Writing in the discipline and reproducible methods: A process-oriented approach to teaching empirical undergraduate economics research

Emily C. Marshall [a]
Anthony Underwood [b]

13 February 2018

(manuscript under review)

Abstract

This paper describes an empirical research project in an upper-level undergraduate economics writing-in-the-discipline course that aims to reduce the high fixed costs associated with designing an empirical research project assignment and encourage more undergraduate economics research. This project has a dual-purpose: to teach students economics writing conventions and reproducible empirical research methods. We present a sequenced project design and replication documentation protocol and posit that requiring students to produce this comprehensive documentation promotes student learning and leads to improvements in organization and coherence throughout the entire research and writing process. Through effective writing, data management, and empirical analysis students learn to *do* econometrics. The project learning objectives, workflow, evaluation criterion, and replication protocol are outlined in detail, along with the pedagogical rationale for each component.

JEL Codes: A22, B4, C8
Keywords: writing in the discipline, reproducibility, replication, teaching economics

[a] Corresponding author. Assistant Professor. Economics Department, Dickinson College, Carlisle, PA 17013-2896. marshaem@dickinson.edu.
[b] Assistant Professor. Economics Department, Dickinson College, Carlisle, PA 17013-2896. underwoa@dickinson.edu

1. Introduction

Research, writing, and empirical skills have become increasingly important in undergraduate education. Recent surveys conducted by the Association of American Colleges and Universities show that "employers want to hire college graduates who can write coherently, think creatively, and analyze quantitative data" (Selingo 2017). There is a broad consensus among economists that enabling students to understand how and when to "think like an economist" is the primary goal of an undergraduate education in economics (Allgood, Walsted, and Siegfried 2015). However, most economists have no formal training in undergraduate writing instruction, despite ample evidence regarding the benefits of teaching writing (Dynan and Cate 2008; Greenlaw 2003; Simpson and Carrol 1999). Based on a survey of economics alumni, Simpson and Carroll (1999) conclude that short papers and analyses of class readings are best at preparing students for graduate programs and future occupations, while quantitative research papers are the best way to enhance student learning of economics. Dynan and Cate (2008) and Greenlaw (2003) provide evidence that writing assignments improve student learning.

Despite this evidence, McGoldrick (2008a) finds that 30% of economics departments have no formal writing requirement for the economics major. Of the institutions that do have a formal writing requirement, the most common type is a "writing-in-the-discipline" (WID) course or senior capstone with a writing component. Additionally, only 15% of economics departments reported offering courses dedicated to the research process, while fewer still (10%) offered a course that was specifically designated as "research methods" (McGoldrick 2008a).

Given the lack of formal training in writing and research instruction, the prospect of designing a course with a large empirical research project or augmenting an existing course with such an assignment is daunting. This paper aims to reduce these high fixed costs and encourage more undergraduate economics research by outlining an empirical research project that can be applied in an upper-level undergraduate economics course. The project serves the dual purpose of teaching disciplinary-specific writing conventions and reproducible empirical research methods in economics.

In developing this project, we recognize both the increasingly empirical nature of the economics discipline (Angrist et al. 2017; Hamermesh 2013) and the increasing popularity of econometrics. In 1980, only about 6% of undergraduate institutions surveyed required econometrics and less than half offered at least one econometrics course (Siegfried and Wilkinson 1982). However, as of 2010, about 81% of all undergraduate economics programs offered at least one econometrics course (Johnson, Perry, and Petkus 2012). Recent estimates show that approximately 50% of institutions require an econometrics course, while only one-fifth of undergraduate economics programs offer no econometrics courses (Siegfried and Walstad 2014). Given these trends, this type of empirical research project should be part of the undergraduate economics curriculum. Even though econometrics has become an increasingly important part of the economics major, few instructors have incorporated a research assignment as part of the learning experience. As of 2010, only 10% of the national universities and 21% of liberal arts

colleges that offer econometrics require an independent empirical research assignment (Johnson et al. 2012).

The purpose of this empirical research project design is to exploit complementarities in satisfying the learning objectives of a WID course and of reproducible empirical research. These goals are mutually beneficial: the meticulous documentation and planning required for data analysis will yield increasingly well-developed empirical research papers. To that end, the purpose of the project is for students to learn how to *do* econometrics through effective writing, data management, and empirical analysis. Greenlaw (2006) suggests that "the best way to learn economics is not to hear about it, or to read about it, but to do it" and that "doing economics means performing economic research." As noted above, economic research has become increasingly empirical and according to the standards of the American Statistical Association, authors of research involving statistical analysis have an ethical responsibility to "promote sharing of data and methods" and "make documentation suitable for replicate analyses" available (ASA 2016). Reproducibility requires documentation that allows an independent researcher to reproduce every step of the data management and analysis process and replicate the results presented in the study. However, many economists, even those conducting their own empirical research, have no formal training in this type of data workflow for reproducibility. Historically, the economics discipline has paid little attention to replicability of empirical research, and evidence on the reproducibility of published papers confirms this inattention (Chang and Li 2015). As instructors, we have the ability to set the standards and incentives that guide the work of our students, emphasizing replicability and documentation in empirical research.

One of the largest fixed costs for instructors incorporating an empirical research project into their courses is the time required to develop the course materials and tools necessary for a successful learning experience. This paper presents an adaptable project template complete with pedagogical rationale in an effort to significantly reduce these costs and facilitate inclusion of empirical research projects and transparent data methods in economics courses. Importantly, this innovation improves student learning of the course content, exposes students to effective research design, and encourages reproducibility (Dynan and Cate 2008; Greenlaw 2003; McGoldrick 2008b; Ball and Medeiros 2012). These resources invested in undergraduate research produce experiential learning that students carry with them throughout their lives and yield human capital accumulation valuable to both graduate programs and potential employers (Hoyt and McGoldrick 2017a, 2017b). In the following section, we discuss the project structure and learning objectives. Section 3 briefly describes the course history and context. Section 4 develops the pedagogical rationale for the project. The evaluation and assessment of the project are discussed in Section 5. Finally, Section 6 concludes with ongoing challenges and suggestions for the promotion of undergraduate empirical research in economics.

2. Project Structure and Learning Objectives

Research has shown that the best way to learn economics is by *doing* it (Greenlaw 2006). Learning econometrics and empirical data techniques means doing economic research (Greenlaw

2006). This project is designed to emphasize the process of empirical research, or what McGoldrick (2008b) calls the iterative research process. The accompanying replication protocol requires students to experience the cycle of data analysis: plan, collect, organize, compute, and document. This cycle mirrors the recursive writing process in many respects: pre-writing, drafting, revising, and editing. When drafting and revising, authors often need to return to pre-writing and planning to develop and expand their ideas, analogous to the situation where econometricians must often revisit the collection or organization phase of data analysis. We argue that the emphasis placed on reproducibility and documentation, combined with the sequenced nature of the project, produces higher quality research papers, significantly enhances learning of the course material, and improves student writing.

The empirical project prompt (included in Sections A and B of the appendix) outlines the learning goals, project sequence, and dual procedure workflow, juxtaposing the writing and data management tasks. In addition, we have developed a detailed series of prompts and rubrics for each of the seven project components, also included in the appendix (Section C). Partitioning the replication documentation and writing tasks into progressive assignments is an important component of student learning and process writing. As discussed in more detail in Section 4.1, this sequencing, combined with the replication requirements, slows students down by facilitating intentionality and reflection. Research has shown that detailed prompts, rubrics, and sequenced writing assignments improve student writing and learning (Ambrose et al. 2010; Dynan and Cate 2008; Greenlaw 2003; Reddy and Andrade 2010).

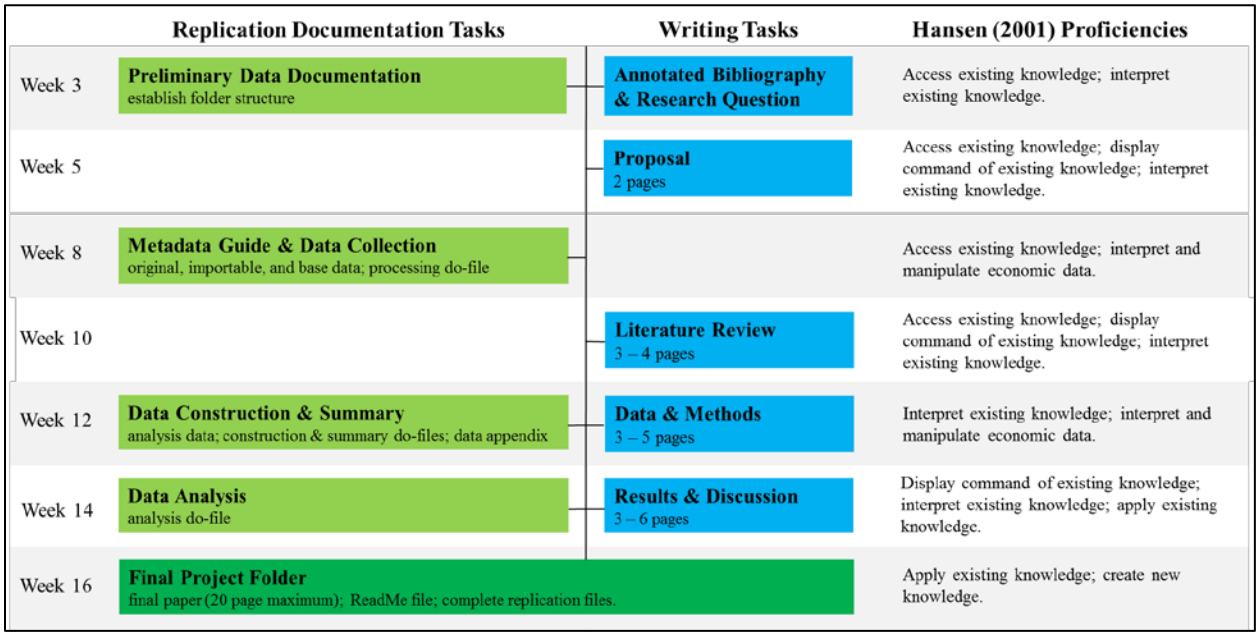
This empirical research project encourages students to develop proficiencies consistent with Hansen (2001): access existing knowledge; display command of existing knowledge; interpret existing knowledge; interpret and manipulate economic data; apply existing knowledge; and create new knowledge. These proficiencies, developed by Hansen (1986, 2001) are routinely used as a template for curriculum, course, and project design (Allgood et al., 2015; Klein, 2013; Li and Simonson, 2016; McGoldrick, 2008b). The project phases, workflow, descriptions, and how they align with Hansen's (2001) proficiencies are illustrated in Figure 1.

Students begin the semester by exploring research topics and creating a replication documentation folder using a file-sharing platform, such as Microsoft OneDrive.[1] This main folder includes a hierarchy of folders, illustrated in Figure 3 and discussed in more detail in Section 4.2, that serve as a repository for writing and data management tasks over the course of the semester. This folder structure provides a straightforward vehicle for both assignment submission, instructor feedback, and replication. By week 3, students submit an annotated bibliography and preliminary research question. Note that students are asked to produce the annotated bibliography prior to submission of the research proposal in week 5. Previously, when students were required to submit proposals prior to finding and evaluating the existing literature, the product was often under-developed, unoriginal, and/or unrealistic. By reading, analyzing, and thinking critically about the literature prior to submitting the proposal, students learn to engage in a scholarly discussion and

[1] This is a highly adaptable feature of the project. In our experiences, we have utilized both email and Microsoft OneDrive. However, many other platforms could be used such as Dropbox, Google Drive, and Open Science Framework.

produce higher quality proposals. After receiving feedback on initial sources and research questions, students write a preliminary thesis statement in their proposal.

Figure 1. Project Structure



The most difficult and time-consuming component of the project tends to be the "Metadata Guide and Data Collection" due in week 8. This subsection requires students to combine several new component skills: finding data, processing data, and learning statistical software syntax (in our case, Stata). These opportunities for deliberate practice are crucial for student learning (Ericsson, Krampe, and Tescher-Romer 1993; Ericsson and Charness 1994; Ericsson and Lehmann 1996).[2] At this phase, students submit their (1) original data files (in unmodified form), (2) importable data files (minimally altered in order to be read by the statistical software), (3) base data file (master data file combining data from all original sources), (4) processing command (.do) file (file including all commands used to produce the base data file), and (5) metadata guide with an entry explaining how to obtain and understand each original data file.

After submitting this section, students focus their attention on a literature review, the objective of which is to motivate their research question, place it in the context of previous work, and provide a foundation for their analysis. The "Metadata Guide and Data Collection" and "Literature Review" are due in isolation due to the rigorous nature of these two project components in order to optimize learning. In the literature review, students are required to update their thesis statement. We use this as an opportunity to show students the evolution of their thesis statement

[2] Along with the project sequencing, a series of in-class workshops complement the project components including topics such as finding sources, finding data, and cleaning data. In addition, students are assigned homework to complete outside of class after the workshop as a means of providing a bridge to their project tasks. These resources are available upon request. Also, see http://geocenter.github.io/StataTraining/ and https://www.aeaweb.org/about-aea/committees/economic-education/econometrics-training-modules/one for published modules and course materials.

since the proposal phase of the project, as they have incorporated realizations based on data availability and the existing literature. This is exactly the type of intentionality and reflection promoted through this project sequencing. Writing and research become a process of discovery.

Over the last month of the semester, students complete the final components of an empirical research paper, working on writing and data management tasks in tandem. At week 12, students submit the "Data Construction and Summary" files and the "Data and Methods" section of the paper. The former includes the (1) analysis data (.dta) file (used to produce the final estimates in the paper), (2) construction command (.do) file (file containing all commands used to produce the analysis data from the base data), (3) data appendix summarizing the analysis data, and (4) summary command (.do) file (file including all commands used to produce the data appendix). The Data and Methods document describes the conceptual or theoretical framework, data, and model specification.

At week 14, students submit the (1) analysis command (.do) file (file containing all commands used to produce final results tables included in the paper) and (2) results and discussion section of the paper. Finally, in the last week of the semester (final exam week, in our case), students submit the (1) final empirical research paper, incorporating all feedback from previous submissions and (2) final replication documentation folder complete with a ReadMe file explaining the contents of the replication folder and providing step-by-step replication instructions. Please see Appendix Sections A, B, and C for more details on project requirements and sequencing.

3. Course History and Context

This project has the potential for implementation in a variety of courses at the undergraduate level. Some courses are natural environments for teaching writing and research methods. The most obvious outlet for this type of project is an upper-level econometrics course. For some institutions, this would be the second course in an econometrics sequence. For others, this would be the first course in an econometrics sequence, but at an upper-level.[3] However, this project could be implemented in a variety of courses.

Another natural environment for this empirical research is a senior seminar or capstone experience. As of 2013, about half of economics programs require a senior seminar or capstone (Siegfried and Walstad 2014). In the same vein, this type of project could be adapted to guide honors or senior thesis writing, both of which have become more popular in recent years. In 1980, one-third of institutions offered honors in economics (Siegfried and Wilkinson 1982) and only about 7% of major programs required a senior thesis. As of 2013, these numbers have risen to 46% and 18%, respectively. Finally, this project could be used in any upper-level economics course with a focus on data analytics and WID.[4] According to the National Census of Writing, 39% of

[3] Currently, at Dickinson College, we implement this project in a 300-level econometrics course.

[4] For example, this project is a major component of a WID course at Dickinson College, which specifies that students will be able to "identify and demonstrate discipline-specific writing conventions" and "understand that writing is a recursive process and develop an effective writing process" (Dickinson College. "Writing-in-the-Discipline Courses." The Writing Culture. http://www.dickinson.edu/info/20158/writing_program/870/the_writing_culture/3). While these learning goals are college-specific, many institutions have similar writing-based courses.

the four-year colleges and universities sampled reported having WID and 62% require all students to take writing-intensive courses (outside of the English or Writing department).

We recognize that this project will not be feasible in its current form at all institutions due to class size limitations. However, below are several suggestions for large classes that would allow implementation of the project, even with a relatively sizable number of students without significant burden on the instructor. First, the most obvious adaptation of this assignment is to reduce grading by assigning group projects. In this case, the remainder of the project structure would be unchanged. We discuss additional strategies for group project implementation in Section D of the Appendix, including formation of groups, techniques to help mitigate free-rider problems, and methods of evaluating individual member contributions. Second, instructors in large classes may wish to shorten the assignment length. A clear variation of this project would be an abbreviated version of each component, with the end goal of constructing an economic note, such as those published in *Economics Letters.* The economics discipline needs to evolve with the rapidly changing educational environment, importance of writing skills, and data analysis in the field. The following section develops the pedagogical rationale for an empirical project supporting this initiative.

4. Pedagogical Rationale

To achieve complementarities in learning outcomes across writing in the discipline and reproducible data analysis, we utilize "process writing" (Schmeiser 2017) or "iterative research" (McGoldrick 2008b), a method of sequencing the project components and an adaptation of the Teaching Integrity in Empirical Research (TIER) protocol (Ball and Medeiros 2012).[5] At the beginning of the semester, we emphasize that a successful project requires a parallel workflow between data management and the writing process. In doing so, the goal is that requiring students to produce comprehensive replication documentation leads to improvements in organization and coherence throughout the entire research and writing process. Improved learning through *doing* economics.

Siegfried et al. (1991) claim that "writing clearly is the acid test of thinking like an economist." Thus, the link between writing and knowledge in the discipline is "doing economics" (Carter 2007). Writing in the discipline "is founded on an integrative relationship between writing and knowing. Its roots lie in rhetoric, in which invention has historically played a critical role in both recovering knowledge and generating new knowledge" (Carter 2007). Using the connections between knowledge, learning, and writing, Siegfried et al. (1991) highlight the potential to integrate these principles in teaching economics. Central to this pedagogy, a writing-to-learn approach, is the belief that instructors must not only devote time and energy to the assessment of students' finished papers but also to their writing processes, "the strategies and procedures followed in the act of writing" (Cohen and Spencer 1993). In this context, "writing is a tool of

[5] Project TIER developed the TIER protocol in an effort to advance the goals of research transparency and reproducibility through disseminating and teaching instructional practices. The guiding principle of this protocol is "that the documentation should allow an independent researcher to replicate every step of the data management and analysis and to generate the same results" (Ball and Medeiros 2012). For more information, please visit http://www.projecttier.org/.

discovery, a way of working through ideas you do not fully understand" (Greenlaw 2003). Below, we describe the pedagogical rationale for this empirical research project in two parts through a discussion of process writing and replication documentation.

4.1. Process Writing

When students are learning how to write, they are often focus solely on the final polished version of a paper, the *product*. This misplaced emphasis frequently leads to procrastination and inadequate drafts. Students regularly dismiss the writing *process* (sequencing, drafting, and revision) and treat research and writing linearly. Even if the assignment includes sequenced deadlines, without the appropriate ancillary materials, these are viewed only as speedbumps on the road to a final product. Development and implementation of this empirical research project helps to improve students' ability to treat research and writing as a recursive process, a tool of discovery. Good analytical writing should have an exploratory component – sharing its discovery process with the reader (Rossenwasser and Stephen 2012). The replication requirements of this project, along with its sequencing and transparency in assessment, encourage this sort of exploration and discovery.

Effective teaching of economics or design of an empirical research project can be viewed as a constrained optimization problem (Hultberg and Calonge 2017): maximize student learning and retention subject to cognitive load constraints and developmental barriers. In economics, as in most disciplines, "simple writing tasks are embedded in complex ones, within the whole process of assembling and communicating knowledge, and these embedded forms often constitute a developmental sequence" (Gottschalk and Hjortshoj 2004). Therefore, in designing an empirical research project, two concepts must be addressed: cognitive load and the developmental stages of mastery. For students to achieve mastery of any domain, they must acquire a set of component skills, practice them to the degree that they can be combined fluently with some degree of automaticity, and understand when and where to apply these skills appropriately (Ambrose et al. 2010). An empirical research paper requires students to combine several component skills, notably: finding data, managing data, utilizing statistical software, finding and evaluating existing literature, and writing technically. Consequently, the total information processing demands imposed by the project (the cognitive load) can often exceed what students can manage without direction. Once a student's cognitive load is exceeded, they are left with insufficient attention and cognitive resources to effectively complete a task (Ambrose et al. 2010; Hultberg and Calonge 2017). Our goal then, in project design, is to manage the cognitive load required of each project task to encourage long-term learning of the research and writing process, in addition to course content and material.

Interestingly, experts (instructors) do not suffer as much as novices (students) when it comes to performing and combining complex tasks. Due to extensive practice of the component skills required, they become more automated, demanding fewer cognitive resources and allowing performance of complex tasks with relative ease (Ambrose et al. 2010). Importantly, this is not because experts have more cognitive resources; rather, they can simply use what they have more

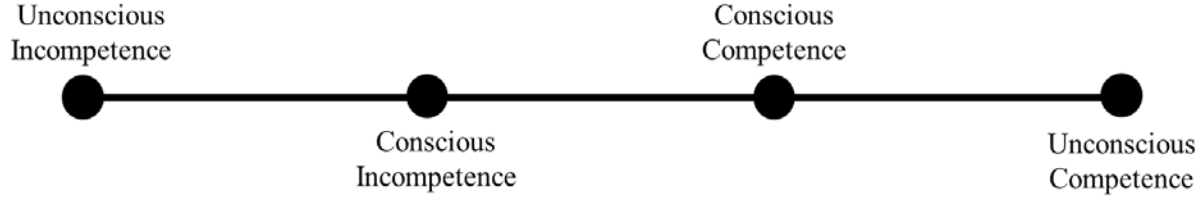
efficiently due to fluency. This difference in the cognitive load required of complex tasks is what often leads to frustration in the form of, "do I really have to put that in the prompt?!" As experts, instructors often develop expectations of autonomy that exceed the cognitive load and prior knowledge constraints of students. Often, this stems, in part, from an *expert blind spot*, where instructors are unaware of the learning needs of students (Nathan and Petrosino 2003), sometimes referred to as the "curse of knowledge" (Camerer, Loewenstein, and Weber 1989; Wieman 2007). As an example of how this may affect students, consider an analogy from Ambrose et al. (2010) of how a master chef may instruct novice cooks to:

> "sauté the vegetables until they are done," "cook until sauce is a good consistency," or "add spices to taste." Whereas such instructions are clear to chefs, they do not illuminate matters to students, who do not know what "done" entails, what a "good consistency" is, or what spices would create a desired taste. … The likely result is that students miss vital information, make unnecessary mistakes, and function inefficiently. They may also become confused and discouraged. Although they might muddle through on their own, it is unlikely that they will learn with optimal efficiency or thoroughness.

As instructors, we are particularly susceptible to the expert blind spot. We often exercise the research and writing skills needed in economics so automatically and instinctively that we may no longer be aware of what we know or do—we just do it. We take shortcuts our students cannot. In general, this works for us in our own research and writing, but it is often an impediment to teaching research skills to our students. Therefore, a key element in this project design is developing writing instructions for students that clearly articulate expectations but is not so exhaustive and detailed that it risks overwhelming them by exceeding their cognitive load. Successfully navigating this trade-off requires that instructors put themselves in the position of their students, many of whom are generating reproducible empirical research for the first time.

To further understand this problem, consider a model of mastery developed by Sprague and Stuart (2000) and illustrated in Figure 2, which describes a developmental trajectory from novice to expert.

Figure 2. Stages in Development of Mastery



Novice students do not know what they do not know; they are in a stage of unconscious incompetence, often leading to inflated self-assessments of their own abilities (Kruger and Dunning 1999). As they gain knowledge and experience, they begin to be aware of what they do not know and what they need to learn—conscious incompetence. As students develop considerable

competence, yet still must think deliberately and consciously, they advance to conscious competence. Finally, as (if) students reach mastery, they enter into a state of unconscious competence, where they exercise skills automatically and instinctively (Ambrose et al. 2010). This developmental divide, with instructors unaware of what they know and students unware of what they do not know, generates confusion for students and frustration for instructors. This challenge can be overcome with effective assignment sequencing and administration of detailed assignment prompts that clearly communicate expectations and goals.

As instructors, writing assignment prompts is one of the two most important forms of writing we do for our students (Gottschalk and Hjortshoj 2004).[6] Expert blind spots and time constraints often lead us to neglect the fact that how we write will affect how our students will write. Ambiguous assignments create unnecessary misalignment of expectations and the resulting papers represent work derived from standards students have previously been exposed to, not the sophisticated knowledge of component skills and their applications that we wish to assess. Comprehensive prompts are less time-consuming for instructors and less over-whelming for students when complex tasks—like an empirical research project—are divided into several smaller tasks that are more manageable. The impact of this type of sequenced project design, process writing (Schemiser 2017) or iterative research (McGoldrick 2008b), on student learning depends upon the clarity of expectations and transparency of boundaries. Decomposing project tasks into component parts allows students to practice component skills before integration, which is essential to student learning (Ambrose et al. 2010).

4.2. Replication Documentation

In recent years, the lack of reproducibility in scientific research has received much attention (National Academies of Sciences 2016). Historically, the economies discipline has done little to promote reproducibility and facilitate replication (Duvendack, Palmer-Jones, and Reed 2017). While many top journals in economics now require authors to submit data and analysis code files as a precondition for publication (including *Econometrica*, *Journal of Political Economy*, and *Journal of Applied Econometrics*), this has not always been the case and still only constitutes a small percentage of economics journals. Prior to 2004, the submission policy for the *American Economic Review* stated that it would only publish papers "if the data used in the analysis are clearly and precisely documented and are readily available to any researcher for purposes of replication" (Bernanke 2004). However, this policy lacked any enforcement provision—authors who failed to provide data and code did not face any penalty (McCollough and McKittrick 2009). In response to the findings of McCullough and Vinod (2003), the *American Economic Review* refined its policy, requiring that authors archive their data and code on the American Economic Association website as a precondition for publication (Bernanke 2004).

The results of recent replication studies in economics are not reassuring. Chang and Li (2015) attempt to replicate 67 empirical papers published in 13 well-regarded economics journals; some of these journals require data and code files and others do not. They were able to replicate

[6] The other is written feedback from the instructor, which is discussed in Section 5.

the key qualitative results for 49% of the papers (Chang and Li 2015).[7] Given these results, promoting transparency and reproducibility in the economics discipline is likely to be an ongoing challenge. There is no clear path to establishing robust incentives for documentation and replicability, beyond perhaps more data and code archiving requirements (Anderson et al. 2008; Ball and Medeiros 2012). As instructors, we *do* have the ability to set the standards and incentives that guide the work of our students, emphasizing replicability and documentation in empirical research. In doing so, we have the potential to create a "trickle-up" effect, as our students enter the workforce (Ball and Medeiros 2012).

For those academic journals that do require archiving of data and code files, most only require that the data and code used to produce the final empirical result be archived. For example, the *American Economic Review* requires that authors archive "the data set(s) and programs used to run the final models," but it does not require authors to submit the original, unprocessed data sets and code with the commands that import and transform the data. This can best be described as *partial replication* (Ball and Medeiros 2012). A higher standard of replicability would require "authors to submit all of the programs used to transform the raw data files into the tables and figures found in the paper," because this "leaves no ambiguity about what procedures the authors conducted to perform their analysis" (Glandon 2011). The TIER protocol is based on this standard and is modeled in this project, what Ball and Medeiros (2012) call a *"soup-to-nuts" replication*. To meet this standard, the replication documentation for a project must include:

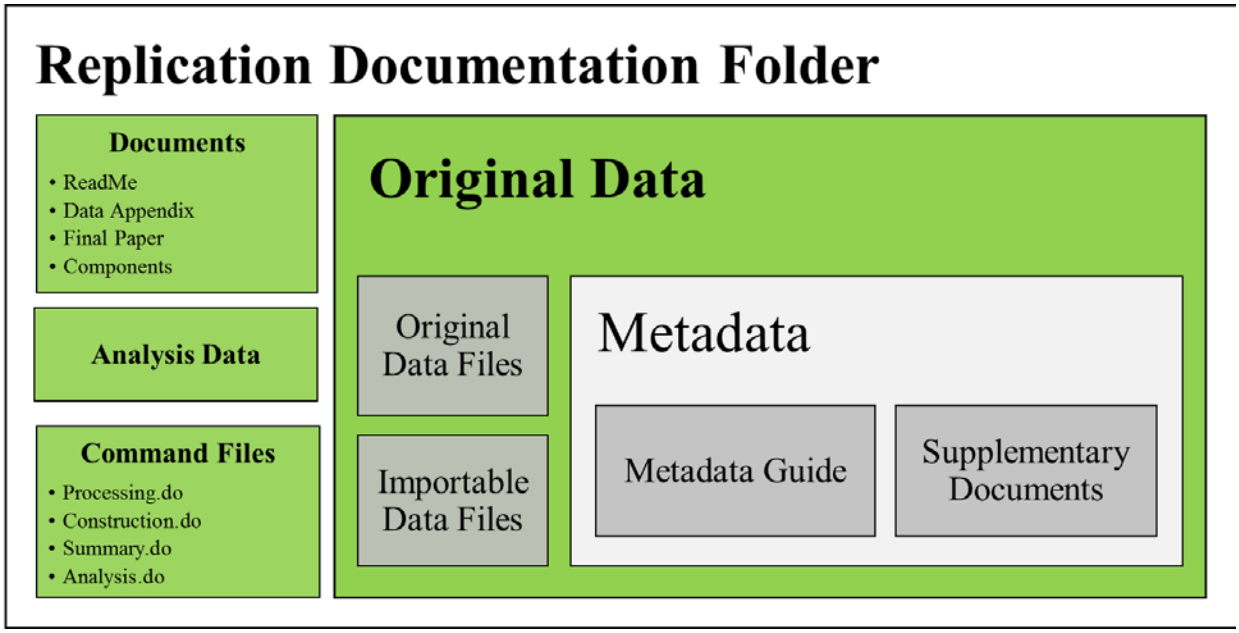
> (1) copies of all raw data files, in their original unmodified form;
> (2) metadata documentation that provides a researcher with all information necessary to obtain and interpret the raw data file(s); and
> (3) easily readable command files that execute all data processing and analysis steps required to reproduce the results of the paper (Ball and Medeiros 2012).

These files should contain all commands required to import raw data files, clean, process, combine, generate new variables, and produce the final empirical estimates in the paper.

As mentioned in Section 2, we require that students create this folder structure at the beginning of the semester. See Figure 3 below for more details. The students then submit the project components in the appropriate folder over the course of the semester—writing tasks in the Documents folder, command files in the Command Files folder, and data files in the appropriate data folder (Original, Importable, or Analysis). Finally, this documentation should be made publically available to facilitate replication. Our guiding principle in adapting the TIER protocol to this empirical project is to come as close to the soup-to-nuts replication standard as possible.

[7] For 22 of the 67 papers (33%), Chang and Li (2015) were able to replicate the results without contacting the authors. For an additional 7 papers, they were able to replicate after contacting the authors. Another 8 papers were excluded that utilized confidential data or unavailable software.

Figure 3: Sample Folder Structure



Note: Adapted from the TIER protocol (Ball and Medeiros 2012). See https://www.projecttier.org/tier-protocol/ for more details.

5. Evaluation and Assessment

For proper implementation of the empirical research project described above, the project should constitute a large proportion of the students' final course grade. The project must comprise a large proportion of the final course grade so that students understand the significance of the assignment for their learning experience and recognize the appropriate amount of time that should be devoted to the project.[8] At the same time, it is important that students are evaluated and given formative feedback at each stage, indicating that each phase is significant and continually providing opportunities for revision.[9] Placing a positive weight on each part of the research project helps reduce the probability that students turn in "draft" work for the individual phases. Assigning a grade to each project stage helps students realize that these are not "drafts;" they are documents that have already been through the drafting process. However, as with any long-term project, the phases of the projects will likely be edited again before submission of the final research paper.

In the evaluation and assessment process, the objectives outlined in the project prompts must align with the criterion for assessment. Practice activities motivated by goals and direction should be coupled with targeted feedback in order to maximize learning.[10] These two activities, practice and feedback, can be viewed as a cycle, where practice results in observed outcomes that can be assessed, which further informs future practice. Goals and learning objectives are at the

[8] In our course, the project is worth 225 points out of 500 total points available, or 45% of the final course grade. At national universities and liberal arts colleges that require a project in econometrics, the projects typically constitute 28-29% of the course grade (Johnson et al. 2012).
[9] Each part of the project is worth 9 to 20% of the project grade, or 4 to 9% of the final course grade. Please see Appendix Section A for additional point distribution details.
[10] In this context, "practice" refers to the components of the final paper and the final paper itself. Practice is defined as any activity in which students apply their knowledge or skills.

center of this cycle and guide every stage, but they must be written in a way that can be monitored and measured (Ambrose et al. 2010). In addition, research has shown that goals must specify exactly what students should learn (Rothkopf and Billington 1979). Explicit requirements about details in the absence of overarching learning goals can result in students failing to meet important benchmarks, further emphasizing the necessity of clear and detailed prompts in conjunction with targeted feedback. In what follows, we focus on the implementation of targeted feedback. For more information on goal-directed prompts, refer to Section 4.1.

Feedback and assessment should align with objectives outlined in well-designed prompts. Goal-directed practice and coordinated targeted feedback through articulating expectations and identifying an appropriate level of challenge support the greatest learning gains (Ambrose et al. 2010). Frequent, timely, and accurate targeted feedback is an important part of the writing process (Ericsson et al. 1993). Feedback should be provided early (Mathan and Koedinger 2005) and often (for a review, see Hattie and Timperley 2007). Furthermore, research shows that any feedback (even minimal) is superior to no feedback (Traxler and Gerensbacher 1992). In order to apply these principles regarding the timing of feedback, the project design necessitates feedback at each stage of the writing process.

The content of feedback given to students dictates the degree to which learning occurs through the revision process. Beason (1993) shows that students respond to about 90% of teacher's concerns in the revision process; so, the quality and direction of instructor comments are extremely important. Research suggests that when the opportunity for revision exists (as in the type of project outlined above), formative assessment relative to stated goals and target criteria offer the greatest opportunities for improvement (Black and William 1998; Cardelle and Corna 1981; McKendree 1990).[11] By providing the opportunity for revision, we give control and ownership of the writing process to the students.

In responding to student writing, while any feedback is better than none, too much feedback can be counterproductive and overwhelm students (Lamburg 1980; Shuman 1979). Students tend to address only those comments that are easy to change (for example, they focus on the details as opposed to the structure). In commenting, instructors must be careful to balance between global-meaning (major) and local (surface-level or minor) comments. Surface-level comments can be important for improving writing skills and facilitating communication of economic content. However, in the presence of over-commenting or only surface-level feedback, students will often ignore areas for global improvement.[12] Several techniques can facilitate this feedback process. First, a file-sharing platform and use of track changes in Microsoft Word can reduce the time it takes to provide comments.[13] Second, "minimal marking" can be used to identify local errors without having to make the same corrections throughout the entire paper. The purpose of minimal marking is to choose one representative paragraph from the student's writing and identify all

[11] For more information on written responses to student writing, see Flanigan and Menendez (1980), Flower (1979), Horvath (1984), Kirby and Liner (1981), Lynch (1982), McDonald (1978), Murray (1972), and Sommers (1982).
[12] In addition, it is tempting to spend more time on weaker student papers and less time providing feedback to strong student papers. Instructors must be careful to balance feedback between weak and strong student papers.
[13] If you are unfamiliar with track changes, see https://support.office.com/en-us/article/Track-changes-in-Word-197ba630-0f5f-4a8e-9a77-3712475e806a.

surface-level comments and areas for improvements (Haswell 1983). Students are then asked to extrapolate and apply knowledge of these errors to the remainder of the paper. From a pedagogical perspective, this technique forces students to struggle with the grammatical rules and understand mistakes. Haswell (1983) finds that, of all errors identified, students are able to correct 61%.[14]

The most straightforward way to articulate expectations and align objectives and outcomes is to provide rubrics. In addition, rubrics serve as an efficiency enhancing mechanism for instructors by lowering the marginal cost of grading. Moreover, rubrics aid instructors in recognizing strengths and weaknesses for individuals and classes, signifying areas of focus for additional learning, and encourage consistency in evaluation (Ambrose et al. 2010). Finally, rubric and rubric sharing have been shown to increase the quality of student work, student knowledge of quality work, and academic performance (Andrade 2001; Osana and Seymour 2004; Reitmeier, Svendsen, and Vrchota 2004; Andrade and Du 2005; Schneider 2006; Reddy and Andrade 2010).

The project structure, outlined in Figure 1, supports implementation of the pedagogical approaches to feedback described above. Approximately every 2 weeks throughout the semester, students are receiving feedback on the practice activities (project components). Appendix Section C provides instructions for how to access rubrics for each project phase to streamline the assessment and evaluation process for instructors.

6. Conclusions and Extensions

In this this paper, we describe an empirical research project for use in an upper-level undergraduate economics writing-in-the-discipline course with goals of reducing the high fixed costs associated with designing this type of assignment and encouraging more undergraduate economics research. In presenting the sequenced project design and replication documentation protocol, we posit that requiring students to produce this comprehensive documentation promotes student learning and leads to improvements in organization and coherence throughout the entire research and writing process.

The adaptable nature of the project structure and resources lends itself to innovation and extension. As previously mentioned, we utilize Stata, Microsoft Word, and Microsoft OneDrive for statistical computing, word processing and providing feedback, and assignment submission and return, respectively. The TIER protocol (currently in Version 3.0), which we adapt in this project, is software neutral.[15] The choice of statistical software is, therefore, determined by the instructor and the available computing resources. The use of programs such as R Markdown can further streamline the process by seamlessly combining LaTeX and R syntax to estimate empirical results and produce a final document.[16] Using batch files to stitch together Stata and LaTeX code can provide similar functionality and is equally reproducible. Depending on the learning objectives of the course, these project components can be scaled up or down to meet the needs of instructors and their students.

[14] Note, there are many variations on minimal marking and only one is described above.
[15] Originally, this protocol was developed for Stata and R. See http://www.projecttier.org/tier-protocol/ for more details.
[16] Stata 15 also now has integration capabilities with Microsoft Word. See https://www.stata.com/new-in-stata/create-word-documents/ for information on this update.

As an additional learning experience, instructors may offer the opportunity for peer review. This can be facilitated in class, outside of class, or electronically. Appendix Section D (b) outlines an example of a peer review activity. In addition, instructors can provide an online platform for students to share project ideas, data sources, and help with coding through a blog or discussion board on a learning management system.

Finally, as more students and instructors begin to utilize such tools in the classroom, opportunities for sharing undergraduate research become available. Regional collaborations among institutions can organize symposia designed to bring undergraduate students together to share their research and bring instructors together to share their innovative pedagogical extensions of these ideas. This type of partnership can contribute to a research culture among undergraduate economics students that yields learning valuable to both graduate programs and potential employers.

Exposure to the research process is a key component in teaching undergraduate students to "think like economists" (Hoyt and McGoldrick 2017a). It fosters a deep appreciation for what it is that economists actually do. The resource-intensive nature of supervising undergraduate research, alongside the rising popularity of econometrics and increasingly empirical economics discipline, has left time-constrained economics departments and faculty unable to provide ample research opportunities for undergraduates. Almost universally, economics faculty believe that exposure to the research process is vital to the development of well-rounded economics majors and cite resource constraints as the primary reason for the inability to provide these opportunities (Hoyt and McGoldrick 2017a). In this paper, we have discussed an empirical research project that can be adapted in variety of undergraduate economics courses. The pedagogical rationale provided and teaching resources included in the Appendix are to designed to reduce the otherwise large fixed costs associated with such a project and promote more opportunities for undergraduate research. The highly sequenced project structure and replication documentation requirements support student learning of course content and expose students to "doing economics" as a tool of discovery.

7. Acknowledgements

We would like to thank Noreen Lape, Gail Hoyt, and Richard Ball for comments, advice, and suggestions throughout the writing process. We also thank attendees of the November 2017 Annual Professors Conference at the Federal Reserve Bank of St. Louis and the November 2017 Annual Economics Teaching Conference who provided valuable comments.

8. References

Allgood, S., Walstad, W. B., and J.J. Siegfried. 2015. Research on Teaching Economics to Undergraduates. *Journal of Economic Literature*, 53(2): 285-325.

Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., and M.K. Norman. 2010. *How learning works: Seven research-based principles for smart teaching*. John Wiley and Sons.

American Statistical Association (ASA), Committee on Professional Ethics. 2016. Ethical guidelines for statistical practice. American Statistical Association. http://www.amstat.org/ASA/Your-Career/Ethical-Guidelines-for-Statistical-Practice.aspx. (accessed September 27, 2017).

Anderson, R. G., Greene, W. H., McCullough, B. D., and H.D. Vinod. 2008. The role of data/code archives in the future of economic research. *Journal of Economic Methodology*, 15(1): 99-119.

Andrade, H. G. 2001. The effects of instructional rubrics on learning to write. *Current Issues in Education*, 4(4).

Andrade, H., and Y. Du. 2005. Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, 10(5): 1–11.

Angrist, J., Azoulay, P., Ellison, G., Hill, R., and S. F. Lu. 2017. Economic Research Evolves: Fields and Styles. *American Economic Review*, 107(5): 293-297.

Ball, R., and N. Medeiros. 2012. Teaching integrity in empirical research: A protocol for documenting data management and analysis. *Journal of Economic Education*, 43(2): 182-189.

Bernanke, B. 2004, March. Editorial Statement. *American Economic Review*, 94(1): 404.

Black, P. and D. William. 1998. Assessment and classroom learning. *Assessment in Education*, 5: 7-74.

Bosshardt, W., Watts, M., and W.E. Becker. 2013. Course Requirements for Bachelor's Degrees in Economics. *American Economic Review*, 103(3): 643-647.

Cardelle, M. and L. Corna. 1981. Effects on second language learning of variations in written feedback on homework assignments. *TESOL Quarterly*, 15: 251-261.

Camerer, C., G. Loewenstein, and M. Weber. 1989. The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy* 97(5): 1232-1254.

Carter, M. 2007. Ways of knowing, doing, and writing in the disciplines. *College Composition and Communication*: 385-418.

Chang, A. C., and P. Li. (2015). *Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say "Usually Not"* (No. 2015-83). Board of Governors of the Federal Reserve System (US).

Cohen, A. J., and J. Spencer. 1993. Using writing across the curriculum in economics: Is taking the plunge worth it?. *Journal of Economic Education*, 24(3): 219-230.

Dickinson College. "Writing-in-the-Discipline Courses." The Writing Culture. http://www.dickinson.edu/info/20158/writing_program/870/the_writing_culture/3 (accessed July 26, 2017).

Duvendack, M., Palmer-Jones, R., & Reed, W. R. 2017. What Is Meant by "Replication" and Why Does It Encounter Resistance in Economics?. *American Economic Review: Papers and Proceedings*, *107*(5): 46-51.

Dynan, L., and T. Cate. 2009. The impact of writing assignments on student learning: Should writing assignments be structured or unstructured?. *International Review of Economics Education*, 8(1): 64-86.

Ericsson, K.A. and N. Charness. 1994. Expert performance: Its structure and acquisition. *American Psychologist*, 49: 725-747.

Ericsson, K.A. and A.C. Lehmann. 1996. Expert and exceptional performance: Evidence on maximal adaptations on task constraints. *Annual Review of Psychology*, 47: 273-305.

Ericsson, K.A., Krampe, R.T., and C. Tescher-Romer. 1993. The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100: 363-406.

Flanigan, M. C., and D. S. Menendez. 1980. Preception and Change: Teaching Revision. *College English* 42(3): 256-266.

Flower, L. 1979. Writer-based prose: A cognitive basis for problems in writing. *College English*, 41(1) : 19-37.

Glandon, P. J. 2011. Appendix to the report of the editor: Report on the *American Economic Review* Data Availability Project. *American Economic Review: Papers and Proceedings* 101(3): 695–99.

Gottschalk, K., and K. Hjortshoj. 2004. *The elements of teaching writing: A resource for instructors in all disciplines*. Macmillan.

Greenlaw, S. A. 2003. Using writing to enhance student learning in undergraduate economics. *International Review of Economics Education*, 1(1): 61-70.

Greenlaw, S. A. 2006. *Doing economics: A guide to understanding and carrying out economic research.* Boston: Houghton Mifflin.

Hamermesh, D. S. 2013. Six decades of top economics publishing: Who and how?. *Journal of Economic Literature*, 51(1): 162-172.

Hansen, W. L. 1986. What knowledge is most worth knowing--for economics majors?. *American Economic Review*, 76(2): 149-152.

Hansen, W. L. 2001. Expected proficiencies for undergraduate economics majors. *Journal of Economic Education*, 32(3): 231-242.

Haswell, R.H. 1983. Minimal Marking. *College English*, 45(6): 600-604.

Hattie, J. and H. Timperley. 2007. The power of feedback. *Review of Educational Research*, 77: 81-112.

Hultberg, P. T., and D.S. Calonge. 2017. Effective teaching of economics: A constrained optimization problem?. *Journal of Economic Education*, 48(4): 265-275.

Hoyt, G. M., and K. McGoldrick. 2017a. Promoting Undergraduate Research in Economics. *American Economic Review: Papers and Proceedings*, 107(5): 655-659.

Hoyt, G. M., and K. McGoldrick. 2017b. Models of undergraduate research in economics: Advice from eight exemplary programs. *Journal of Economic Education*, 48(4): 288-289.

Johnson, B. K., Perry, J. J., and M. Petkus. 2012. The status of econometrics in the economics major: A survey. *Journal of Economic Education*, 43(3): 315-324.

Klein, C. C. 2013. Econometrics as a capstone course in economics. *Journal of Economic Education*, 44(3): 268-276.

Kruger, J., and D. Dunning. 1999. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology* 77(6): 1121-1134.

Lamburg, W. 1980. Self-provided and peer-provided feedback. *College Composition and Communication*, 31(1): 63-69.

Li, I., and R. Simonson. 2016. Capstone senior research course in economics. *Journal of Economic Education*, 47(2): 161-167.

Mathan, S.A., and K.R. Koedinger. 2005. Fostering the intelligent novice: Learning from errors with metacognitive tutoring. *Educational psychologist* 40(4): 257-265.

McKendree, J. 1990. Effective feedback content for tutoring complex skills. *Human-Computer Interaction*, 5(4): 381-413.

McCullough, B. D., and R. McKitrick. 2009. *Check the numbers: The case for due diligence in policy formation*. Toronto, Canada: Fraser Institute.

McCullough, B. D., and H.D. Vinod. 2003. Verifying the solution from a nonlinear solver: A case study. *American Economic Review*, 93(3): 873-892.

McCullough, B. D., and H.D. Vinod. 2004. Verifying the solution from a nonlinear solver: a case study: reply. *American Economic Review*, 94(1): 391-396.

McGoldrick, K. 2008a. Writing requirements and economic research opportunities in the undergraduate curriculum: Results from a survey of departmental practices. *Journal of Economic Education*, 39(3): 287-296.

McGoldrick, K. 2008b. Doing economics: Enhancing skills through a process-oriented senior research course. *Journal of Economic Education*, 39(4): 342-356.

Nathan, M. J., and A. Petrosino. 2003. Expert blind spot among preservice teachers. *American Educational Research Journal*, 40(4): 905-928.

National Census of Writing. "Four-year Institution Survey." https://writingcensus.swarthmore.edu/survey/ (accessed October 25, 2017).

National Academies of Sciences, Engineering, and Medicine. 2016. *Statistical Challenges in Assessing and Fostering the Reproducibility of Scientific Results: Summary of a Workshop*. Washington, DC: The National Academies Press. https://doi.org/10.17226/21915.

Osana, H.P., and J.R. Seymour. 2004. Critical thinking in preservice teachers: A rubric for evaluating argumentation and statistical reasoning. *Educational Research and Evaluation*, 10(4–6): 473–98.

Petkus, M., Perry, J. J., and B.K. Johnson. 2014. Core requirements for the economics major. *Journal of Economic Education*, 45(1): 56-62.

Reddy, Y.M. and H. Andrade. 2010. A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35(4): 435–448.

Reitmeier, C.A., L.K. Svendsen, and D.A. Vrchota. 2004. Improving oral communication skills of students in food science courses. *Journal of Food Science Education*, 3: 15–20.

Rosenwasser, D., and J. Stephen. 2012. *Writing Analytically*. 6th edition. Boston: Wadworth, Cengage Learning.

Rothkopf, E. Z., and M.J. Billington. 1979. Goal-guided learning from text: inferring a descriptive processing model from inspection times and eye movements. *Journal of Educational Psychology,* 71(3): 310-327.

Schmeiser, K. 2017. Teaching writing in economics. *Journal of Economic Education*, 48(4): 254-264.

Schneider, J.F. 2006. Rubrics for teacher education in community college. *The Community College Enterprise,* 12(1): 39–55.

Shuman, R.E. 1979. How to grade student writing. In G. Stanford (Ed.). *Classroom practices in teaching English 1979-1980: How to handle the paper load*. Urbana, IL: National Council of Teachers of English.

Selingo, J. 2017. "Why can't college graduates write prose?" *The Washington Post*, August 11.

Siegfried, J. J., Bartlett, R. L., Hansen, W. L., Kelley, A. C., McCloskey, D. N., and T.H. Tietenberg. 1991. The status and prospects of the economics major. *Journal of Economic Education*, 22(3): 197-224.

Siegfried, J. J., and W.B. Walstad. 2014. Undergraduate coursework in economics: A survey perspective. *Journal of Economic Education*, 45(2): 147-158.

Siegfried, J. J., and J.T. Wilkinson. 1982. The economics curriculum in the United States: 1980. *American Economic Review*, 72(2): 125-138.

Simpson, M. S., and S.E. Carroll. 1999. Assignments for a writing-intensive economics course. *Journal of Economic Education*, 30(4): 402-410.

Sprague, J., and D. Stuart. 2000. *The Speaker's Handbook*, Ft. Worth, TX: Harcourt College Publishers.

Traxler, M.J. and M.A. Gerensbacher. 1992. Improving written communication through minimal feedback. *Language and Cognitive Processes*, 7: 1-22.

Wieman, C. E. 2007. APS News – The back page. The "curse of knowledge" or why intuition about teaching often fails. *American Physical Society News*, 16(10).

Appendix

Below, we present supplementary material associated with the empirical research project design and intended for direct implementation with adaptations. Throughout, we have highlighted in yellow the course-specific information that needs to be filled in by each instructor.

Table of Contents

A. Empirical Paper Prompt

Course Title
Section
Empirical Paper Assignment
Due: Date

Prompt:

The goal of this project is for you to develop an understanding of how economists conduct applied empirical research. To this end, you should demonstrate:

> (1) an understanding of Stata syntax, data management skills, and best coding and documentation practices for reproducibility;
> (2) the ability to place a research question in the context of existing scholarly discourse through an effective literature review; and
> (3) an understanding of the necessary components of a well-written empirical research paper and the economics discipline formatting and style conventions.

You are to choose your own topic and develop a well-defined, innovative research question. This question, should, in general, have some implied causality. That is, based on expectations derived from economic theory and existing empirical research you should explore whether changes in *X* lead to changes in *Y*, ceteris paribus. Your paper should examine an issue related to current macroeconomic or microeconomic discourse, broadly defined. Papers that simply rehash class material will receive poor grades; good papers will apply the empirical tools in a rigorous and thoughtful manner. Your chosen econometric methodology need not be groundbreaking, just well done and complete, but you should be identifying and filling a gap in the literature and/or contributing to a scholarly discussion.

The reader should be able to easily recognize the role or purpose of your paper. The audience of your essay should be an informed reader, colleague, or peer. You should be a participant in a scholarly conversation. As such, your writing should be academic, technical, maintain a critical distance from the examination, and refrain from personal narratives. Econometricians are concerned with analyzing data to test the plausibility of economic theory or develop new understandings of human behavior and its consequences for social well-being. Often, econometricians use their analyses to provide policy prescriptions and suggestions for improving welfare. As such, your essay should explain to the reader why the topic is important and the motivation for the analysis.

The other guiding principle of this project is the belief that providing comprehensive replication documentation for research involving statistical data should be as ubiquitous and routine as providing a list of references. Thus, over the course of the semester, you will be required to produce detailed replication documentation for your project. You should view this documentation as an essential component of how econometricians communicate their research to other scholars. This workflow should be an integral part of your research throughout the entire process – not a discrete task that you postpone until the end. You should begin constructing your documentation before

you even start working with data and add to it incrementally as your research progresses. More information regarding this documentation will be distributed early in the semester.

There will be several deadlines throughout the semester, as outlined on the course schedule and presented below. On the day the prompt is administered, you will be given a complete description of the assignment. A rubric will be administered after the prompt to ensure that students do not "write to the rubric." Instead, the rubric should be a final checking point to confirm that all the relevant information is included.

Table 1: Assignment Due Dates

Task	Prompt Administered	Due Date
Final Paper		
Research Question & Annotated Bibliography		
Proposal		
Data Collection & Metadata Guide*		
Literature Review		
Data, Methods, & Data Appendix*		
Results & Discussion*		
Research Paper Compilation & Replication Documentation*		

Cells highlighted in red represent components of the final grade.
Cells highlighted in gray are components of the final paper.
Cells highlighted in blue are NOT components of the final paper but are counted in the project grade.
*denotes that data (.xlsx or .dta) files and/or command (.do or log) files are due at this stage also.

Additional Information:

Here are some points to keep in mind:

- When in doubt, narrow your topic. It is far easier to extend a seemingly small subject than to adequately summarize a vast subject.
- Good papers may be either normative (advocating a particular position/policy) or positive (describing an issue).
- Your grade will largely depend on the contribution that your paper makes beyond what was discussed in class or in the existing literature.
- Do not neglect your other coursework. If your training in other economics courses or other disciplines provides insight to your topic, then include it in your paper.
- The quality of your writing may affect your paper grade. Keep in mind that this assignment is due during a very busy time of the year. I make no apologies for the fact that a clearly presented paper is likely to score better than a poorly written paper, even if the underlying quality of the econometric analysis is the same. Please utilize campus resources including the ones listed below.
- Remember that your final empirical paper should constitute professional writing.

Evaluation Criterion:

A separate prompt and rubric will be given for each task in the empirical research project. Your project grade is distributed as follows, according to the syllabus:

Table 2: Point Distribution

Task	Points	Percentage of Grade
Final Paper	225	45%
Research Question & Annotated Bibliography	30	13% (of project) 6% (of final grade)
Proposal	20	9% (of project) 4% (of final grade)
Data Collection & Metadata Guide*	35	16% (of project) 7% (of final grade)
Literature Review	30	13% (of project) 6% (of final grade)
Data, Methods, & Data Appendix*	35	16% (of project) 7% (of final grade)
Results & Discussion*	30	13% (of project) 6% (of final grade)
Research Paper Compilation & Replication Documentation*	45	20% (of project) 9% (of final grade)

Cells highlighted in red represent components of the final grade.
Cells highlighted in gray are components of the final paper.
Cells highlighted in blue are NOT components of the final paper but are counted in the project grade
*denotes that data (.xlsx or .dta) files and/or command (.do) files are due at this stage also.

I will provide feedback on each section using track changes in Microsoft Word. If you are unacquainted with this resource, please familiarize yourself with the track changes tool: https://support.office.com/en-us/article/Track-changes-in-Word-197ba630-0f5f-4a8e-9a77-3712475e806a.

Writing Resources:

In addition, please take full advantage of the writing and research resources available to you at Institution Name, including the following website:

- Website link

Librarian Name is the economics department contact in Library and Information services. For help with research, finding sources, collecting data, etc., please see:

- Library Website link

Individual sessions at the Writing Center are provided throughout the semester. See the following link for information on how to make an appointment, among other things:

- Writing Center link

A guide to citing sources is available at:

- Library citing sources link

Usually, economists use Chicago or APA citation style. However, if you are familiar with another style that is a common format (i.e. ASA, GSA, and MLA), please talk with me outside of class.

You likely have read several academic articles this semester in economics, either as required readings for other courses or in collecting literature for the annotated bibliography/literature review. These should give you a good idea of how economists write and the format that I expect for this assignment. Of course, you will have to tailor your style and sections to your paper topic. I will also take this time to direct you to

several writing guides that I have found to be useful. The following are links to a Guide to Scientific Writing, a Guide to Writing in Economics, and writing tips for economics research papers respectively:

- http://www.aacc.org/publications/clin_chem/ccgsw/Pages/default.aspx#
- http://lupus.econ.duke.edu/ecoteach/undergrad/manual.pdf
- http://www.people.fas.harvard.edu/~pnikolov/resources/writingtips.pdf

In addition, these two books may be of use:

- McCloskey, Deirdre N. The writing of economics. No. E10-1470. 1987.
- Thomson, William. A guide for the young economist. MIT Press, 2001.

Academic Integrity:

All papers will be electronically analyzed to verify that they constitute original work. I reserve the right to submit your paper to electronic databases that will check for academic misconduct both during this term and in the future. You are required to use appropriate (in-text) citations. This includes giving credit whenever you quote or paraphrase another person's work, or when you borrow their ideas. Restating a sentence does not absolve you of your obligation to give credit. If you need help or have questions, then ask. Failing to properly cite is a common form of unintentional plagiarism and will adversely affect your grade.

Intentional plagiarism is the deliberate or unreasonably careless representation of another's work as your own, either as a portion of your paper or as its entirety. This includes, but is not limited to, purchasing a paper that someone else wrote, using a friend's paper, downloading a paper from the Internet, or knowingly aiding in another student's intentional plagiarism. Intentional plagiarism represents academic misconduct, and I intend to fully pursue all instances.

If you are planning on using a paper from another class as the basis for this assignment, you must clear it with me beforehand. I am not entirely opposed to your doing this, provided that there is a substantial and new contribution to the paper that is related to this course. Failing to clear this with me, however, may adversely affect your grade. Extreme cases (e.g. handing in exactly the same paper from another class) may be treated as intentional plagiarism.

B. Project Timeline and Workflow

Just as good writing requires forethought and planning, so, too, does effective data workflow and construction of replication documentation. Thus, you should view this project as two interrelated projects happening simultaneously: the writing process AND data workflow. This workflow is designed to help you keep organized and enhance your own understanding of the data processing and analysis you do.

Table 3: Project Timeline and Data Workflow

Dates	Writing Process	Data Workflow
	Brainstorm Research Questions/Topics related to your interests.	Explore potential data sources related to these topics. Which variables do you need?
	Find sources. Read them. Engage with them. How can your research contribute to this conversation?	**Complete "Pre-Data" work.** Construct a hierarchy of empty folders in Office 365 and create three blank documents: • A ReadMe file • A Metadata Guide • A Data Appendix
	Write Annotated Bibliography and Research Question. Refine the focus of your paper.	Find data. Explore it. Based on the existing literature do you have all the variables you need? If not, find them.
	Submit your Annotated Bibliography and Research Question	Begin "Data Work"
	Evolve your thesis statement. Narrow your focus.	Each time you obtain a new file containing data you will use for your project, you should save a copy in your Original Data folder, and record some information about it in your Metadata Guide and ReadMe file.
	Write your Proposal.	Begin writing your "Processing.do" command file as you obtain these original data files. Save this in your Command files folder.
	Submit your Proposal	
	Further engage with the sources from your annotated bibliography and all other additional sources as you continue to narrow your focus. Where does your research question fit into the existing conversation? How does it contribute?	After receiving feedback on your proposal, verify you have all the data that you need to address your research question. Begin cleaning and processing in order to generate your base data file, which you

		should save in your Original Data folder.
	Begin writing your Literature Review. Read it. Does it achieve its goal? Are you engaging with the sources critically and meaningfully or just summarizing? Is it beginning to inform your empirical methodology?	Record information about any additional data in your Metadata Guide and your ReadMe file as necessary. Be sure to document any cleaning decisions that were made during the generation of the base data in your Metadata Guide.
		Submit original data files, base data file, Processing.do command file, and Metadata Guide
	Revisit your Literature Review. Is it consistent with your thesis? Does your thesis need to evolve? Does your focus need to be further narrowed or refined?	Begin constructing your analysis data file(s) from your base data file. Document these commands in your Construction.do command file and save it in your Command Files folder. Any new variables, transformed variables, etc. should be generated during this phase.
	Again revisit your Literature Review. Does it motivate your empirical methodology? How is what you are doing different from what others have done? How is it the same? You should have a clear sense of you model specification by this time.	Once you've completed data construction, save your analysis data file(s) in your Analysis Data folder and immediately begin working on your Data Appendix and compile these commands in the Summary.do command file and save in your Command files folder.
	Using information already compiled in your Metadata Guide, ReadMe file, and Data Appendix begin writing the Data section of your paper.	Get to know your data. What stands out? Which aspects of the composition of your sample are most relevant?
	Submit your Literature Review	Give your documentation a "check-up". Is everything there? Have you worked on your ReadMe file?
	Revisit the Data section of your paper. Which details about your data does your reader need to know to understand your methodology and the meaning of your results?	Complete your Data Appendix. Be sure that all commands necessary to generate the descriptive statistics, tables, and figures needed for the Data Appendix are included in the Summary.do file. What is most relevant in describing your data? Does your data make sense? Is it in line with expectations? Do you have

		outliers? What does your reader need to know?
	Develop your regression specification. Write out the equation using the equation editor in Word. What is your explanatory variable of interest? What are your expectations for your estimates? Why?	Which tables, charts, or graphics from your Data Appendix may be helpful to the reader in understanding your data? Which type of chart is most effective in making the point?
	Submit Data & Methods	**Submit analysis data files, Construction.do command file, Summary.do command file, and your Data Appendix**
	After ensuring that you have estimated your model correctly and obtained results, begin interpreting your results. Are they in line with expectations? Are they statistically significant? Are there any surprises?	Begin analyzing your data. All commands used to generate descriptive statistics, graphics, regression results, and hypothesis testing should be compiled in the Analysis.do command file. Every command that generates any of your results should be preceded by a comment that states which result the command generates.
	What can you say about your research question? Provide context for your reader. How should they understand the results?	Generate tables of your results using outreg2. Include these commands in your Analysis.do command file.
	Think big picture about your results. Revisit your results in the context of your research question and thesis statement. Have you answered the question? Likely, you have in some ways and not in others; discuss these. What limitations do you see in your analysis? Were you limited by your data? Did you need more variables? More observations? Are you concerned about unobserved heterogeneity and omitted variable bias? Were you able to effectively deal with problems of heteroskedasticity or serial correlation?	Finish your ReadMe file. You should already have recorded one part of the required information, namely notes explaining any modifications you made to the original data files when you made importable versions of them and generated the base data. To finish your Read Me file, you should add: (1) an overview of all the files included in the replication documentation and the structure of the folders in which they are stored and (2) Step-by-step instructions for using the replication documentation to replicate the study.
	Submit Results & Discussion	**Submit Analysis.do command file**
	Begin construction of your final paper. Most empirical papers include these sections:	Edit all of your command files to be sure they are accurate, concise, and free of clutter. Have you provided

	• Introduction • Literature Review • Data & Methods • Results/Discussion • Conclusion • References	sufficient comments in your command files? Could someone else follow what each command is doing?
	For the most part, you need to simply compile the work you've already done, but do not simply copy/paste them into one document. Your introduction is likely new, perhaps it draws from your proposal, but it should motivate your thesis statement. You'll need to add in transitions between these sections to make the paper more readable and cohesive.	Test your command files to be sure that they all run without error and that they successfully reproduce the results you report in your paper. Try following the instructions for replicating your project that you wrote in the ReadMe file to be sure that all your command files run without a hitch and produce the intended output.
	Check your citations and reference list. Be sure that all sources cited in the paper are listed in the reference list and vice versa. Check to be sure you are consistent with your citation style (Chicago or APA).	Check to be sure your replication documentation is complete. Are all the required files included in your replication documentation, and are they are stored in the correct folders? Finally, delete any extraneous files that are not required and unnecessary for replication.
	Read your paper. Read it again. Does it flow well? Have someone else read it. Do they agree? Revise. Revise. Revise.	Check all documents for accuracy and readability. The ReadMe file, Metadata Guide, and Data Appendix should be well-written, proofread, and formatted.
	Submit Final Paper	**Submit Final Replication Documentation**

C. *Project Component Prompts and Rubrics*

We have designed these prompts and rubrics as a way of reducing the fixed costs associated with inclusion of this type of project in your classroom. Please use these materials as a starting point. They are highly adaptable. Password protected electronic copies of the prompts and rubrics can be downloaded here: http://blogs.dickinson.edu/underwood/teaching-resources and here: http://sites.google.com/site/emilycorinnemarshall/teaching-resources.

Please email either Emily Marshall (marshaem@dickinson.edu) or Anthony Underwood (underwoa@dickinson.edu) from your institutional email account to get the password.

a. Annotated Bibliography and Research Question
b. Proposal
c. Data Collection and Metadata Guide
d. Literature Review
e. Data, Methods, and Data Appendix
f. Results and Discussion
g. Research Paper Compilation & Replication Documentation

D. *Adaptation Guide*

a. Large Classes

We recognize that this project will not be feasible in its current form at all institutions due to class size limitations. However, here we provide several suggestions for large classes that would allow the project to be implemented, even with a relatively sizable number of students without significant burden on the instructor.

i. Group Projects

The most obvious adaptation of this assignment is to reduce grading by assigning group projects. In this case, the remainder of the project structure would be unchanged. Group projects can increase the amount of feedback students receive (from peers) and reduce the amount of feedback that instructors need to provide. In addition, group projects tend to increase the frequency of deliberate practice activities, as students working together must access the material more often. Strategies for group formation, techniques to help mitigate free-rider problems, and methods of evaluating individual member contributions can be downloaded here: http://blogs.dickinson.edu/underwood/teaching-resources and here: http://sites.google.com/site/emilycorinnemarshall/teaching-resources.

ii. Economic Note

Instructors in large classes may wish to shorten the assignment length. A clear variation of this project would be to ask students to write an abbreviated version of each component, with the end goal of constructing an economic note, such as those published in *Economics Letters.*

b. Peer Review

There are several opportunities for peer review throughout all stages of the project sequence. Below is an example of a peer review assignment for the literature review. Instructors may also want to consider assigning referee reports as a form of peer review.

Course Title
Section
Peer Review – Literature Review
Due: Date

Focus on the areas that you and your partner decide need attention first. You are not required to answer the questions in order. Make sure you are using class time to discuss the feedback with the writer, as opposed to spending the entire time writing out the responses. You can finish the written responses after class; it is more important for you to use the class time to interact with your partner. Partner assignments are below:

Group	Name	Name
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		

Writers, after the peer review and before you revise, write a paragraph in which you state your revision goals – that is, what you plan to do in the next draft. Hand in this paragraph, as well as the peer review response sheets with the Literature Review. Both of these responses should be typed.

1) What do you find engaging about the introduction?

2) Evaluate the quality of the thesis statement. Remember a good thesis is arguable, novel, and compelling.

3) What points or paragraphs seem unrelated to the thesis? How might the writer make the connection clearer?

4) How does the writer's evidence support his/her claims? Are all claims supported with evidence? Does your partner interact well with the evidence, using it clearly and adding analysis to connect it back to their argument? Explain.

5) How would you describe the type of research the writer did? Does the writer need to do additional research to fill in gaps or address unanswered questions?

6) What missing evidence or counter-evidence might the writer need to address?

7) Where does the writer let the evidence speak for itself? (Remember, it is the writer's job to explain "So what?" when he provides evidence for the reader.)

8) On a scale of 1-6 (6 being the highest), how well organized is the draft? How would you describe the organization as you perceive it?

9) Are your partner's body paragraphs clear, with one clear point being made in each?

10) Where might the writer need to re-order points or insert transitions?

11) On a scale of 1-6 (6 being the highest), how often were you distracted by mechanical (usage, punctuation, spelling) errors? What kinds of errors did you notice? (You need not be the writer's proofreader or editor, but you can help by telling the writer the types of errors you noticed.)