EMPIRICAL METHODS IN LAW

Robert M. Lawless
Jennifer K. Robbennolt
Thomas S. Ulen

Outline
January 22, 2009

Introduction (1st draft reviewed)

Part I—Why Gather, Organize, and Evaluate Data: An Overview of Empirical Methods

- Introduction to Part I
- Chapter 1—Thinking Empirically (1st draft reviewed)
- Chapter 2—Research Design (1st draft reviewed)

Part II—How to Gather Data: Empirical Research Methodologies

- Introduction to Part II (1st draft reviewed)
- Chapter 3—Asking Questions: Surveys and Interviews (1st draft reviewed)
- Chapter 4—Experiments (JKR is drafting)
- Chapter 5—Archival Methods (JKR is drafting)
- Chapter 6—Sampling (1st draft reviewed)
- Chapter 7—Coding (RML is drafting)

Part III—How to Evaluate the Data: Statistical Techniques

- Introduction to Part III (1st draft reviewed)
- Chapter 8—Distributions and How to Describe Them (RML is drafting)
- Chapter 9—Hypothesis Testing (RML is drafting)
- Chapter 10—Inferential Statistics (1st draft reviewed)
- Chapter 11—An Introduction to Correlation and Regression Analysis (TSU is drafting)
- Chapter 12—Advanced Regression Techniques (TSU is drafting)

Part IV—Communicating Your Results

- Chapter 13—Reading, Presenting, and Writing About Empirical Matters
- Chapter 14—Conclusions

Annotated Outline

Introduction

(1st draft reviewed)

In the introduction we will explain our view of the audiences to whom we are addressing this book, why we are doing this book, and what we hope that students will get out of the book. Importantly, we are directing this material at making law students, professors, judges, and practitioners more sophisticated consumers of empirical materials in law, both legal scholarship and material presented as part of a legal case or controversy. It is our hope that readers will find themselves more comfortable deposing an expert witness about her survey techniques offered to support a claim of trademark infringement or that they will read an article on the causes and consequences of executive compensation with a more critical eye for the strengths and weaknesses of the empirical techniques and data offered in support of the author's viewpoint. If, in addition to becoming more sophisticated consumers, readers also become adept producers of empirical materials themselves, so much the better. We also hope to make clear our view that having a background in empirical study as an undergraduate or employee before coming to law school is not necessary, that a prior background in mathematics or statistics is not necessary, and that it is our intent to make this material accessible to anyone and to make this book a comprehensive introduction to the topic. We will use an example to contrast traditional legal analysis from the empirical approach which will be our focus.

The introduction will also introduce the plan of the book. Part I introduces discussion of the role and importance of empirical research in the law (Chapter 1) and sets the conceptual framework for the discussion by introducing a number of underlying concepts upon which an empirical approach is based, regardless of the specifics of its form (Chapter 2). Part II has two goals. First, it describes several general methodological approaches for addressing research hypotheses and collecting data such—surveys (Chapter 3), experiments (Chapter 4), and archival data (Chapter 5). Second, it address two issues that cut across all methods of data gathering—sampling (Chapter 6) and the coding of data (Chapter 7). Part III covers the major statistical techniques that are common to empirical research of these varying types, including principles of computing and presenting descriptive statistics (Chapter 8), the basics of hypothesis testing (Chapter 9), basic inferential statistics and the concept of statistical significance (Chapter 10), and regression analysis (Chapters 11 and 12). Part IV addresses issues involved in writing about empirical matters and the presentation of empirical data (Chapter 13) and concludes with general lessons and further topics that the interested reader might explore on her own (Chapter 14).

PART I—WHY GATHER, ORGANIZE, AND EVALUATE DATA: AN OVERVIEW OF EMPIRICAL METHODS

Introduction to Part I

Chapter 1—Thinking Empirically

(1st draft reviewed)

In Chapter 1, we will provide a general discussion of the role of empirical work in law. We will explain what we mean when we talk about "empirical" research and provide some discussion of the relationship between theory and empirical testing. Most importantly, we intend to demonstrate the utility of empirical research in law by providing readers with some compelling examples of legal issues (e.g., capital punishment, crime, eyewitness testimony and DNA exonerations) that have had clear light thrown on them by the application of empirical techniques.

Chapter 2—Research Design

(1st draft reviewed)

The purpose of Chapter 2 is to provide a vocabulary and a framework for the rest of the discussion by introducing some very general propositions about empirical research, such as how to frame research hypotheses, triangulation, reliability, internal and external validity, and so on. We will describe the variety of choice one has when one is designing an empirical study. While the randomized field experiment is the gold standard for establishing causal relationships, such experimental are not always feasible. Thus, simulations, studies using comparison groups, and correlational approaches are often used. We will introduce the notion that what method one chooses (for example, survey, experiment, or archival data) will depend in large part on the research question posed and on how one balances the pros and cons of the methods (discussed in detail in the following chapters) given the purpose of the research. We will include a box introducing and describing the role of Institutional Review Boards (IRBs) in human subjects research.

PART II—HOW TO GATHER DATA: EMPIRICAL RESEARCH METHODOLOGIES

Introduction to Part II

(1st draft reviewed)

Chapter 3—Surveys

(1st draft reviewed)

Chapter 3 will discuss principles of survey design and administration. Topics will include the nitty-gritty of question (and answer) drafting and the different methods of administration (such as phone surveys, mail surveys, face-to-face surveys or interviews, and Internet-based surveys). The pros and cons of each of these mechanisms for administration will be considered as well as the pros and cons of this method more generally.

Possible examples: trademark surveys, surveys in cases involving issues of pre-trial publicity and motions to change venue, recurring large national surveys such as the General Social Survey (GSS) and the Current Population Survey (CPS).

Chapter 4—Experiments

Chapter 4 will discuss the basic principles of randomized experiments—the gold standard of research design. The purpose of much empirical research is to demonstrate a causal relationship between variables, and this chapter will demonstrate how experiments seek to establish such relationships using random assignment to isolate the particular construct of interest. We will pay particularly close attention to experiments conducted in field settings (such as courts) and to simulation experiments. We will also briefly discuss the virtues and drawbacks of "natural experiments." And we will conclude by discussing the pros and cons of these various approaches to experimentation.

Possible examples: jury reform experiments in the field, any of a number of simulation experiments (e.g., eyewitness identifications and line-up procedures).

Chapter 5—Archival Methods

Chapter 5 will discuss issues related to obtaining and working with archival data. We will first describe a variety of different sources (for instance, court records). Material will include discussion of existing data sets and methods of working with and recording data from original sources. In addition, we will discuss the sometimes vexing issues related to gaining access to and the accuracy of archival data sources.

Possible examples: PACER, court files, contracts.

Chapter 6—Sampling

(1st draft reviewed)

Chapter 6 will discuss principles of sampling units for study from a larger population. These units might be cases, people (litigants, voters, consumers, judges, etc.), documents (wills, contracts, etc.), or something else. We will have already discussed the principles that are important when using any of the methods already discussed (surveys, experiments, and archival methods). We will describe and compare a variety of sampling methods, including random samples, stratified samples, convenience samples, snowball samples, and so on.

Possible examples: U.S. Census, electoral polls.

Chapter 7—Coding

Chapter 7 will discuss issues related to coding, the systematic recordation of data that has been gathered. We will present and evaluate principles of coding (and codebooks) of data. Although coding may seem obvious, there are a number of important principles that will minimize errors. We will discuss these techniques and discuss how to construct a database that will work for the investigator. We also will discuss proofreading and cleaning of data to ensure accuracy. Statistical software and other software that might aid in empirical studies will be given attention.

PART III—HOW TO EVALUATE THE DATA: STATISTICAL TECHNIQUES

Chapter 8—Distributions and How to Describe Them

The purpose of Chapter 8 is to introduce basic descriptive statistics—statistics that are used to describe the characteristics of a set of data. These include measures of central tendency (e.g., mean, median, and mode) and the differences among them, and measures of dispersion (e.g., range and standard deviation). Other possible topics include the lognormal distribution (which is useful for multiplicative processes and those that have left-skewed distributions), extreme data and the generalized extreme distribution, and zero-inflated damages. We will describe the material and concepts in this chapter using examples; boxes will be included that contain the mathematical details.

Chapter 9—Hypothesis Testing

In Chapter 9, we will cover the basics of hypothesis testing. We will provide background on hypothesis testing as preliminary to our discussion on inferential statistics. In this chapter we will be concerned with discussing at a conceptual level how to determine whether a difference is reliable and when it is due to chance fluctuations. We will discuss the null hypothesis and its comparison to the research hypothesis as well as the concept of statistical significance. In addition, we will address the topics of effect sizes and power analysis.

Chapter 10—Inferential Statistics

(1st draft reviewed)

In Chapter 10, we will discuss principles of inferential statistics—statistical techniques that are used to make inferences about a population from a sample. We will discuss basic inferential statistical techniques (such as the χ^2 and *t*-statistics), the assumptions underlying these techniques, and the circumstances under which each is appropriate to use. We will describe and critique the concept of statistical significance. Throughout the chapter we will use examples as frequently as possible, in the hope of making this material quickly intelligible and not intimidating. Where there are mathematical details, we will strive to put those in boxes as supplemental or advanced material.

Chapter 11—An Introduction to Correlation and Regression Analysis

Chapter 11 will cover the principles of regression analysis. This chapter (in contrast to the following one) will focus on the most basic principles of regression analysis, such as explaining how and why this technique works and what particular problems one needs to be wary of (such as multicollinearity and the problem of omitted variables). We will discuss the differences among time-series, cross-section, and panel data, with examples of each in a legal setting.

We hope to provide some very simple examples of data sets and to explain—in a clear, step-by-step fashion—how to take that data and perform simple and multiple regressions. The examples may involve the use of actual, publicly available data sets or imaginary data designed to illustrate a relatively simple but legal issue (such as, for example, how the salaries of law firm associates and partners might be determined). Included in the demonstration will be instructions on how to use a widely available spreadsheet program like *Microsoft Excel* to perform regressions and, also, how to use a more powerful program, such as *Stata*, which is devoted to regression analysis.

Possible examples: see the discussion of an example (the death penalty) in Chapter 1.

Chapter 12—Advanced Regression Techniques

Chapter 12, which may be one that some classes will skip but which will serve as a useful reference for later, will develop some additional, advanced topics in regression analysis. We will cover such topics as logit, probit, and Tobit models, difference-in-differences techniques, the "bootstrap regression" method, fixed effects, and instrumental variables. The emphasis will be when the research question dictates the use of these more advanced techniques and how to use them. We will illustrate these advanced techniques with actual data and real legal issues and also show how computer software packages can help to run these more advanced regressions.

PART IV—COMMUNICATING YOUR RESULTS

Chapter 13—Reading, Presenting, and Writing About Empirical Matters

The topic of Chapter 13 will be how to write effectively about empirical results. Most practitioners and students of the law will have had little or no experience in presenting empirical results. There are better and worse ways to do so, ranging from issues about what form of chart (never a pie chart!) or other visual representation of the data to use to the formatting of tables of data and the results of regression analysis. We will use illustrations from the literature of inapt and particularly good methods of presenting empirical analysis.

Chapter 14—Conclusions

This chapter will draw the lessons of the previous 13 chapters and give some guidelines to students on how they might further advance their understanding of empirical techniques.