2

Writing Papers

A good film makes you forget that you are watching a film. Similarly, a good research paper makes you forget that you are reading a research paper. The authors tell you a story. They 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 for most graduate field courses 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 most important.

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 reviewed more papers—but reading only the best papers is a double-edged sword. To be sure, those are the papers we learn the most from 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. But it is 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 those films on the British Film Institute's list of the 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 top-50 films any 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 in his classic *How to Read a Book* as inspectional reading (Adler and Van Doren 2014). 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 item on one's reading list. Reading papers that way is a good way to develop one's knowledge of the literature on a given topic, but it is hardly a recipe for learning how to write good papers.[1]

The goal of this chapter is thus to help readers write applied papers for eventual submission and publication in peer-reviewed journals. 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.

2.1 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 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 chapter 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 adopting a minimum-wage policy on unemployment) might not require a Theoretical Framework section 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 avant-garde albums like *A Love Supreme* (1965), he first learned how to operate within the (much more regimented) 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 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 economist should make sure to have written enough papers 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 chapter will not follow the structure just given. Though it would certainly be easier for me to write a chapter whose subsequent headings 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.

2.2 Theoretical Framework

Since the goal of empirical 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 a theoretical hypothesis about the relationship between *x* and *y*.

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 empirical relationship of interest a given article will focus on. In other words, the best empirical articles 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.[2]

The question, then, is how to take your hypothesis and convert it into a proper theoretical framework for an empirical economics article. Here, there are two possible scenarios: (i) you are investigating a question which has already been studied by theorists, or (ii) you are investigating a question which 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 not 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 fine to use someone else's theoretical framework—provided that you clearly state that you are doing so and cite the source of your theoretical framework.

A closely related option is to adapt somebody else's theoretical framework to suit your needs—say, by incorporating an additional variable, or by relaxing or 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 the theoretical reasons behind your hypothesis. In some cases, this may require a formal theoretical model. In other cases, it is enough to merely present a verbal conceptual framework.[3] 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.

Generally when writing papers that combine theory and empirics, you should make sure that your empirics actually test the testable predictions of your theory. A prediction that is not tested empirically should not be included in your theoretical framework, and your hypothesis tests should be grounded in your theoretical framework. In other words, avoid any disconnect between your theory and your empirics.

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 an 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. In that sense, having a theoretical framework in your paper often only serves as a signal 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.

2.3 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, and Gerber and Green 2012 or Glennerster and Takavarasha 2013 for randomized controlled trials), so this section will not

discuss where the data come from, and assume that you already have them. Rather, this section will focus on how to present your data in the context of an economics article.

The best Data and Descriptive Statistics sections answer all of the reader's questions about the data. 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, firms, 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 if the data are longitudinal, and 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, people often derive their income from many difference sources. So if an "income" variable 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, imagine what it is like to the reader—but it can nevertheless contain crucial information.

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 2.1 shows one such table. This allows presentation of 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.

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 non-agricultural 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 = 0.01 hectare = 100 square meters.)
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 2.1
Example table of variable descriptions from Bellemare (2012).

At this point, 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, namely tables where each line is a variable retained for analysis, where means 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 only 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 5 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 pairwise comparisons return too many systematic differences, one should ideally control for the relevant covariates in a regression or matching context when estimating treatment effects.[4]

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 binary.

When writing a Data and Descriptive Statistics section, there are a few mistakes you should avoid making. The first is for the writeup to present a boring 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 your data or results, and just as you should avoid the passive voice, you should also avoid the past tense in research papers, except when summarizing and concluding. Indeed, the past tense should be largely kept for when you discuss what other researchers have done before you, and the future tense for what you are planning on doing or what others should be doing in the future. The present tense is ideal because it refers to that which occupies the reader right now, which is your paper.[5]

Finally, 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 difficult to interpret in tables, such as $1.37e + 8$, or anything other than units readers are used to dealing 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 logarithm of income. Ultimately, although a lot of what goes into 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 estimates of interest, and to get an idea of how those estimates are likely to vary across a given conditioning domain.

2.4 Empirical Framework

After discussing the data and presenting descriptive statistics, you normally turn to discussing your empirical framework, that is, 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).

2.4.1 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 therein, 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), and the error term (labeled).

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, and so forth, 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 subsection features several different specifications of the same equation, coefficients should also have subscripts. In other words, one should not reuse estimand notation. If is used to denote the coefficient of interest in a regression of y on D, it should not be reused 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 $_0$ and in the latter, $_1$. Or it can be done by adding letter subscripts to each coefficient, so that for example $_r$ and $_s$ can respectively refer to reduced-form and structural estimates of the same coefficient.
- The estimation strategy subsection 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

$$y = + D + x + , \textbf{(2.1)}$$

for instance, the relevant hypothesis test would be of the form $H_0 : = 0$ versus $H_A : \ 0$. 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 subsection 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).

2.4.2 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? Briefly, 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 from cause D to effect y.

Although an unbiased coefficient estimate implies an identified—that is, causally identified—coefficient estimate, the converse is not true. 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.

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 kept short, as your results are causally identified by virtue of experimental assignment. In other words, you can estimate what Pearl (2009) denotes $E(y|do(x))$; that is, 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. This chapter cannot dive into causal identification with observational data, but there are nevertheless certain things that can be discussed as being necessary 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.[6]
- Discuss in turn the three following sources of statistical endogeneity:[7] (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, and 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. Be honest about what your paper can and cannot do.
- Once that is done, there is one more threat to internal validity to be considered, namely 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 important: if your results are not causally identified, *be honest* 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 how much you wish those claims to be true. Editors and reviewers would much rather deal with manuscripts wherein the author candidly admits to the limitations of their findings than with manuscripts wherein the author tries to deceive the reader. In plain English: the former kind of manuscript has a much better chance of not being rejected than the latter.

2.5 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 any lingering questions, 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.

2.5.1 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 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).[8]

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).

2.5.2 Robustness Checks

After presenting the core results in a paper, it is time to turn to robustness checks. There was a time when 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), which has led to greater emphasis being placed on causal identification and inference,

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 cherry-picked, which means that establishing the robustness of a finding involves its own set of complications.

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, and so forth. If you have access to all seven of those measures of "welfare," one first step toward establishing that your result is robust might be simply to re-estimate your core equation for each of those measures, showing that the result holds across all of them.

Similarly, you may have different measures of the treatment variable. In most randomized controlled trials (RCTs), there is one (and only one) treatment variable (unless there are several treatment arms, and unless those treatment arms 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 one of those specifications, that goes a long way toward establishing that the 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. In difference-in-differences studies—a methodology where the frontier has been evolving rapidly over the past few years—one should analyze trends.

Yet another kind of robustness check comes in the form of looking at different estimators. Most empirical 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.

2.5.3 Treatment Heterogeneity

It is rare 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 subgroups (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: 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. 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.

When I write that exploring treatment heterogeneity can salvage a null finding, you should not conclude from that statement that when you have a null finding, you should explore treatment heterogeneity. If you wish to explore treatment heterogeneity, you need to have a good reason (usually stemming from your theoretical framework) for doing so. Anything else will reek of *p*-hacking (i.e., the phenomenon whereby researchers slice their data until they find something significant to report in their paper), which leads to plain bad science.

2.5.4 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.

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.[9] 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*?"

2.5.5 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 subsection 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.

Second, external validity may be limited as well. This is often the case with lab or a lab-in-the-field experiments,[10] or with RCTs. 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.

Finally, 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.

2.5.6 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:

- 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 Years of Education on Wage by Gender.” The titles should thus tell us what is being estimated (e.g., OLS), what the relationship of interest is (i.e., the effect of participation in education on wage), 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.[11]
- 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 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; for completeness and transparency, none should be omitted), and additional symbols if necessary.[12] 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 an 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, 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 left with more questions than one has answered after looking at the tables?

2.6 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 down 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, "tell 'em what you're going to tell 'em, tell 'em what you want to tell 'em, and tell 'em what you just told 'em." 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—a story about the paper's contribution, and the gap it fills.
- *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.
- *Real-World Implications*. Presumably, your work has some sort of implication for policy, business strategy, or something else 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. 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 probably where the phrase "back-of-the-envelope calculation" most often comes up in 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 what your results suggest. Also discuss how easy or hard it would be to implement.
- *Implications for Future Research*. No work is 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 an 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.

2.7 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 empirical papers, wherein authors rarely advance the frontier of knowledge theoretically or methodologically. Indeed, 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 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. Those same readers are unlikely to have that kind of patience for empirical papers, which are about the sign and magnitude of an empirical relationship, and how the authors estimated that relationship. Consequently, the following subsections focus on these components of an applied economics paper.

2.7.1 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* (*PNAS*) often insist that we should state our results in our titles. While that may be true for the papers we submit to those general-science journals, titles conform to 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 empirical economics paper—that is, a paper that asks an empirical question of the form "What is the effect of *D* on *y*?"—it is safe to go with a title of the form "The Impacts of *D* on *y*: Evidence from [the Context You Are Studying]." A variant on this theme is a title of the form "*D* and *y*," with or without the subtitle after the semicolon.

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, adages, dicta, proverbs, short biblical passages, or titles of famous films, books, or TV shows. Ultimately, if you are going to take the clever route, make sure the cleverness is warranted, and that the clever part of your title perfectly fits your paper.

2.7.2 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 them 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 . . ." Then why should anyone put up with your attempt to make that literature 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. To carry on with the example in the previous bullet point: "What is the impact of participation in contract farming on the welfare of those who participate?" The more clearly this is stated, the better, because 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. Kim (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*. Finally, you may also be bringing a small methodological improvement to the table. This is not necessary, as even papers with fewer than three contributions deserve to be published, provided at least one of their contributions is important enough.
- *Roadmap*. Finally, 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 sections 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 had 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.

As mentioned earlier in this chapter, 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.

In addition to Keith Head, Claudia Sahm and David Evans both provide advice on how to structure the introduction of an economics article. In her advice to job-market candidates, Sahm (2019) suggests the following order of sections for papers in macroeconomics:

1. Motivation
2. Research question
3. Main contribution
4. Method
5. Findings
6. Robustness checks
7. Roadmap

Though her suggested structure is close to Keith Head's, it is worth noting that Sahm suggests a more precise content for Head's "value added" section, namely, you should tell your reader what empirical methodology you use to answer your research question as well as how you make sure your results are robust.

Similarly, after analyzing the introduction of "the most recent empirical microeconomic development papers from a range of top journals," Evans (2020) suggests the following structure, as well as a suggested length for each "section" of the introduction:

1. Motivations (1–2 paragraphs)
2. Research question (1 paragraph)
3. Empirical approach (1 paragraph)
4. Results (3–4 paragraphs)
5. Value added (1–3 paragraphs)
6. Robustness checks, policy relevance, limitations (optional)
7. Roadmap (1 paragraph)

Although Evans' advice stems from analyzing empirical development articles, it broadly overlaps with the advice given by both Keith Head and Claudia Sahm, and there is a payoff to adopting a common structure for your introduction: McCannon (2019) analyzed the papers published in the *American Economic Review* from 2000 to 2009 by looking at their readability score, and found that the papers that were hardest to read suffered a statistically significant decrease in their citation count of 0.20 standard deviation.

2.7.3 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 that except for the requisite terminology (e.g., RCT) 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 empirical paper in economics. 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. In other words, do not make the mistake of confusing lack of intelligibility with intellectual rigor; this is economics, not French postmodern philosophy (Sokal and Bricmont, 1999).

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.[13] 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 considerably expanded the scope of your citations—for better or for worse, a lot of people cite a lot of articles they have only read the abstract of.

2.8 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 a 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 or method. 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 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 thought about it for a long time—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 economics article, the background section should tell the reader what she needs to know—no more and no less.

2.9 Writing for the Right Journal

All of the foregoing is geared toward writing papers that can be submitted (and hopefully published) in peer-reviewed journals. Though I will have more to say in chapter 4 about where to submit on navigating the peer-review process, this naturally raises the question of whether you should write with a specific journal in mind.

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 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,[14] 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 often.

If you want to write for a specific journal, however, here are a few general guidelines, in no particular order.

If you plan on submitting 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 the *Journal of Development Economics*, or *Labour Economics* for the *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 likely 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 of interest to readers), 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 topics 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 manuscript 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 before. And therein lies the rub: one of the unfortunate, unstated truths about this profession is that 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 had an opportunity to ask their questions about the work.

If you plan on submitting to a field journal, it is thus important to cite articles that have been published recently in that or closely related journals. How about general journals? Here, opinions differ. When submitting to a top-five journal,[15] 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.

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.

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, 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.

2.10 The Act of Writing

A chapter on writing papers would be incomplete without discussing the act of writing itself.

I cannot claim to be a good writer, but I have managed to become a competent one over the years. To carry my earlier analogies further, no one can become a good film director who has not watched a lot of films, good and bad, and no jazz musician can become proficient at improvising over chord changes who has not listened to a lot of other musicians doing it. It is the same with writing, and the best way to become a competent writer is to spend time reading. Here, almost anything will do—just find something you enjoy reading, and which has been professionally produced. English is my second language, so when I made the conscious decision to improve my written English during my first year of college, I settled on taking the *The Economist* and the novels of Robertson Davies with me on my daily subway ride to campus. Reading regularly will provide a solid foundation upon which to build your own writing skills.

Beyond that, there are a number of excellent resources on how to improve your writing. Among the ones I have benefited from are Strunk and White's classic *The Elements of Style*, Zinsser's *On Writing Well*, best-selling author Stephen King's *On Writing*, and Phillips' *Ernest Hemingway on Writing*.

Generally, however, the following rules of thumb can help you become a better writer:

1. *Briefly embrace mediocrity*. When trying to write anything, anyone but the most self-delusional of narcissists will typically hear a voice in their head immediately criticizing anything they write once they write it. In a now-famous section of her 1995 book *Bird by Bird* titled "Shitty First Drafts," Anne Lamott outlined a very useful strategy to become more productive as a writer, which is as follows: When we write the first draft of anything, we should do so fully expecting that nothing good will come out of it, and knowing that the first draft will be, well, shit. But once you have a first draft, you can improve upon it—no one will ever see how bad your first draft was, and you can just keep polishing it until you have a product you like. The same cannot be said of a draft that never gets written and remains an idea in your mind because you could not briefly embrace mediocrity.
2. *Writing is rewriting*. Speaking of polishing, as much as people like to believe urban legends about this or that famous writer who wrote their *magnum opus* in one fell swoop,[16] the bulk of writing anything consists in rewriting it to make it better. As I have mentioned when discussing above how to write introductions, this act of rewriting is what ultimately can take a mediocre first draft and make it good.
3. *Write every day*. Though this seems like a tall order—who actually has time every day to dedicate to writing?—it really is not when you think about it. Unless you make a point of not responding to email (say, because you are on vacation), every day you spend in this profession will bring an occasion for you to make the decision to deliberately write well in some form, even if that form is responding to email. Though some productivity tips and other "life hacks" encourage you not to worry about proper capitalization and grammar when writing email (presumably in an effort to add a whole five minutes of productivity to your day), I would encourage you to see everything you write as an occasion to write clearly and concisely—in other words, to write competently. If you want more occasions to write, the habit of keeping a daily journal builds writing time into your day.

2.11 Writing Papers for General-Science Journals

Increasingly, economists interested in improving the visibility of their work submit to and publish in general-science journals such as *Nature*, *PNAS*, and *Science*. Writing for those journals, however, is dramatically different than writing articles for economics journals. I thus briefly discuss how to do so to close this chapter.

General-science journals care much more about your findings and their implications than they do about methods. Methods are important, but they are relegated to the end of the paper (or to supplemental materials), for specialists who may be interested in seeing them.

Structure-wise, a paper in a general-science journal has four sections: (i) introduction and motivations, (ii) results, (iii) discussion, and (iv) methods. In the results section, you should only report results. The interpretation of those results should be kept for the discussion section, where they can be put in context. You can even afford to speculate, since unlike in an economics journal, not everything you state needs to be backed up by a number of robustness checks. But even in general-science journals, your core results need to be robust, and speculation should only be relied on to explain the implications of these results.

Understand: it is very challenging to write for general-science journals, as you will need to appeal to researchers beyond economics. As a colleague with experience publishing in both top general-economics and general-science put it: "Publishing in *PNAS* requires just as much effort as publishing in a general interest economics journal. Don't wing it."

Notes

1. 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 of a reader's time is high, so the average reader is more likely to give up on reading a paper than to hunt 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).
2. This is not to say that there are no good papers looking 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.
3. 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. 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.
5. In his classic *On Writing Well*, Zinsser (2006) writes: "[T]he whole purpose of tenses is to enable a writer to deal with time in its various gradations, from the past to the hypothetical future." My editor at MIT Press also tells me that analysis of literature and film also uses the present tense when describing a book or a film.
6. This presumes that your research design *is* the best thing out there. In cases where your research design 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.
7. I talk explicitly of statistical endogeneity—what makes $Cov(D, |x)$ 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 $Cov(D, |x) = 0$ 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 empirical economist who says his results are causally identified because his treatment variable is theoretically exogenous.
8. 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 (analysis of covariance) setup (McKenzie 2012).
9. This section's placement in this chapter in no way indicates that you should think about mechanisms *ex post*. It merely indicates that the section in which you empirically test for mechanisms should come after your main results. Ideally, the presumed

mechanisms whereby your treatment variable affects your outcome of interest should be discussed as part of your theoretical framework.

10. Lab-in-the-field experiments are lab experiments that are conducted with "real" subjects (e.g., firm managers) in the field, outside of the experimental lab.

11. 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).

12. That being said, at the time of writing, the AEA journals have moved away from relying on asterisks (i.e., "stars") or other typographical symbols to denote statistical significance. It remains to be seen, however, how many other journals will follow suit. Here, the best thing to do is to read a journal's formatting guidelines before submitting.

13. 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.

14. Field journals are defined in relation to general journals. Whereas the latter might be open to publishing articles on most if not all areas of economics, the former publish articles in a given field (e.g., agricultural economics, economic history, monetary economics).

15. 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*.

16. The most famous of such urban legends is probably that surrounding Jack Kerouac's *On the Road*. From a story on National Public Radio (Shea 2007): "Legend has it that Kerouac wrote *On the Road* in three weeks, typing it almost nonstop on a 120-foot roll of paper. The truth is that the book actually had a much longer, bumpier journey from inspiration to publication, complete with multiple rewrites, repeated rejections and a dog who—well, *On the Road* wasn't homework, but we all know what dogs do."

3

Giving Talks

After writing papers, presenting them—giving talks—is the second most common form of communication in which economists engage. Even an economist who has no intention of ever joining academia will almost surely have to prepare at least one research presentation in her career—her job talk, or a presentation of her research for an audience of people who are considering hiring her as a colleague.

Like writing papers, how to give talks is something the average economist is expected to learn on her own, learning both from others and by doing, with little to no formal guidance except the odd blog post. This chapter provides guidelines for giving talks, from your first presentation at your department's graduate-student brown-bag seminar to full-fledged invited seminars, and from technical disciplinary research talks to outreach talks given to people who know little or nothing about economics.

Presenting a paper involves the distillation of the results in that paper, so much of the advice given in the previous chapter for writing papers (e.g., structure, content) naturally extends to giving talks. One difference between presenting and writing—indeed, one advantage of giving talks over writing papers—is that giving a talk is an occasion to engage in a dialogue with would-be readers, and maybe even with potential editors and reviewers.[1]

3.1 Invited Seminars

I use invited seminars as the benchmark against which I will compare other types of talks, because an invited seminar is the ideal format for a talk in that it allows you to present a paper fully, or nearly so. Moreover, it is a format most economists will be familiar with from presenting at departmental brown-bag seminars.

An invited seminar lasts anywhere from 60 to 90 minutes,[2] and the presenter is expected not only to present a paper but also to answer the audience's questions about that paper. As such, it is the ideal setting to communicate the contents of a paper from start to finish.

Thus, the first step involved in successfully giving invited seminars is to know precisely what the norms are in the department that has invited you to present your work. Though this obviously includes asking how much time you will have to present, it also involves asking what the ground rules are regarding questions. It used to be a distinguishing trait of economics seminars that questions would be asked and answers given in a free-for-all format,[3] but that is no longer the case as some seminar series have moved to a format where no questions are allowed, say, for the first 15 minutes of presentation to allow the speaker to set the stage for her work, or to a format where only clarification questions are allowed throughout the talk, with more substantial questions being relegated to a formal question-and-answer session at the end.

Generally speaking, an updated version of the one-slide-per-minute rule of thumb applies to giving talks. Though a 75-minute seminar will rarely need as many as 75 slides, a 15-minute conference presentation should have no more than 15 slides. Thus, the updated version of the rule is something like "You should only have as many slides as you have minutes to present, and the shorter the talk, the more likely this constraint is to be binding."

Although the advice given to people making their first deck of slides is invariably of the "less is more" variety when it comes to making each individual slide, economists tend to be more comfortable with more text on slides as well as with fewer images than, say, business executives. It is thus perfectly acceptable for slides to have three to four bullet points, each containing a full sentence. This serves a dual purpose: First,