How to Write Applied Papers in Economics*

Marc F. Bellemare†

September 7, 2020

Abstract

How does one write good research papers? Though many research economists instinctively know how to do so, most have spent too little time thinking about how one writes good research papers, and even the most successful economists often have a hard time articulating a clear answer to that question. This is due to both (i) what economists read and (ii) how they read it. The goal of this paper is to teach readers how to write applied economics papers that will eventually be published in a peer-reviewed journal. The various components of an applied economics paper are discussed in as much detail as possible, roughly in the order in which they are tackled in the context of the typical research project.

Keywords: Applied Economics, Applied Microeconomics, Methodology

JEL Codes: A2, B4

*This paper a chapter in progress from a forthcoming MIT Press book; as such, reader comments are eagerly solicited. I thank Sara Savastano from the International Fund for Agricultural Development's Research and Impact Assessment division for commissioning the first version of this paper as a guide for her staff. I also thank Jeff Bloem, Liz Davis, Paul Glewwe, Eva-Marie Meemken, and Joe Ritter for useful comments and suggestions. All remaining errors are mine.

†Northrop Professor, Department of Applied Economics, and Director, Center for International Food and Agricultural Policy, University of Minnesota, 1994 Buford Avenue, Saint Paul, MN 55108, Web: http://www.marcfbellemare.com, Email: mbellema@umn.edu.

"I believe that what we become depends on what our fathers teach us at odd moments, when they aren't trying to teach us. We are formed by little scraps of wisdom."

— Umberto Eco (1988), *Foucault's Pendulum*.

1 Introduction

A good film makes you forget that you are watching a film. Similarly, a good research paper paper makes you forget that you are reading a research paper. The authors take you on a tour of what they have done: What they have asked themselves, how they have answered it, how they have made sure their answer was robust and what, if anything, we can learn from their results for policy or for business.

But just as a good film immerses you in the world it creates and makes you forget the various tropes and techniques used in the making of it, a good research paper is one that makes you forget to notice its overall structure as well as the various rhetorical devices employed by the authors.

How do you write a good research paper? In my experience, most research economists have spent too little time thinking about that question, and even the most successful economists would have a hard time articulating a clear answer to the same question.

This state of affairs is due both to *what* economists read, and to *how* they read it.

Regarding what economists read, the syllabus of most graduate field courses (e.g., behavioral, development, industrial organization, labor) usually consists of a "best of" for each topic covered—those papers that have shaped how people working in the field think and what they know about that topic. For instance, the syllabus for a graduate development economics course will almost surely include Foster and Rosenzweig (1995) and Suri (2011) in its reading list under the topic of technology adoption. In that literature, those two articles are widely understood to be among the best.

This applies mainly to more junior readers—the more senior one gets, the more one

has been exposed to bad papers by virtue of having refereed more papers—but reading only the best papers is a double-edged sword. To be sure, those are the papers we learn from the most when it comes to how our peers think about a given topic. At the same time, those papers tend to be the most polished ones—those nearest to perfection—on a given topic.[1] But it is rather difficult for one to learn what makes a paper good if all one ever reads is perfect papers. To carry the film analogy further, if all you ever watch are the films on the British Film Institute's (BFI) list of 50 greatest films of all time, and you never watch any bad (or even average) movies, it will be difficult for you to discover what actually makes those films on the BFI Top 50 good.

Regarding how economists read, the syllabi of most graduate field courses often lists so many articles as to cause graduate students to quickly develop a skill Mortimer Adler referred to as inspectional reading in his classic *How to Read a Book*. When reading academic papers, inspectional reading involves reading the introduction, looking at the methods and results, and (maybe) reading the conclusion before moving on to the next article on one's reading list. Reading papers that way is good to develop one's knowledge of the literature on a given topic, but it is hardly a recipe to learn how to write good papers.[2]

When they do know how to write a good research paper, economists have typically gleaned that information in unstructured, unintentional ways, often from throwaway comments by advisors and senior colleagues and coauthors. In other words, and as in the opening quote by Umberto Eco, they have learned it "at odd moments ... by little

[1] That said, if everyone reads the same papers and fails to read other papers, an unhealthy group-think dynamic settles in among the scholars working on a given topic which may lead to years of scholarly effort wasted on answering the wrong questions, or to attempts at answering questions using the wrong methods.

[2] Knowing is half the battle, so knowing that many readers will read your paper inspectionally can make you a more effective writer, because it forces you to put more thought into writing your introduction, your methods and results sections, and your conclusion. If you know that many of your readers are unlikely to bother with reading, say, your background section, you should state the most important facts of that section in your introduction. The greatest sin an academic writer can commit is the sin of omission, which consists in leaving important information out of a paper. The second greatest sin an academic writer can commit is one of commission, and it consists in forcing the reader to rifle through the paper hunting for a specific bit of information. The opportunity cost a reader's time is high, so the average reader is more likely to give up on reading a paper than on hunting for information. This is especially true when a relatively junior writer (e.g., a PhD student or an assistant professor) is writing to impress more senior readers (e.g., a faculty advisor, a journal editor, or journal reviewers).

scraps of wisdom," when those same advisors, colleagues, and coauthors were not trying to teach them.

The goal of this paper is thus to teach its readers how to write applied economics papers that will eventually be published in peer-reviewed journals.[3] To do so, the various components of a research paper are discussed in as much detail as possible, roughly in the order in which they are tackled in the context of a research project.[4]

The remainder of this paper is organized as follows. Section 2 lays out and discusses the structure of a typical applied economics article. Section 3 explains how an applied paper's theoretical framework should be presented, keeping in mind that readers of applied articles are rarely theorists. Next come the truly applied sections of an applied paper: Section 4 explains how to present data and descriptive statistics, Section 5 explains how to present the empirical framework, and Section 6 explains how to discuss the empirical results, including their limitations. Section 7 explains how to write a proper conclusion. Because the title, the abstract, and the introduction of a paper are subject to change until the author has figured out what her results are and what they mean, Section 8 then explains how to go about choosing a good title, writing a good introduction, and writing a good abstract. Section 9 discusses sections that are more often than not optional, viz. the literature review and background sections. Section 10 discusses the choice of where to submit an article for publication. Section 11 concludes.

2 Structure

Before producing any kind of work, it helps to know what the typical structure of such work looks like, and to write down a rough sketch of that structure. In its most abstract

[3] Though I am referring mainly to articles written so as to estimate a causal relationship of interest by reduced-form methods when I talk of "applied papers," much of the advice in this paper applies to other types of empirical papers, such as descriptive or structural papers.

[4] By "research project," I refer to the entire research process involved in answering a given research question, from idea to publication.

sense, the structure of the typical economics paper—applied or otherwise—is as follows:

1. Title
2. Abstract
3. Introduction
4. ...
5. Summary and Concluding Remarks
6. References

Depending on the type of paper one writes, the fourth item will change. Since this paper is geared toward writing applied papers, the structure above will typically be modified as follows:

1. Title
2. Abstract
3. Introduction
4. Theoretical Framework
5. Data and Descriptive Statistics
6. Empirical Framework
7. Results and Discussion
8. Summary and Concluding Remarks
9. References
10. Appendix

This structure is not set in stone. A frequent departure from the sequence above is when items 5 and 6 are switched around so that the Empirical Framework section comes before the Data and Descriptive Statistics section—something that is often a matter of taste, if not of expositional clarity. Similarly, a paper investigating a question that has

often been asked (e.g., the effect of stronger property rights on agricultural productivity) might not require a theoretical framework at all because the theory behind that question is well-known and is the stuff of textbooks. Or there might be a Background section after the Introduction where important contextual details are given that fit neither in the Introduction nor in the Descriptive Statistics section.

Some papers might require a major overhaul of that structure. Even in such cases, it helps to be familiar with the usual structure. A good analogy in this case is this: Before jazz legend John Coltrane ever thought of recording his more avant-garde compositions on albums like *A Love Supreme* (1965), he first learned how to operate within the (much more rigid and conventional) structure of bebop on albums like *Blue Train* (1957) and, before that, as a sideman on albums by Miles Davis.

What does this mean for an applied economist? It means that before you break the rules, you have to learn them. So before thinking of writing a paper whose structure is barely recognizable to the average reader, an applied economist should make sure to have written enough paper that follow the usual structure laid out above. In other words, freedom from structure tends to be a privilege granted to more experienced researchers, who have accumulated enough good will from their readers that they are allowed to bend the rules a little bit.

The remainder of this paper will not follow the structure just given. Though it would certainly be easier for me to write a paper whose subsequent chapter headings to follow that exact structure in order, it turns out that the structure in which we present our work in a research paper tends to be very different from the structure in which we actually do the work.

3 Theoretical Framework

Since the goal of applied economics is generally to answer questions of the form "Does x cause y?" or "If x increases by one unit, how many units does y change by?," most applied work in economics begins with an idea that comes from some theory of change.

Though many impact assessment reports look at whether and how a myriad of outcomes change in response to a specific intervention, the best research articles tend to focus on a single question (e.g., "What is the impact of having a land title on agricultural productivity?") or on the mechanisms behind a given question (e.g., "If land titles improve agricultural productivity, do they do so because land titles allow landowners to use their land as collateral?") Thus, a first decision has to be made about what causal relationship of interest a given article will focus on. In other words, the best applied economics article tend to be rather narrowly focused on a single question, and so you will almost inevitably have to leave some material on the cutting room floor.[5]

The question, then, is how to take your theory of change and convert it into a proper theoretical framework for an applied economics article. Here, there are two possible scenarios: (i) you are investigating a question for which the theory of change has already been studied by theorists, or (ii) you are investigating a question for which the theory of change has not already been studied by theorists.

In the first scenario, there are two options. The first option is to include a theoretical framework in the article by incorporating or adapting somebody else's theoretical framework lock, stock, and barrel. Though it often feels like a research paper has to innovate on all fronts, that is rarely the case. For applied papers in particular, what matters is that the research question, the empirical strategy, or both be novel. In most cases, it is perfectly fine to use someone else's theoretical framework—provided that you clearly state that you are doing so.

[5]This is not to say that there are not papers that look at several research questions at once. But at this point in time, what tends to be rewarded by the economics profession is answering a single, relatively narrow research question well.

A closely related option is to adapt somebody else's theoretical framework to suit your needs—say, by incorporating an additional variable, or by making additional assumptions to suit the needs of your application.

In the second scenario, when you are investigating a question for which the theory of change has not already been studied by theorists, you have to clearly state your theory of change. In some cases, this may require a formal theoretical model. In other cases, it is enough to merely present a verbal conceptual framework.[6] In all cases, your theoretical framework—be it mathematical or verbal—should start from the primitives and make the necessary assumptions to generate the result "x causes y through mechanism m," no more and no less.

One could write an entire book on how to write economic theory (and some have; see Thomson 2011), so nothing more will be said on this topic save for the following: Writing theoretical models in economics is an art form, and if you have not learned how to do it in graduate school, it is perhaps best to work with someone who has as a coauthor. When it comes to publishing applied economics article, better an informal, chatty conceptual framework than a bad formal theoretical model.

That said, even if your working paper includes an elegant theoretical model, it will sometimes happen that you will be asked by reviewers or by an editor to get rid of your theoretical model before your paper can be published, or to put said theoretical model in an appendix. If that happens, know that this is not uncommon, and do not get offended. In that sense, having a theoretical framework in your paper often only serves as a signal (in the Spence, 1973 sense) that you know what you are doing. This is especially true for job-market papers, which are used to show the breadth of their author's skills in addition to making a contribution to research.

[6] One possibility is to make a theoretical argument without math, in words, and to leave the math to an appendix. See, for instance, Sánchez de la Sierra (2020).

4 Data and Descriptive Statistics

After developing your theory of change, you have presumably gone in search of data to test the predictions of that theory. As with writing formal theoretical models, entire books have been written about the dos and don'ts of data collection (see Deaton 1997 or Glewwe and Grosh 2000 for survey data; see Gerber and Green 2012 or Glennerster and Takavarasha 2013 for randomized controlled trials), so this section will not discuss where the data come from to assume instead that you have them. Rather, this section will focus on how to present your data in the context of an applied economics article.

The best Data and Descriptive Statistics sections answer all of the reader's questions about the data themselves. Specifically, a good Data and Descriptive Statistics section first discusses where the data come from, when they were collected, by whom, how the observations that compose the sample were chosen for inclusion (i.e., the survey methodology, or how regions, communities, households, individuals, etc. were all chosen), what population the sample is representative of, what the target sample size was and how that sample size was determined (e.g., via power calculations), what the actual sample size is, what the nonresponse rate was, what the attrition rate is in case the data are longitudinal, how missing data were dealt with (e.g., whether observations were simply dropped, or whether some values were imputed and, if so, how the imputation was done). Broadly speaking, the information presented here allows the reader to judge the external validity of the results contained in a paper (and sometimes their internal validity, as is the case when the data suffer from attrition), or how those results might be used for out-of-sample predictions.

After presenting those basics, a good Data and Descriptive Statistics section introduces all the variables used in the paper (and no variable not used in the paper) by precisely and concisely explaining what they measure, and how they do so. For instance, in many rural areas of developing countries, people derive their income from many difference sources. So if "income" is included in the analysis, the reader needs to be told what the various

income sources are.

This may seem tedious—and if it seems tedious to you as writer, you can imagine what it seems like to the reader—but it can nevertheless contain crucial information. For instance, an age-old question in the literature on agricultural development, and one on which I have done quite a bit of work is whether participation in agricultural value chains (via contract farming, as a grower) makes participating households better off (see Bellemare and Bloem 2018 for a review). This is usually assessed by regressing a measure of household income (as a proxy for welfare) on a dummy for whether the household participates in contract farming. Without knowing what the components of household income are, however, it is impossible to know whether it includes income from contract farming (in which case there is an obvious reverse causality problem) or not (in which case reverse causality is much less of a problem).

The good news is that it is relatively easy to present that information when one has access to the survey questionnaires that were used to collect the data, which is almost always the case. Moreover, one way of presenting that information optimally is by creating a table of variable descriptions, where each line is a specific variable retained for analysis, where the first column gives the name of that variable (and the unit of measurement in parentheses), and where the second column gives precise measurements. Figure 1 shows one such table. This allows presenting a lot of required but tedious information in a compact manner, which minimizes reader discontent: Those who want to know all there is to know about the data can read the table, and those who do not can just skip it to focus instead on variable names.

After presenting the foregoing, it is time to present and discuss descriptive statistics. Here, whereas it used to be sufficient to simply present a table of means and standard deviations, it has become practically necessary in cases where the variable of interest (i.e., the treatment variable) is composed of a small number of categories to show the results of balance tests, viz. tables where each line is a variable retained for analysis, where means

Data Description for Selected Variables

Variable	Description
Dependency ratio	Percentage of individuals under 15 and over 64 within the household
Assets (100,000 ariary)	Sum of the values of the household's assets (i.e., animals, house, television, radio, car, and bank account balance) and agricultural equipment (i.e., hoe, harrow, cart, plow, tractor, and small tractor)
Income (100,000 ariary)	Sum of the proceeds from animal sales, agricultural and nonagricultural wages, and proceeds from leases of cattle and equipment
Liquidity constraint dummy	Dummy for whether the household is liquidity constrained
Plot size	Area covered by the plot in ares (1 are = 001 ha = 100 m^2)
Plot value (100,000 ariary)	Price expected by the landowner if she were to sell her plot
Formal title dummy	Dummy for the presence of a formal title
Relationship length	Number of years the landlord and tenant have been contracting with one another
Kin dummy	Dummy for a contract signed between kin
Tenant introduced by kin	Dummy for a contract signed with a tenant whom the landlord met through a member of her extended family
Introduced by other than kin	Dummy for a contract signed with a tenant whom the landlord met through someone who is not a member of her extended family
Tenant is friend	Dummy for a contract signed with a tenant who is a friend of the landlord
Tenant chosen for his wealth	Dummy for whether this particular tenant was chosen because of his wealth
Tenant chosen for his honesty	Dummy for whether this particular tenant was chosen because of his honesty
Tenant chosen for his ability to bear risk	Dummy for whether this particular tenant was chosen because of his ability to bear risk
Tenant chosen to return a favor	Dummy for whether this particular tenant was chosen because the landlord wanted to return a favor
Time spent looking for a tenant	Number of days spent looking for a potential tenant
Other potential tenants considered	Number of other potential tenants considered when looking for a tenant

FIGURE I: Example Table of Variable Descriptions from Bellemare (2012).

and standard errors are shown conditional on treatment status, and wherein one assesses whether the mean of each variable systematically differs across treatment statuses by reporting p-values for a test of difference in means. Though the textbook example involves two treatment statuses—treatment and control—it is increasingly common for studies to include more than two treatment arms, and so any meaningful balance test must be reported for each pairwise comparison of means. With two treatment arms, this means (i) treatment 1 versus control, (ii) treatment 2 versus control, and (iii) treatment 1 versus treatment 2.

With experimental data, the idea behind such balance tests is to show the reader that randomization was done properly. With observational data, where we would not expect the data to be balanced, the idea behind such balance tests is to assess how unbalanced the data are—an idea which comes from the matching literature (Morgan and Winship 2015). With perfect random assignment across treatment and control groups, there should be fewer than 1 in 10 pairwise comparisons differing at less than the 10 percent level of statistical significance, fewer than 1 in 20 pairwise comparisons differing at less than the 20 percent level of statistical significance, and fewer than 1 in 100 pairwise comparisons different at less than the 1 percent level of statistical significance. In cases where pair-

wise comparisons return too many systematic differences, one should ideally control for the relevant covariates in a regression or matching context when estimating treatment effects.[7]

Beyond the usual table of means and standard deviations and one or more tables showing the results of balance tests, a good Data and Descriptive Statistics section can also be used to explore the data nonparametrically by showing kernel density estimates of the relevant variables (i.e., outcome and treatment variables at a minimum, but also controls suspected to be the source of treatment heterogeneity) when they are continuous, histograms of the relevant variables when they are categorical, or cross-tabulations (i.e., two-by-two tables) in cases where both the treatment and the outcome are both binary.

When writing a Data and Descriptive Statistics section, there are a few important mistakes you should avoid making. The first such mistake is for the write-up to present a bland enumeration of means. If a gender variable is merely used as a control in the analysis, there is little use to stating in the text that "37.4 percent of respondents are female"when the reader can look that up for herself; the only variables that typically deserve discussion here are the outcome and treatment variables, any variable that is used for identification (e.g., an instrumental or forcing variable), or anything that really stands out. Generally, a good rule of thumb is to keep the discussion of the descriptive statistics to a few sentences.

The second such mistake is the use of the past tense in discussing the data and descriptive statistics. The example above stated how "37.4 percent of respondents are female," and not how "37.4 percent of respondents were female." Scientific communication in English is more effective when using the present tense to discuss one's data or results, and

[7]Comparing means across treatment and control groups is the strict minimum when it comes to testing for balance. A more restrictive approach consists in running a joint test (i.e., F-test) of whether all means are simultaneously the same across groups. Another more restrictive approach consists in conducting tests of equality of distributions for pairwise comparison using a Kolmogorov-Smirnov test or using Bera et al.'s (2013) smooth test for equality of distributions.

just as the passive voice should be avoided, so should the past tense, except when summarizing and concluding.

Lastly, another mistake is to present numbers that either have too many decimal places because they are too small (usually, three decimal places is more than enough and, at any rate it is always possible to rescale a variable to make its magnitude fit with that of the other variables) or to present numbers that are hard to interpret in tables, such as $1.37e+8$, or anything other than units readers are used to deal with (for instance, it is always possible to express a dollar amount in thousands or hundreds of thousands if need be). In other words, even if the empirical work regresses the logarithm of income on the treatment variable, the table of descriptive statistics should report the mean of the income level, not the mean of the log of income.

Ultimately, although a lot of what goes in a Data and Descriptive Statistics section might seem like useless posturing, as stated before, a good Data and Descriptive Statistics section should allow the reader to form reasonable expectations about the sign and the magnitude of the causal effect of interest, and to get an idea of how that effect is likely to vary across a given conditioning domain.

5 Empirical Framework

After discussing the data and presenting descriptive statistics, you normally turn to discussing your empirical framework, i.e., the research design you use to empirically answer your research question.

An empirical framework consists of two related components: (i) an estimation strategy (i.e., what is estimated, how it is estimated, and how statistical inference is conducted), and (ii) an identification strategy (i.e., what feature of the data allows making a causal statement or, if that is not possible, how we know we are getting close to making such a statement).

5.1 Estimation Strategy

An estimation strategy typically consists of the equations to be estimated in an effort to answer a research question. Though it may be possible for a savvy reader to recover the estimated equations in a paper by looking at the tables in a paper, that is not always possible. At any rate, the amount of work a reader should have to do should be kept to a minimum, so presenting the equations to be estimated is very much the norm.

Ideally, those equations will be as parsimonious as possible. Although a regression might include 10 to 15 control variables, it is best to put all of those into a vector x of control variables. What deserves its own variable in an equation to be shown in an estimation framework? For starters, the dependent variable (labeled y) should be included along with the treatment variable (labeled either D or T), the (vector of) controls (labeled x), an intercept term (labeled α), and the error term (labeled ϵ).

Here are, in no particular order, a few other norms that are best followed:

- All variables should have the proper subscripts, usually labeled i, j, k, ℓ, etc. from the smallest (e.g., individual) to the largest level (e.g., region).
- Latin letters should denote variables. Greek letters should denote coefficients.
- If the estimation strategy sub-section features several different specifications of the same equation, coefficients should also have subscripts. In other words, one should not re-use estimand notation. If β is used to denote the coefficient of interest in a regression of y on D, it should not be re-used to denote the coefficient of interest in a regression of y on D and x as well—the two estimands being different, the notation used to denote them should also be different. This is best done by adding numerical subscripts to each coefficient, so that in the former specification, the coefficient on D would be denoted β_0 and in the latter, β_1. Or it can be done by adding letter subscripts to each coefficient, so that for example β_r and β_s can respectively refer to reduced-form and structural estimates of the same coefficient.

- The estimation strategy sub-section should also specify what estimation method is used to estimate each estimable equation. We are generally interested in $E(y|x)$, but $E(y|x)$ could be estimated in a number of different ways parametrically, semiparametrically, or nonparametrically. With a binary outcome variable, the reader needs to know whether a linear probability model, a probit, or a logit is estimated. In cases where it is ambiguous, the estimator (e.g., least squares, maximum likelihood, or generalized method of moments) also needs to be specified.

- After presenting the estimable equations, it is a good idea to discuss the relevant hypothesis tests. In a regression of the form

$$y = \alpha + \gamma D + \beta x + \epsilon, \tag{1}$$

  for instance, the relevant hypothesis test would be of the form $H_0 : \gamma = 0$ versus $H_A : \gamma \neq 0$. Here, note that a hypothesis test always tests for an equality sign. So while a paper might test the (theoretical) hypothesis that changing D from 0 to 1 causes an increase in y (and further assesses by how much y increases in response to the change in D), statistically speaking, the same paper tests the (null) hypothesis that the association between D and y is not statistically significantly different from zero.

- The estimation strategy sub-section also needs to discuss inference, meaning whether and how the standard errors are robust (and if so, robust to what; it is not enough to say that the standard errors are robust if the Huber-sandwich-White correction is used, but it is warranted to say that they are robust to heteroskedasticity), whether and how they are clustered (and if so, at what level and why; see Abadie et al. 2017 for a primer), and whether sampling weights were used to bring the sample closer to the population of interest (and if so, how they were constructed; see Solon et al. 2015 for a primer).

5.2 Identification Strategy

After showing and discussing what equations are estimated, there needs to be a discussion of how the coefficient pertaining to the causal relationship of interest is identified.

The term "identification" has gone through several meanings over time (Lewbel, 2019). For better or for worse, the term more often than not refers to *causal* identification nowadays in applied papers. What is causal identification? In a few sentences, it refers to situations where a coefficient is more than just a partial correlation between the dependent variable y and some variable of interest D, and where the estimated coefficient instead reflects a relationship that is causal.

Although an unbiased coefficient estimate implies an identified—that is, causally identified—coefficient estimate, the converse is not true. Indeed, there are situations where one knows a coefficient to be biased, but where a statistically significant coefficient estimate can still be used to denote a causal relationship.

For instance, imagine that D is continuous and randomly assigned and you have data on an outcome variable y and a vector of controls x. Imagine further that subjects perfectly comply with D, so that you can in principle estimate an average treatment effect, but that D is measured with error—when the data for D were entered, they were entered with some degree of error. With classical measurement error on D (i.e., with random mistakes in the entering of D), we know the treatment effect will be biased toward zero because of attenuation bias.[8] In such cases, when $H_0 : \gamma = 0$ is rejected, we can still say that we have found evidence of a causal effect of D on y, adding the caveat that that effect is biased toward zero (or, alternatively, that we have estimated the lower bound in absolute value on the true effect). Sometimes, the same can be said with systematic measurement error on D, as there are cases where you know the estimate $\widehat{\gamma}$ of γ is biased toward zero because of systematic measurement error, but that is much less common.

[8] As a colleague noted, if D is binary, any measurement error cannot be classical, for it the observed D will be negatively correlated with the true value of D.

If you are fortunate enough (i) to have experimental variation in your treatment variable, and (ii) balance tests suggest the experimental assignment of observations to treatment and control groups was truly random, your identification strategy section can be mercifully short, as your results are causally identified by virtue of experimental assignment. In other words, you can estimate what Pearl (2009) denotes $E(y|do(x))$, i.e., the (causal) effect of treatment x on outcome y.

If you have (i) experimental variation in your treatment variable but (ii) balance tests suggest the experimental assignment of observations to treatment and control groups was not truly random, your identification strategy section can also be short, as you only need to explain how you will add in control x on the right-hand side of your equation of interest to help rectify the situation, but only somewhat as unobservables are also likely to be unbalanced when the observables are unbalanced.

If you do not have experimental variation in your treatment variable, there is yet more work to be done. In the interest of brevity, this paper cannot and will not provide a deep dive into causal identification with observational data (for a complete introduction, see Morgan and Winship 2015). There nevertheless are certain things that can be discussed as being necessary items in any good identification strategy section:

- Explain intuitively why your results have a shot at causal identification. Practically speaking, this means that you have to tell your reader why your results bring us closer than ever before to making a causal statement about the relationship of interest. In the best-case scenario, this will be because you have a research design (e.g., a strictly exogenous instrumental variable such as a lottery) which clearly allows thinking of treatment as if it were randomly assigned. In less-than-ideal scenarios (e.g., an instrumental variable that is only plausibly exogenous; cf. Conley et al. 2012), you need to explain why, even though your research design does not yield clean and clear causal identification, your results are the best in the literature.[9]

[9]This presumes that your research design *is* the best thing out there. In cases where your research design

- Discuss in turn the three following sources of statistical endogeneity:[10] (i) reverse causality, (ii) unobserved heterogeneity, and (iii) measurement error, explaining whether each of those sources of statistical endogeneity is a concern in your application, how it is dealt with in your application. Here, if there are issues, admit to them, and explain how they might bias your estimate of the coefficient of interest. Do not lie about what your paper can and cannot do!

- Once that is done, there is one more source of problems to be considered, viz. violations of the stable unit treatment value assumption (SUTVA). What SUTVA means is specific to each application, but in short, if you observe the effect of a treatment D_{it} on outcome y_{it}, where i denotes an individual unit of observation and t denotes a time period, it has to be the case that the value of D_{it} does not affect the value y_{-it}, $y_{i,-t}$, or $y_{-i,-t}$. In other words, there cannot be any spillovers from one unit being treated to another unit's outcome, and there cannot be any spillovers from one unit being treated at a given point in time to that same unit's outcome in the future, nor can there be any spillovers from one unit being treated at a given point in time to other unit's outcome in the present or in the future. The SUTVA can be extremely difficult to satisfy. That said, one can often test for SUTVA violations; see Burke et al. (2019) for an example of a paper where the authors deal with SUTVA violations very well.

- Again, because this is extremely important: If your results are not causally identified, *do not lie* about what they can and cannot do! And generally, do not make claims that are not backed up by your research designs of your results, no matter

is second- (or third-, or n^{th}-) best, unless you significantly improve on external validity, you will need to adjust your set of target journals downward.

[10] I talk explicitly of statistical endogeneity—what makes $Cov(D, \epsilon|x) \neq 0$—because many research economists still confuse theoretical and statistical exogeneity. Theoretical exogeneity is when a given variable is determined outside of a given theoretical framework (e.g., prices and income in the typical utility-maximization problem). Statistical exogeneity is when $Cov(D, \epsilon|x) = 0$ in the regression framework we have been considering. Though the two share the same "exogeneity" and "endogeneity" terminology, there is little overlap in their respective meanings. It is a poor applied economist who says her results are causally identified because her treatment variable is (theoretically) exogenous!

how much you wish those claims to be true.[11] Editors and reviewers would much rather deal with manuscripts wherein the author candidly admit to the limitations of their findings than with manuscripts wherein the authors try to deceive the reader. In plain English: The former kind of manuscript has a much better chance of not being rejected than the latter.

6 Results and Discussion

The section of an applied economics article that discusses the paper's findings is obviously the most important section of the paper. Somewhat paradoxically, it is perhaps also the least read section of a paper: After a reader has read the title, the abstract, the introduction, looked at a few tables, and maybe looked at the empirical framework section to answer her lingering question, your reader knows whether she can trust you and your findings, and she is often only interested in your core finding. Only reviewers and critical readers (e.g., graduate students reading your paper for a class, should your article end up on someone's syllabus, or for their dissertation) will read the entirety of the results section. Nevertheless, results sections have their own structure, which is discussed below.

6.1 Order of Results

There is a certain logical order in which results should be presented. Typically, results progress from most parsimonious (e.g., a simple, bivariate regression of y on D) to least parsimonious (i.e., a regression of y on D and a full set of control variables x). With experimental variation in D, this is not as useful as with observational variation in D. In the former case, adding controls on the right-hand side of the equation of interest will in

[11] The proper conduct of social science dictates that one should not wish that one's findings should go in a particular direction. Research economists should strive to be first and foremost scientists, and not advocates for a certain position.

principle not change the sign and the magnitude of the estimated treatment effect. Rather, it will only make the estimate of the treatment effect more precise (i.e., it will reduce the standard error around it).[12]

In the latter case, where one cannot assume that $E(y|x) = E(y|do(x))$, the most-to-least-parsimonious approach is one first step toward assessing the robustness of one's results: If the sign and the magnitude do not change much or at all as one adds in control variables on the right-hand side, this suggests that one's results are already somewhat robust. This is in the spirit of Altonji et al.'s (2005) approach to robustness (although Oster 2019 critiques Altonji et al. 2005 and suggests a new method aimed at assessing how important unobserved heterogeneity is in a given application).

6.2 Robustness Checks

After presenting the core results in a paper, it is time to turn to robustness checks. Though there was a time where it was sufficient to present one or two tables of empirical results to convince the reader that there was a "there" there, times have changed, and as a consequence of the Credibility Revolution (Angrist and Pischke, 2010), standards of evidence are considerably higher than they were in the early to mid-2000s. Authors now have to work hard to convince readers that their results were not cherrypicked, which means that establishing the robustness of a finding involves its own song and dance.

In many cases, the outcome we are interested in has more than one measurement. "Welfare," for instance, can be measured in a number of ways: household income, household income per capita, household income per adult equivalent, household consumption expenditures, household consumption expenditures per capita, household consumption expenditures per adult equivalent, subjective well-being of the respondent, etc. If you have access to all seven of those measures of "welfare," one first step toward establishing

[12] A colleague who has run numerous RCT notes that in his experience, adding controls tends to have almost no effect on the standard errors, with the only exception being when the baseline value of the outcome is added as a control variable in an ANCOVA setup (McKenzie 2012).

that your result is robust might be simply to re-estimate your core equation for all seven of those measures, showing that the result holds across all seven of them.

Similarly, you may have different measures of the treatment variable. In most randomized controlled trials, there is one (and only one) treatment variable (unless there are several treatment arms, and unless those treatment arm are interacted). But with observational data, it might be possible to look at different measures of the treatment variable. In the contract farming literature, for example, one can look at whether a household participates in contract farming (i.e., contract farming at the extensive margin), but one could also look at the proportion of one's crop acreage that is under contract (i.e., contract farming at the intensive margin).

Now imagine that you have those two measures for the treatment variable, and the aforementioned seven measures for the outcome variable. This allows estimating 14 different specifications of the core equation of interest! If the finding holds for each and every one of those specifications, that goes a long way toward establishing that a finding is robust.

One can also check for robustness by conducting placebo and falsification tests. In the former case, a "fake" treatment (i.e., a variable that is correlated with the treatment, but which presumably does not cause the outcome) is used in lieu of the actual treatment. In the latter case, a "fake" outcome (i.e., a variable that is correlated with the outcome, but which presumably is not caused by the treatment) is used in lieu of the actual outcome. In both cases, robustness comes from the lack of a statistically significant finding, since a statistically significant finding hints at the fact that the core results might be spurious.

Yet another kind of robustness check comes in the form of looking at different estimators. Most applied economics articles, for instance, rely on some linear, fully parametric regression. If the treatment is continuous, it might be useful to estimate specifications that allow for a more flexible functional form (e.g., a restricted cubic spline), which would allow one to determine whether the relationship between y and D is generally

monotonic. Very often, robustness checks of this kind are where modest methodological contributions—a paper's third contribution, as listed in the introduction—come from.

6.3 Treatment Heterogeneity

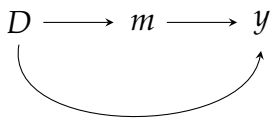
It is almost never the case that the treatment effect we are interested in estimating is homogeneous across the population of interest. After assessing the robustness of your results, you may be interested in looking at whether the treatment varies for various sub-groups (e.g., men vs. women, rural vs. urban, black vs. white, by income quintile, etc.) This section is where this is assessed. Keeping with the contract farming example, suppose you were interested in whether the impacts of contract farming differ between male and female respondents. This alone would bring the number of estimated specifications up to 28 (i.e., seven measures of welfare, two treatment measures, and male vs. female respondents). From this, it is rather it easy to see why the average applied paper is now typically 50 pages—if not longer.

One good thing about exploring treatment heterogeneity is this: Very often, doing so can salvage a null finding (i.e., an effect that is statistically insignificant) because average effects can mask a tremendous amount of heterogeneity. One of my students, for instance, was interested in looking at the effects of introducing soup kitchens on food insecurity in Mexico. Looking at the whole sample yielded nothing interesting, as her estimated effects were not statistically significantly different from zero. It was only when a colleague suggested that she split up the sample in income quintiles that she found that soup kitchens were associated with a decrease in food insecurity—but only for the poorest income quintiles. So before calling it quits, saying that an intervention or treatment has had "no effect," and abandoning an entire research project, it is well worth thinking about whether the treatment effect might be heterogeneous, and whether said heterogeneity is of interest for policy or business.

6.4 Mechanisms

As a result of the Credibility Revolution (Angrist and Pischke, 2010), applied microeconomists have been answering questions of the form "Does D cause y?" or "What is the effect of D on y?" first and foremost.

FIGURE II: Direct and Indirect Mechanisms



In recent years, however, much has been written in the quantitative social science literature about how to test for whether a given variable m is a mechanism whereby some other variable D causes some outcome y—what is called mediation analysis—and this remains a very active area of research.

To start with, readers should read the paper by Acharya et al. (2016), which develops a method which, under certain assumptions, allows determining for the directed acyclic graph (Pearl 2009) in Figure I (i) what is the indirect effect of D on y, i.e., $D \longrightarrow m \longrightarrow y$, (ii) what is the direct effect of D on y, i.e., $D \longrightarrow y$, and (iii) whether m is the only mechanism whereby D causes y, or whether there are multiple such mechanisms.

A good section on mechanisms does its best to investigate potential mechanisms. In the best-case scenario, this involves a proper mediation analysis. In many cases, this means doing what one can do with the data at hand, such as presenting descriptive (i.e., not causally identified) regressions or correlations. In other cases, this means simply admitting that there are some mechanisms one cannot test for, not even with imperfect proxies. When anything but the ideal is feasible, you should clearly explain why you cannot test for specific mechanisms to leave no doubt in your readers' minds that you have thought about the question "How does D cause y?"

6.5 Limitations

A good empirical results section should be honest about what it can and cannot do. Though this is often discussed quickly in the conclusion, it should be discussed more fully in a separate sub-section of the estimation results section.

What limits one's results? Typically, limitations come in three varieties. First and foremost, internal validity may be limited. In other words, one might not be able to make a causal statement, but instead only get close to doing so relative to the literature. For instance, your instrumental variable might only be plausibly exogenous, but not strictly so. This would be a good time to remind the reader that this is so (and you should also have assessed the effects of departures from strict exogeneity in the robustness checks section using the various methods laid out in Conley et al. 2012).

Second, external validity may be limited as well. This is often the case with lab or a lab-in-the-field experiments,[13] or with randomized controlled trials. Or you may have a strictly exogenous instrumental variable, but it is not entirely clear who the compliers and defiers are, and so who the local average treatment effect applies to is a nebulous subset of the sample.

Lastly, the variables you use as your treatment or your outcome variable might only be proxies for what you are truly interested in. For instance, though you may be interested in looking at whether economic shocks push people to commit suicide, data on suicides may not be available (or suicides may be significantly under-reported), and so you might have to resort to using mortality rates instead.

6.6 Tables

Before closing out this section, I would like to discuss some miscellaneous pieces of advice regarding tables of empirical results. In no particular order:

[13] Lab-in-the-field experiments are lab experiments that are conducted with "real" subjects (e.g., farmers or managers) in the field, outside of the experimental lab.

- The titles of your tables should be self-explanatory: "OLS Results for the Effect of Participation in Contract Farming on Household Income," or "OLS Results for the Effect of Participation in Contract Farming on Income by Gender." The titles should thus tell us what is being estimated (e.g., OLS), what the causal relationship of interest is (i.e., the effect of participation in contract farming on household income), and what subset of your sample, if any, it applies to (i.e., male and female respondents separately).

- Coefficient estimates and standard errors should be reported with the same number of decimal places throughout your tables—usually two or three.

- Some people like to omit control variables, preferring instead to include a line that says "Controls? Yes" in the second (i.e., bottom) half of the table. Though this is fine to save space in a published article, a working paper should show everything to the readers (especially the reviewers and the editor). The obvious exception is for individual, household, or community fixed effects, of which there are usually too many to list. If you must include a line at the bottom that says "Controls? Yes," make sure the notes to the table (i.e., right under the table) include a detailed list of which controls are included—a careful reader will want to know whether you condition on colliders or include as control a variable that lies on the causal path between the treatment and outcome variables.[14]

- The last lines of the table should list the number of observations, the R^2 (I prefer the usual R^2 to the adjusted one, because this tells me how much of the variation in y is explained by the variables on the right-hand side, without any arbitrary correction for the number of observations and parameters), maybe the results of a test of joint significance of the variables on the right-hand side, and various lines indicating

[14] A collider is a variable caused by two separate, possibly unrelated variables. Conditioning on a collider or on a variable that lies on the causal path between the treatment and outcome variables is problematic because it introduces bias (Morgan and Winship 2015).

which controls are included (e.g., state fixed effects, a linear time trend, year fixed effects, state-specific linear trends, state-specific quadratic trends, region–year fixed effects, and so on).

- Finally, the notes to the table should present all symbols for statistical significance (typically, * for statistical significance at less than the 10 percent level, ** at less than the 5 percent level, and *** at less than the 1 percent level; none should be omitted for completeness and transparency), and additional symbols if necessary. For instance, you may have adjusted your p-values for multiple comparisons, bootstrapped your standard errors, or done some randomization inference, all of which would lead to different inferences and critical levels of statistical significance, in which case you might use the symbols †, ††, and ‡ ‡‡ to denote significance at less than the 10, 5, and 1 percent level for this additional version of the standard errors.

- Present estimation results for the same estimation sample. That is, as the number of control variables increases, the sample size is nonincreasing due to missing variables. If the sample size decreases as you throw controls on the right-hand side, this involves and apples-to-oranges comparison (different estimation samples are representative of different populations). Instead, take your smallest sample size (as dictated by missing observations) and use that sample for all specifications.

- For variable names, use plain English words like "Years of education," "Age squared," and "Female" and not Stata or R codenames like "Edu," "AGE_2," or "SEX."

- Ultimately, it always helps to put yourself in your reader's shoes,[15] and the right question to ask yourself (or a friend who owes you a favor) is this: When given only the tables, can one write down the exact regression that was estimated? Or is one

[15]This can be difficult, which is why the best way to learn is to review as many papers for journals as possible. Many scholars—economists, in particular—see refereeing as an unfortunate tax they need to pay in order to get their own papers reviewed and published. Unlike a tax, however, there is almost always something to be learned from refereeing, and from refereeing bad papers in particular.

left with more questions than one has answered after looking at the tables?

7 Summary and Concluding Remarks

Many economics papers title their conclusion "Summary and Concluding Remarks," which is a pretty good indication of how a conclusion should proceed. What I learned in high school was that a good conclusion should have two main parts: (i) a summary of what you have spent the several pages before the conclusion doing, and (ii) the way forward.

The following guidelines should help cut on the transaction costs one faces when writing a conclusion by providing a roadmap. Strictly speaking, a conclusion should be structured as follows:

- *Summary*. You have surely heard that when writing a research paper, you should "tell them what you're going to tell them, tell them what you want to tell them, and tell them what you just told them." Writing this part of a conclusion is tedious—you have just spent 40 or more pages telling them—but it needs to be there, and it needs to be different enough from the abstract and the introduction. This does not mean this part must say something new; it just needs to be different enough. If possible, tell a story.

- *Limitations*. Some people like to have a "Limitations" section at the end of their results section; I like to have that myself, as discussed above. But even then, the conclusion should (re-) emphasize the limitations of your approach.

- *Implications for Policy.* Presumably, your work has some sort of implication for how policy is made in the real world. This will not always be the case—some papers make a purely technical point, or a point that is only ancillary when it comes to making other policy-related points—but I would guess that since you are reading a paper about writing applied economics papers, there is a high likelihood that what

you are working on has some real-world implications. Discuss what those implications are. Do not make claims that are not supported by your results. Try to assess the cost of what you propose in comparison to its benefits. You can do so somewhat imperfectly (this is where the phrase "back-of-the-envelope calculation" most often comes up in applied economics papers), since the point of your work was presumably about only one side of that equation—usually the benefits of something, sometimes its costs, but rarely both. In two or three sentences, identify the clear winners and losers of a given policy solution. Briefly discuss its political feasibility. How easy or hard would it be to implement?

- *Implications for Future Research*. Finally, your work is not perfect. Your theoretical contribution could be generalized or broadened by relaxing certain assumptions. Your empirical contribution could probably benefit from better causal identification for better internal validity. Even with a randomized controlled trial (RCT) with perfect compliance and a perfect average treatment effect estimate, you are likely to have some treatment heterogeneity that is not accounted for, or you might want to run the same RCT in additional locations for external validity. If you are writing a follow-up paper, this is a good place to set the stage for it.[16]

8 Title, Abstract, and Introduction

The title, abstract and introduction of a paper are, in order, the three most important marketing tools for any paper. This probably is doubly true for applied economics papers, wherein authors rarely advance the frontier of knowledge theoretically or methodologically. Readers are probably more likely to put up with a bad title, a poorly written abstract, a meandering introduction—or all three—if they know that a paper will change

[16]That said, it is generally fruitless to look for research ideas in the conclusions of the papers you read, because those conclusions will either list (i) ideas that are so difficult to execute that the authors of the papers you are reading did not think it was worth exploring them, or (ii) ideas that those same authors are already working on.

their understanding of how the world works, or if they know that it will give them new tools they can use in their own research. Consequently, the following sub-sections focus on these components of an applied economics paper.

8.1 Title

It is difficult to pinpoint exactly what makes a good title. Much like US Supreme Court Justice Potter Stewart famously said of hard-core pornography in *Jacobellis v. Ohio*, when it comes to a good title, "I know it when I see it." Colleagues who tend to publish in general-science journals like *Science*, or *Proceedings of the National Academy of Sciences* often insist that we should state our results in our titles. While that may be true for those papers we submit to those general-science journals, titles follow a certain norm in economics which is best followed if you want your papers to look like they fit in.

To that end, it is perhaps easier to define what makes for a bad title. For starters, any title which emphasizes the technique you are using is sure to turn off most readers, unless you develop said technique. Spare your would-be readers titles—especially subtitles—of the form "A Semiparametric Investigation" or "Nonparametric Evidence from [Your Context]."Long titles also tend to make readers not want to read your paper. That is probably why there is an inverse relationship between the length of a paper's title and the number of times that paper gets cited (Letchford et al., 2015).

For an applied economics paper—that is, a paper that asks an empirical question of the form "What is the effect of D on y?"—it is safest to go with a title of the form "The Impacts of D on y: Evidence from [the Context You Are Studying]." My own work features a number of examples of this: "The Welfare Impacts of Contract Farming," "The Welfare Impacts of Commodity Price Volatility: Evidence from Rural Ethiopia," or "The Welfare Impacts of Rising Quinoa Prices: Evidence from Peru."[17]

[17] From, that list, it should be obvious that my coauthors and I tend to care quite a bit about welfare. For an applied economist, looking at how certain things affect people's welfare is always a good place to start.

A variant on this theme is a title of the form "D and y," with or without the subtitle after the semicolon. My own work features a number of examples of that, too: "Contract Farming and Food Security" or "Farmers Markets and Food-Borne Illness." In both cases, there is no subtitle after a semicolon. In both cases, my coauthor and I made the conscious decision to go with a big, bold title that did not specify the context, and to be quite honest, in both cases, we were surprised reviewers did not demand that we narrow down our title by specifying "Evidence from Madagascar" in the former case an and "Evidence from the United States, 2004-2013" in the latter case. It sometimes pays to go big and bold.[18]

There is also the question of whether you should be cute or funny—for lack of a better term, let's refer to either as "clever"—in your title. If you are going to have a clever title, make sure it appeals to as many people as possible, and make sure it actually makes sense. What often works here is common sayings (e.g., "Fair Enough?"for a paper on whether the international trade of a given commodity is fair), dicta, proverbs, biblical passages (e.g., "As You Sow, So Shall You Reap"for a paper on whether participation in agricultural value chains increases the welfare of participants), or titles of famous films ("Star Wars" for a paper looking at p-hacking, "Look Who's Talking" for a paper looking at the effect of the intrahousehold allocation of cell phones), books, or TV shows (e.g., "All in the Family"for a paper on bequest motives in East Africa).

I recently handled a manuscript whose title began with "One Size Does Not Fit All." Because (i) saying that out loud is a mouthful, and (ii) the actual, familiar expression is "one size fits all," I asked the authors to reformulate the first half of their title as "One Size Fits All?" Ultimately, if you are going to take the clever route, make sure the cleverness is warranted, and that the clever part of your title fits your paper like a glove.

Lastly, I recall reading in grad school that your paper titles should never start with "On," because it makes it sound as though you are merely writing a comment on someone else's paper. That is pretty bad advice: Hirshleifer's (1956) paper titled "On the Theory

[18]That, or those are examples of the privilege that come with being "senior" in a given profession.

of Transfer Pricing" has been cited over 900 times, and it should be pretty clear to anyone reading that paper that Hirshleifer is making an original point. From personal experience, I have written two such articles: one won an award, the other has garnered over 100 citations in less than two years, and no one seems to have mistaken either for a comment on anyone else's work. I therefore think it is perfectly fine to write a paper "on" a given topic, especially if you are the first to take a serious look at said topic.

8.2 Introduction

I list the introduction second after the title and before the abstract because writing an abstract is much easier once the introduction has been written. The best way to write an introduction is to follow Keith Head's (2020) introduction formula.

I remember coming across that formula while in graduate school (and so it has been around since at least 2006) and thinking "I know how to write, I don't need this." Do not make that mistake. Even if you (think you) know how to write, the beauty of Head's formula is that it removes all uncertainty as to the order in which an introduction's sections should be presented.

The formula—and really, all credit goes to Keith Head for articulating it—is as follows:

- *Hook*. A good introduction starts with a good "hook," i.e., something that grabs the reader's attention and makes her want to keep reading. Here, the closer one can get to the reader, the better. Likewise, the broader one can go, the better. Bad hooks tend to appeal to the literature: "A long literature in economics has looked at ..." If that is the case, do you really want to make it any longer? Good hooks tend to relate to the real world: A lot of the food we buy at the grocery store is grown in the context of long value chains. What does the first link in that value chain look like? What does participating in those value chains do for the people who actually grow the food we eat? The hook should be one or two paragraphs long.

- *Research Question*. After hooking the reader in and setting the stage, it is time to state your research question as clearly as possible. I like to do so by stating my actual research question as the first sentence of this part of my introductions. "What is the impact of participation in contract farming on the welfare of those who participate?" The clearer this is stated, the better, because the fewer are the occasions for the reader to be disappointed. This should be one paragraph long.

- *Antecedents*. After stating your research question, it is time to relate it and what you are doing to the existing literature. Here, relate your work to the five to ten closest studies (the closer to five, the better) in the literature. What the relevant literature—the antecedents—is will obviously depend on the question at hand. If you are lucky enough to work in a literature that has seen a lot of activity, you may have a hard time narrowing it down, and you will need to judiciously pick the five to ten closest studies. If you are working on a problem that no one has really looked at, or that no one has looked at in a long time, you might have to go back in time a bit further or expand your parameters for what counts as antecedents. Here, what counts is to tell a bit of a story; no one wants to read a bland enumeration of studies: "Johnson (2011) found this. Wang (2012) found that. Smith (2013) found something else. Patel (2015) found something else altogether." For every topic, the intellectual history of that topic can be told in an interesting way.

- *Value Added*. This is where you need to shine. What is your contribution? How does your paper change people's priors about your topic? Ideally, your paper will have three contributions. For instance, you may be improving on the internal validity front for the question you are looking at by having a better identification strategy. You may also be improving on the external validity front by having data that cover a broader swath of the real world; or you may be performing a mediation analysis that allows identifying what mechanism m the treatment variable D operates through

in causing changes in y. Lastly, you may also be bringing a small methodological improvement to the table. But even papers with one contribution deserve to be published, provided that contribution is important enough.

- *Roadmap*. Lastly, you should provide your reader with a roadmap to your paper. This section usually starts with "The remainder of this article is organized as follows," and it lists section and what they do in order. So for a typical paper, it would go: "The remainder of this paper is organized as follows. Section 2 presents the theoretical framework used to study the research question and derives this paper's core testable prediction. In section 3, the empirical framework is presented, first by discussing the estimation strategy, and then by discussing the identification strategy. Section 4 presents the data and discusses some summary statistics. In section 5, the empirical results are presented and discussed, followed by a battery of robustness checks and a discussion of the limitations of the results. Section 6 concludes with policy recommendations and suggestions for future research." I have seen some economists on social media state that they have gotten papers rejected for many reasons, but never for want of a roadmap section. Fair enough. In most cases, however, it is simply easier to include such a roadmap section and delete it at a reviewer's request than to not have one and have to write one when asked to revise and resubmit a paper, not to mention the fact that some readers will simply expect there to be a roadmap, since the majority of applied economics articles include them. Anything that signals that you know what the unspoken rules and norms of the profession are is a good thing for your article's chances of getting published.

It is best to start writing a paper's introduction as soon as there are some empirical results. After the title and the abstract, the introduction is where most people will decide whether (i) they think your work is interesting enough to keep on reading, and (ii) whether they think your work is of a good enough quality for them to believe your findings. I would guess that the fate of at least 75 percent of articles—whether they get sent

out for review, or whether a revision is solicited by the journal when they do get sent out for review—is driven by the introduction. As such, the introduction should be rewritten every time the file is worked on by any of the authors. I would guess that, for most of my papers, I have gone over the introduction at least a few hundred times.

A good introduction works because it sets your readers' expectations just right. If there is one thing that will make a reviewer recommend a rejection, it is a bait-and-switch (i.e., when an introduction overpromises and the rest of the paper underdelivers), or when an introduction is unclear as to what the paper does, and how it does it.

A busy reader will typically read: (i) your title, (ii) your abstract, (iii) your introduction, then skip to (iv) your tables of results, then read (v) your conclusion, going back to the other sections if and only if they have questions about what you are doing, or how you do it. A good introduction minimizes (or eliminates altogether) a reader's need to flip through the paper in search of answers to her questions.

8.3 Abstract

Having chosen a good title and having written a good introduction, the task of writing your abstract should be relatively easy. Typically, it is possible to write a solid draft of your abstract by keeping only the first sentence of the hook, research question, and value added sections of your introduction, and by polishing up the resulting paragraph some.

A good piece of advice I received from a senior colleague early on in my career was this: Except for the requisite terminology (e.g., randomized controlled trial, difference-in-differences, regression discontinuity), your abstract should be intelligible to any smart, college-educated person who is not an economist. This is especially true for an applied economics paper. After all, we are writing about real-world phenomena that are of interest to policy makers or business managers, so your abstract should be intelligible to someone with a master's degree in public policy or in business administration, depending on what you are doing. Do not make the mistake of confusing lack of intelligibility

with intellectual rigor; this is economics, not French postmodern philosophy.

Ultimately, your goal is not only to get your papers published, but to get them read, and to get them cited. The measure of a scholar's impact in any discipline is her number of citations.[19] If your title is not repellent, and if your abstract is intelligible to people who are not experts in your field and to people in other disciplines, you have just expanded the scope of your citations tenfold, because whether one likes it or not, a lot of people cite stuff they have only read the abstract of.

9 Literature Review and Background Sections

You may have noted that Keith Head's introduction formula includes its own (mini) literature review. Although master's theses or doctoral dissertation chapters should include a separate section reviewing the literature to signal that the student is clearly familiar with the literature she is working in, such a section is almost always entirely unwarranted in an applied economics paper to be submitted to an economics journal.

The reason is simple: Most readers have only very little time on their hands, and most readers will want to get to a paper's contribution sooner rather than later. As a result, a mini literature review discussing how a given paper relates to the five to ten closest studies in the literature is much more effective than a separate section reviewing an entire literature.

Moreover, most people are not good enough writers to pull off writing a literature review section that is worthy of being read, which requires telling a compelling story about the development of an idea. Though most researchers know their topic well enough to be able to identify all or almost all of the relevant related studies, few are able to aggregate the knowledge derived therefrom and coherently write up the intellectual history of the

[19] One is tempted to add "in any discipline *except* economics," as it is only in economics that it is seemingly more important to please five to ten gatekeepers so as to get in the right journals than it is to actually have an impact.

topic at hand. In any case, literature reviews are best written by senior scholars—who are more likely to offer a unique perspective on a topic because they have actively worked on it—and to theses and dissertation chapters. For the majority of applied economics articles, unless a reviewer asks for a separate literature review section, a mini literature review in the introduction is enough.

What about background sections? Those are a different story. When a topic requires a good amount of background knowledge, a separate background section can be very useful. This is especially the case when the details of some legislation need to be kept in mind when assessing the effects of some part or all of that legislation on some outcome of interest. Likewise, in empirical industrial organization studies, it is common for authors to include a background section that describes the industry they are studying. As with anything else in an applied economics article, the background section should tell the reader what she needs to know—no more and no less.

10 Where to Submit?

All of the foregoing was geared toward writing papers that can be submitted (and hopefully published) in peer-reviewed journals. This naturally raises the question of how to choose where to submit an article once that article is "finished."[20]

Over the years, I have heard differing advice about whether one should write for a specific target journal or not. Some people say that you should write with a specific journal in mind; others say you should just write the paper, see how it turns out, and then think about where to submit.

I do not really have an opinion on the matter, except for the following: I strive to write for an (imaginary) audience composed of PhD economists, but an audience of PhD economists who are not familiar with my field. Here, think of your classmates during

[20] I use quotation marks before our work is never finished. When I am unfortunate enough to have to look at my own published papers, I always find something I wish I could have done better.

your first-year core courses, most of whom probably ended up in different fields. As such, I tend to write for a more general reader. I am convinced that even when you end up submitting an article to a field journal, writing for a general audience helps. The editor, for instance, might be in your field, but might not be familiar with your specific topic, so writing for a general audience can help convince her that your work is of general enough interest within your field. Likewise, writing for a general audience might help you attract readers who would otherwise not read your article by making it accessible to them, which ultimately leads to your work being cited more.

Going back to the question of where to submit, here are a few general guidelines, in no particular order.

If you choose to submit to a field journal, make sure that you actually cite a good number of articles published in that journal (or close substitutes, e.g., *Economic Development and Cultural Change* for *Journal of Development Economics*, or *Labour Economics* for *Journal of Labor Economics*) and in that field over the last five years, and more recently if possible. This does two things. First, citing articles published in that journal serves to convince the editor, who has to decide whether to desk reject your paper or send it out for review, that your paper should be sent out for review because it is a good fit with what the journal publishes. Second, citing articles recently published in that journal helps the editor select reviewers for your paper.

If you only cite older articles published in your target journal, odds are the journal has moved on from publishing on that topic (probably because the topic is no longer in fashion), which makes it more likely that the editor will desk reject. If she does choose to send your paper out for review, it might be difficult for her to find the right reviewers, because the people who have published on that topic in her journal are likely to have moved on to other topic and to get cranky about having to review papers on it.

If you do not cite articles in your target journal, even if the editor decides that it is a good fit for that journal, you run the risk of getting reviewers suggested by a keyword

search. For instance, I once had to handle a trade manuscripts which only cited the works of Jagdish Bhagwati, Paul Krugman, Marc Melitz, and so on, without citing any work in the journal I was handling it for (or in any close substitute journal, for that matter). When they are not familiar with a given topic, editors start thinking about reviewers by looking at the references of a paper. Here, the issue is that Bhagwati, Krugman, and Melitz probably do not have time to referee for field journals, especially field journals that are not ostensibly about international trade. So how did I get reviewers? By doing a keyword search (e.g., "international trade") in the editorial system. This returned a few hundred candidate reviewers, and I selected two or three of them. But I am pretty sure none of those reviewers had seen the paper. And therein lies the rub: One of the unfortunate, unstated truths about this profession is that no matter how much we like to think we are doing Very Serious Science, we are not. Network effects sadly matter, and reviewers are more likely to be favorable toward your paper if they have seen it before, preferably in a seminar or at a conference where they got to ask their questions about the work.

When submitting to a field journal, it is thus important to cite things that have been published recently in that or closely related journals. How about general journals? Here, opinions differ. I have heard that, when submitting to a top-five journal,[21] it is best to minimize the number of citations to field journals, because some general-journal editors conclude when they see that an article citing too many articles in field journals that that article also belongs in a field journal. I personally think that this is misguided, but I am also not an editor at a general economics journal, so my opinion is just that.

Given the foregoing, two approaches work reasonably well. The first approach is that you write your paper with a specific target journal in mind, because you know that that journal has recently been publishing articles on your topic. For instance, suppose you are writing a paper on the inverse relationship between farm or plot size and productivity. In

[21] Traditionally, the top five journals in economics have been, in alphabetical order, the *American Economic Review*, *Econometrica*, the *Journal of Political Economy*, the *Quarterly Journal of Economics*, and the *Review of Economic Studies*. Other top general journals include the *American Economic Journal: Applied Economics*, the *Economic Journal*, the *Journal of the European Economic Association*, and the *Review of Economics and Statistics*.

recent years, the *Journal of Development Economics* (JDE) has published a number of articles on the topic, and so that would be a natural outlet for your paper if and only if (i) your methodology clearly improves on that used in those recent articles in JDE on this topic,[22] or (ii) your methodology is as good as the methodology in those articles, but your results go against conventional wisdom on this topic and you can explain why that is the case.

The second approach is to just write the paper without a specific outlet in mind, but still keeping the average economist in mind. Once you are "done" writing your paper, you then look at your list of references. If there are some field journals you cite more than three times each, those are all good candidates regarding where to submit. Once again, if your work improves on both the internal validity and external validity fronts, you should start with a more general economics journal. Know, however, that even the very best papers have a low probability of getting into those journals, as the competition is fierce—and it is getting fiercer.

11 Conclusion

I have sought to write down the unspoken rules and norms that respectively guide and surround the writing of papers in applied economics. Starting with the typical structure of an economics article, I have first discussed each of the constituent sections of that structure. Rather than tackle these sections in the order in which they appear in a working paper or published article, I have instead tackled them in the order in which they are actually written in the process of writing a research paper. Among other things, this means that the discussion of how to pick a title, how to write an introduction, and how to write an abstract as been relegated toward the end of this paper.

Ultimately, my goal is to de-mystify the paper-writing process for applied economists.

[22] By "clearly improves," I mean that it improves from the perspective of internal validity. If you can improve on both the internal validity and the external validity front, you may want to try submitting to a more general economics journal first.

Science-fiction author William Gibson famously said that "The future is already here—it's just not evenly distributed." Likewise with a lot of the information contained in this paper: that information is publicly available; it's just not uniformly distributed. This is especially so given that many graduate programs in economics, applied economics, business, or public policy—programs that purport to train people in the art of writing academic papers—do not have a class or seminar dedicated to that art. It is my belief that, when it comes to technical skills, most PhD programs train their students in ways that are roughly equal (and really, anyone with enough cash can buy her own copy of leading graduate textbooks in microeconomic theory, econometrics, and so on, and go through them). Where the quality of training differs between "good" and "bad" PhD programs is in whether students are taught interstitial knowledge of the kind presented in this paper, and whether they are taught to discriminate between good and bad research questions. My objective in writing this paper has been to play to my comparative advantage and equip anyone who has the time to read this paper with the right tools for the former. For the latter, the reader will have to look elsewhere, as I am only starting to find out now, more than 13 years after finishing my PhD, what constitutes a good research question.

References

[1] Abadie, A., S. Athey, G. Imbens, and J. Wooldridge (2017), "When Should You Adjust Standard Errors for Clustering?," NBER Working Paper.

[2] Acharya, A., M. Blackwell, and M. Sen (2016), "Explaining Causal Findings without Bias: Detecting and Assessing Direct Effects," *American Political Science Review* 110(3): 512-529.

[3] Altonji, J., T.E. Elder, and C.R. Taber (2005), "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy* 113(1): 151-184.

[4] Adler, Mortimer J., and C. Van Doren (2014), *How to Read a Book*, New York: Simon and Schuster.

[5] Angrist, J. D., and J.S. Pischke (2010), "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con Out of Econometrics, *Journal of Economic Perspectives* 24(2): 3-30.

[6] Bellemare, M.F. (2012), "Insecure Land Rights and Share Tenancy: Evidence from Madagascar," *Land Economics* 88(1): 155-180.

[7] Bellemare, M. F., and J.R. Bloem (2018), "Does Contract Farming Improve Welfare? A Review," *World Development* 112: 259-271.

[8] Bera, A.K., A. Ghosh, and Z. Xiao (2013), "A Smooth Test for the Equality of Distributions," *Econometric Theory* 29(2): 419-446.

[9] Burke, M., L. Falcao Bergquist, and Edward Miguel (2019), "Sell Low and Buy High: Arbitrage and Local Price Effects in Kenyan Markets," *Quarterly Journal of Economics* 134(2): 785–842.

[10] Conley, T.G., C.B. Hansen, and P.E. Rossi (2012), "Plausibly Exogenous," *Review of Economics and Statistics* 94(1): 260-272.

[11] Deaton, A. (1997, 2018), *The Analysis of Household Surveys*, Washington, DC: World Bank Group.

[12] Foster, A. D., and M.R. Rosenzweig (1995), "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy* 103(6): 1176-1209.

[13] Gerber, A.S., and D.P. Green (2012), *Field Experiments*, New York: W.W. Norton & Co.

[14] Glennerster, R., and K. Takavarasha (2013), *Running Randomized Evaluations*, Princeton, NJ: Princeton University Press.

[15] Glewwe, P.W., and M. Grosh (2000), *Designing Household Survey Questionnaires for Developing Countries*, Washington, DC: World Bank Group.

[16] Head, K. (2020), "The Introduction Formula," http://blogs.ubc.ca/khead/research/research-advice/formula last accessed January 2, 2020.

[17] Hirshleifer, J. (1956), "On the economics of transfer pricing," *Journal of Business* 29(3), 172-184.

[18] Jacobellis v. Ohio (1964), 378 US 184.

[19] Letchford, A., H.S. Moat, and T. Preis (2015), "The Advantage of Short Paper Titles," *Royal Society Open Science* 2(8): 150266.

[20] Lewbel, A. (2019), "The Identification Zoo: Meanings of Identification in Econometrics," *Journal of Economics Literature* 57(4): 835-903.

[21] McKenzie, D. (2012), "Beyond Baseline and Follow-Up: The Case for More T in Experiments," *Journal of Development Economics* 99(2): 210-221.

[22] Morgan, S.L., and C. Winship (2015), *Counterfactuals and Causal Inference*, Cambridge: Cambridge University Press.

[23] Oster, E. (2019), "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business and Economics Statistics* 37(2).

[24] Pearl, J. (2009), *Causality*, 2nd ed., Cambridge: Cambridge University Press.

[25] Sánchez de la Sierra, R. (2020), "On the Origins of the State: Stationary Bandits and Taxation in Eastern Congo," *Journal of Political Economy* forthcoming.

[26] Solon, G., S.J. Haider, and J.M. Wooldridge (2015), "What Are We Weighting For?," *Journal of Human resources* 50(2): 301-316.

[27] Suri, T. (2011), "Selection and Comparative Advantage in Technology Adoption," *Econometrica* 79(1): 159-209.

[28] Thomson, W. (2011), *A Guide for the Young Economist*, 2nd. ed., Cambridge, MA: MIT Press.