSTRUCTOR:
Robbie Moon, PhD
Assistant Professor
Scheller College of Business
Georgia Institute of Technology
Cell: +1-678-896-2552
Email: robbie.moon@scheller.gatech.edu

COURSE OBJECTIVES:
This course is designed for students interested in performing archival-based academic research. The purpose of this course is to help students understand, design, and ultimately to implement many common (e.g., linear regression) and advanced (e.g., natural language processing, machine learning) analytical techniques used by archival researchers and apply these methods to a variety of data (e.g., financial statements, text, social media, etc.). We will focus on three main themes: (1) the types of questions accounting researchers ask, (2) the data sources (both structured and unstructured, free and subscription-based) that can be used to answer these questions, and (3) the methods archival researchers use to answer each question. With respect to (3), we will discuss how to minimize the likelihood that documented statistical associations are not endogenous (i.e., to help in establishing a causal link in archival studies). In addition, you will be exposed to some programming techniques that can be used to collect and analyze data from nontraditional sources, including web scraping and textual analysis.

Stated formally, the course objectives are:

- Help students understand the basics of conducting accounting archival research including: major areas of research, basic design choices, data analysis techniques, and common pitfalls in archival research.
- Improve student’s abilities to lead and participate in research seminar discussions.
- Discuss ethical issues in archival research.
- Be familiar with data collection techniques, particularly related to unstructured data.

EXPECTATIONS:
I assume no significant prior knowledge of archival accounting research. For each assigned paper, thoroughly read the abstract and introduction to understand the paper’s primary motivation, prediction, results, and contribution (how the paper alters readers’ prior beliefs). You may skim through the background/literature review/hypothesis development sections (if applicable). Then, closely read the methodology section to try to understand what the authors attempt to do methodologically (or how they reach the conclusions discussed in the introduction). In the beginning, you might have to search Google to get a feel for what some of the terms mean, but hopefully as time goes by, you’ll become more familiar with some of the lingo and technical language.

There will be a set of questions to go with each reading assignment. Hopefully, the questions will help you identify the key aspects of each paper.

What will you do prior to coming to each class session?
• Read the assigned paper. (Copies of all papers are available in the Dropbox folder.)
• Answer a set of questions using the response template posted on Dropbox. The first six questions are always the same:
  1. What is the paper’s primary research question?
  2. In the eyes of the authors, why is this paper important?
  4. What kind of data is this paper using (standard datasets, textual data, unique data, etc.)?
  5. What is one concern you have with the study?
  6. What new idea does this study spark in your mind?
** Any remaining questions will relate to the specific tool being highlighted that week.
• You are encouraged to use any resource at your disposal to answer the questions (Google, classmates, textbooks, other research papers, etc.). However, each student must provide answers to the questions written in their own words.

What will we do during class?
• This class follows the seminar method, which is common in PhD programs. Each session will focus on one or two papers. I will select papers that expose you to a variety of (1) topics in financial accounting, auditing, tax, and capital markets settings, (2) data sources (e.g., traditional financial statement and stock return data, narrative disclosures, and social media posts), and (3) analysis techniques.
• Active student participation is critical to the success of any academic seminar. Attendance and being both attentive and participative are necessary to get the most out of this class and your graduate education.
• During class, we will work through all of the pre-class questions. To assist in this process, one student will be assigned to start the discussion on each question. These student assignments will be random.

SCHEDULE:
The schedule on pp. 3-4 outlines reading assignments for each meeting. Please watch your e-mail for changes that may arise as we go forward.
**Tentative Schedule**

*(Note: Schedule is subject to change, and classes 2 & 3 and 7 & 8 will be held on the same day; full paper citations can be found following this schedule and papers will be provided on Dropbox)*

<table>
<thead>
<tr>
<th>Day</th>
<th>Topic and reading assignment</th>
<th>Other assignments and information</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Course introduction:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Read syllabus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Lecture introducing Empirical Research and “The Experimental Ideal”</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Introduction to linear and non-linear (logistic) models [Financial Reporting]:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Call, Campbell, Dhaliwal, and Moon (JAE 2017)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Whited, Swanquist, Shipman, and Moon (WP 2019)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Fixed effects [Audit] &amp; Event Studies [Tax]:</td>
<td>Optional:</td>
</tr>
<tr>
<td></td>
<td>• Moon, Shipman, Swanquist, and Whited (CAR 2019)</td>
<td>• Joint Hypothesis (f-test) Explanation</td>
</tr>
<tr>
<td></td>
<td>• Wagner, Zeckhauser, and Ziegler (JFE 2018)</td>
<td>• *This paper talks about CARs versus BHARs: Barber and Lyon (JFE 1997).*pdf</td>
</tr>
<tr>
<td></td>
<td>Optional:</td>
<td>• *This paper talks about event days windows and whet Berkman and Truong (JAR 2008).*pdf</td>
</tr>
<tr>
<td>4</td>
<td>Interactive effects [Financial Reporting, Capital Markets]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Lafond and Watts (TAR 2008)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Khan and Watts (JAE 2009)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Natural Experiments &amp; Difference-in-Differences Analysis [Capital Markets]:</td>
<td>Optional:</td>
</tr>
<tr>
<td></td>
<td>• Brown, Stice, and White (JAR 2015)</td>
<td>• Bertomeu, Beyer, and Taylor (2015)</td>
</tr>
<tr>
<td></td>
<td>• Blankespoor, deHaan, Wertz, and Zhu (JAR 2019)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Propensity-score Matching [Audit]:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Minutti-Meza (JAR 2013).pdf</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Shipman, Swanquist, and Whited (TAR 2017)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Introduction to Textual Analysis [Capital Markets &amp; Financial Reporting]</td>
<td>Optional:</td>
</tr>
<tr>
<td></td>
<td>• Li (JAE 2008)</td>
<td>• Loughran and McDonald (JAR 2016)</td>
</tr>
<tr>
<td></td>
<td>• Loughran and McDonald (JF 2011)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Machine Learning Applications [Capital Markets, Auditing]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Frankel, Jennings, and Lee (JAE 2016)</td>
<td></td>
</tr>
</tbody>
</table>
| 9 | Social Media and New Disclosure Venues [Capital Markets, Voluntary Disclosure]  
- Hales, Moon, and Swenson (AOS 2018)  
- Campbell, DeAngelis, and Moon (RAST 2019) | Optional:  
- Blankespoor (AOS 2018) |
|---|---|---|
| 10 | Final Class Assignments:  
This last class session will work a little differently.  
Each student will be assigned one working paper from the list given. For that working paper, students should read the paper closely and prepare responses to the following requirements:  
1. Briefly explain the research question of the assigned paper, the research design (tools used, data, etc.), the key results, and the contribution.  
2. Identify concerns the author(s) identify (if any) in the paper and explain how they address them in their design.  
3. If applicable, discuss tools they use that we did not cover in class (e.g., regression discontinuity designs, hazard model, etc.). Highlight any concerns you still have about the paper’s conclusions.  
4. Do one of the following:  
   a. Explain why the research tool chosen by the author was best for that paper, OR  
   b. Propose an alternative method that could be used to answer this question.  
5. (Optional) Propose a follow-up research question that you thought of based on this paper.  

In order to answer these questions, you are encouraged to do some outside reading and to come see me with questions. Often it helps to read some of the papers the authors cite as motivation for their study.  

In addition to a written assignment, you will each be given 10-15 minutes to talk a little bit about each paper and teach the tool to the class. You do not have to cover all 5 points above (but you can), but make sure everyone understands both the purpose of the paper and any nuances of the design that make it interesting. You do not have to bring PowerPoint slides, but are welcome to do so if you think it would help you or the class. | Working paper list to be determined at a later date. |
References (all papers will be provided in a Dropbox folder)


